

(19)



Евразийское  
патентное  
ведомство

(21) 202191033 (13) A1

## (12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки  
2021.12.21(51) Int. Cl. C12N 9/22 (2006.01)  
C12N 9/78 (2006.01)  
C12N 15/10 (2006.01)  
C12N 15/113 (2010.01)(22) Дата подачи заявки  
2019.10.15

## (54) ПРОГРАММИРУЕМОЕ РЕДАКТИРОВАНИЕ ОСНОВАНИЙ ДНК ПОСРЕДСТВОМ СЛИТЫХ БЕЛКОВ NME2CAS9-ДЕЗАМИНАЗА

(31) 62/745,666

(72) Изобретатель:

(32) 2018.10.15

Зонтаймер Эрик Дж., Гао Синь,

(33) US

Мир Аамир, Эдраки Алиреза, Вульф

(86) PCT/US2019/056341

Скот А., Лю Пэнпэн (US)

(87) WO 2020/081568 2020.04.23

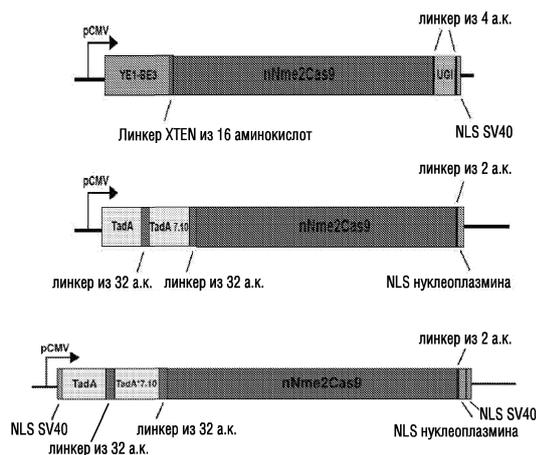
(74) Представитель:

(71) Заявитель:

Медведев В.Н. (RU)

ЮНИВЕРСИТИ ОФ МАССАЧУСЕТС  
(US)

(57) Настоящее изобретение относится к области редактирования генов. В частности, редактирование генов направлено на редактирование единичных нуклеотидных оснований. Например, такое редактирование единичных нуклеотидных оснований приводит к преобразованию пары оснований С-Г в пару оснований Т-А. Высокая точность и аккуратность описанного в настоящем описании однонуклеотидного редактора достигается посредством нуклеазы NmeCas9, которая слита с белком нуклеотиддезаминазой. Компактная природа NmeCas9, сопряженной с более высоким количеством совместимых прилегающих к протоспейсеру мотивов, обеспечивает наличие у слитых конструкций Cas9, описанных в настоящем описании, окна редактирования генов, которое может редактировать участки, на которые не нацеливаются другие общепринятые редактирующие основания платформы SpyCas9.



A1

202191033

202191033

A1

**ПРОГРАММИРУЕМОЕ РЕДАКТИРОВАНИЕ ОСНОВАНИЙ ДНК  
ПОСРЕДСТВОМ СЛИТЫХ БЕЛКОВ NME2CAS9-ДЕЗАМИНАЗА  
ПЕРЕКРЕСТНАЯ ССЫЛКА НА РОДСТВЕННУЮ ЗАЯВКУ**

Настоящей заявке испрашивается приоритет по предварительной патентной заявке США № 62/745666, поданной 15 октября 2018 года, включенной в настоящее описание в качестве ссылки в полном объеме.

**Область изобретения**

Настоящее изобретение относится к области редактирования генов. В частности, редактирование генов направлено на редактирование единичных нуклеотидных оснований. Например, такое редактирование единичных нуклеотидных оснований приводит к преобразованию пары оснований С-Г в пару оснований Т-А. Высокая точность и аккуратность описанного в настоящем описании однонуклеотидного редактора достигается посредством нуклеазы NmeCas9, которая слита с белком нуклеотиддезаминазой. Компактная природа NmeCas9, сопряженной с более высоким количеством совместимых прилегающих к протоспейсеру мотивов, обеспечивает наличие у слитых конструкций Cas9, описанных в настоящем описании, окна редактирования генов, которое может редактировать участки, на которые не нацеливаются другие общепринятые редактирующие основания платформы SpyCas9.

**Уровень техники**

Многие заболевания человека возникают вследствие мутации одного основания. Возможность корректировать такие генетические aberrации является первостепенной при лечении этих генетических нарушений. Короткие палиндромные повторы, регулярно расположенные кластерами (CRISPR), вместе с CRISPR-ассоциированными (Cas) белками составляют РНК-направляемую адаптивную иммунную систему у архей и бактерий. Эти системы обеспечивают иммунитет посредством нацеливания на и инактивации нуклеиновых кислот, которые происходят из чужеродных генетических элементов.

Платформы редактирования оснований SpyCas9 не могут быть использованы для нацеливания на все мутации единичных оснований вследствие их ограниченных окон редактирования. Окно редактирования ограничивается частично необходимостью в PAM NGG и необходимостью, чтобы редактируемое основание(я) находилось на строго определенном расстоянии от PAM. SpyCas9 также по своей природе ассоциированы с частыми нецелевыми эффектами при редактировании генома.

В данной области необходима высокоточная платформа редактирования единичных оснований Cas9, имеющая программируемую специфичность к мишени вследствие распознавания разнообразного набора участков PAM.

**Сущность изобретения**

Настоящее изобретение относится к области редактирования генов. В частности, редактирование генов направлено на редактирование единичных нуклеотидных оснований. Например, такое редактирование единичных нуклеотидных оснований

приводит к преобразованию пары оснований С-Г в пару оснований Т-А. Высокая точность и аккуратность описанного в настоящем описании однонуклеотидного редактора достигается посредством нуклеазы NmeCas9, которая слита с белком нуклеотиддезаминазой. Компактная природа NmeCas9, сопряженной с более высоким количеством совместимых прилегающих к протоспейсеру мотивов, обеспечивает наличие у слитых конструкций Cas9, описанных в настоящем описании, окна редактирования генов, которое может редактировать участки, на которые не нацеливаются другие общепринятые редактирующие основания платформы SpyCas9.

В одном варианте осуществления настоящее изобретение относится к мутантному белку NmeCas9, содержащему слитую с ним нуклеотиддезаминазу и связывающую область для нуклеотидной последовательности N<sub>4</sub>CC. В одном варианте осуществления указанный белок представляет собой Nme2Cas9. В одном варианте осуществления указанный белок дополнительно содержит белок сигнала ядерной локализации. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу. В одном варианте осуществления белок дополнительно содержит ингибитор урацилгликозилазы. В одном варианте осуществления указанный белок сигнала ядерной локализации включает, но не ограничивается ими, нуклеоплазмин (NLS), и/или NLS SV40, и/или NLS С-мус. В одном варианте осуществления указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера. В одном варианте осуществления указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию. В одном варианте осуществления указанная мутация представляет собой мутацию D16A. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит CBE4. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит линкер. В одном варианте осуществления указанный линкер представляет собой линкер из 73 а.к. В одном варианте осуществления указанный линкер представляет собой 3хНА-метку.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой оптимизированную конструкцию nNme2Cas9-ABEmax.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой nNme2Cas9-CBE4.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой YE1-BE3-nNme2Cas9 (D16A)-UGI.

В одном варианте осуществления настоящее изобретение относится к аденоассоциированному вирусу, содержащему мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для нуклеотидной последовательности N<sub>4</sub>CC. В одном варианте

осуществления указанный вирус представляет собой аденоассоциированный вирус 8. В одном варианте осуществления указанный вирус представляет собой аденоассоциированный вирус 6. В одном варианте осуществления указанный белок представляет собой Nme2Cas9. В одном варианте осуществления указанный белок дополнительно содержит белок сигнала ядерной локализации. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу. В одном варианте осуществления белок дополнительно содержит ингибитор урацилгликозилазы. В одном варианте осуществления белок сигнала ядерной локализации включает, но не ограничивается ими, нуклеоплазмин (NLS), и/или NLS SV40, и/или NLS С-тус. В одном варианте осуществления указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера. В одном варианте осуществления указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию. В одном варианте осуществления указанная мутация представляет собой мутацию D16A. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит CBE4. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит линкер. В одном варианте осуществления указанный линкер представляет собой линкер из 73 а.к. В одном варианте осуществления указанный линкер представляет собой 3×НА-метку.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой оптимизированную конструкцию nNme2Cas9-ABEmax.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой nNme2Cas9-CBE4.

В одном варианте осуществления настоящее изобретение относится к конструкции, где указанная конструкция представляет собой YE1-BE3-nNme2Cas9 (D16A)-UGI.

В одном варианте осуществления настоящее изобретение относится к способу, включающему: а) предоставление; i) нуклеотидной последовательности, содержащей ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC; ii) мутантного белка NmeCas9, содержащего слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC; b) приведение указанной нуклеотидной последовательности в контакт с указанным мутантным белком NmeCas9 в таких условиях, чтобы указанная связывающая область связывалась с указанной нуклеотидной последовательностью N<sub>4</sub>CC; и c) замену указанного мутантного единичного основания основанием дикого типа посредством указанного мутантного белка NmeCas9. В одном варианте осуществления указанный белок представляет собой Nme2Cas9. В одном варианте осуществления указанный белок дополнительно содержит белок сигнала ядерной локализации. В одном

варианте осуществления указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу. В одном варианте осуществления белок дополнительно содержит ингибитор урацилгликозилазы. В одном варианте осуществления белок сигнала ядерной локализации включает, но не ограничивается ими, нуклеоплазмин (NLS), и/или NLS SV40, и/или NLS C-мус. В одном варианте осуществления указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера. В одном варианте осуществления указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию. В одном варианте осуществления указанная мутация представляет собой мутацию D16A. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит CBE4. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит линкер. В одном варианте осуществления указанный линкер представляет собой линкер из 73 а.к. В одном варианте осуществления указанный линкер представляет собой 3×HA-метку. В одном варианте осуществления указанный ген кодирует тирозиназу. В одном варианте осуществления указанный ген представляет собой *Fah*. В одном варианте осуществления указанный ген представляет собой *c-fos*.

В одном варианте осуществления настоящее изобретение относится к способу, включающему: а) предоставление; i) пациента, имеющего нуклеотидную последовательность, содержащую ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC, где указанный мутантный ген вызывает медицинское состояние на генетической основе; ii) аденоассоциированный вирус, содержащий мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC; б) лечение указанного пациента указанным аденоассоциированным вирусом в таких условиях, чтобы указанный мутантный белок NmeCas9 заменял указанное мутантное единичное основание на единичное основание дикого типа, так чтобы указанное медицинское состояние на медицинской основе не развивалось. В одном варианте осуществления указанный ген кодирует белок тирозиназу. В одном варианте осуществления указанное медицинское состояние на генетической основе представляет собой тирозинемию. В одном варианте осуществления указанный вирус представляет собой аденоассоциированный вирус 8. В одном варианте осуществления указанный вирус представляет собой аденоассоциированный вирус 6. В одном варианте осуществления указанный белок представляет собой Nme2Cas9. В одном варианте осуществления указанный белок дополнительно содержит белок сигнала ядерной локализации. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу. В одном варианте осуществления указанная нуклеотиддезаминаза представляет собой

аденозиндезаминазу. В одном варианте осуществления белок дополнительно содержит ингибитор урацилгликозилазы. В одном варианте осуществления белок сигнала ядерной локализации включает, но не ограничивается ими, нуклеоплазмин (NLS), и/или NLS SV40 и/или NLS С-мус. В одном варианте осуществления указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера. В одном варианте осуществления указанный домен, взаимодействующий со вспомогательным мотивом протоспейсера, содержит указанную мутацию. В одном варианте осуществления указанная мутация представляет собой мутацию D16A. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит CBE4. В одном варианте осуществления указанный мутантный белок NmeCas9 дополнительно содержит линкер. В одном варианте осуществления указанный линкер представляет собой линкер из 73 а.к. В одном варианте осуществления указанный линкер представляет собой 3×НА-метку. В одном варианте осуществления указанный ген кодирует тирозиназу. В одном варианте осуществления указанный ген представляет собой *Fah*. В одном варианте осуществления указанный ген представляет собой *c-fos*.

В одном варианте осуществления настоящее изобретение относится к способу, включающему: а) предоставление; i) пациента, имеющего нуклеотидную последовательность, содержащую ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC, где указанный мутантный ген вызывает медицинское состояние на генетической основе; ii) оптимизированной конструкции nNme2Cas9-ABEmax, содержащий мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC; б) лечение указанного пациента указанной оптимизированной конструкцией nNme2Cas9-ABEmax в таких условиях, чтобы указанный мутантный белок NmeCas9 заменял указанное мутантное единичное основание на единичное основание дикого типа, так чтобы указанное медицинское состояние на генетической основе не развивалось.

В одном варианте осуществления настоящее изобретение относится к способу, включающему: а) предоставление; i) пациента, имеющего нуклеотидную последовательность, содержащую ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC, где указанный мутантный ген вызывает медицинское состояние на генетической основе; ii) конструкции nNme2Cas9-CBE4, содержащей мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC; б) лечение указанного пациента указанной конструкцией nNme2Cas9-CBE4 в таких условиях, чтобы указанный мутантный белок NmeCas9 заменял указанное мутантное единичное основание на единичное основание дикого типа, так чтобы указанное медицинское состояние на генетической

основе не развивалось.

В одном варианте осуществления настоящее изобретение относится к способу, включающему: а) предоставление; i) пациента, имеющего нуклеотидную последовательность, содержащую ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью  $N_4CC$ , где указанный мутантный ген вызывает медицинское состояние на генетической основе; ii) конструкции YE1-BE3-nNme2Cas9 (D16A)-UGI, содержащей мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности  $N_4CC$ ; б) лечение указанного пациента указанной конструкцией nNme2Cas9-CBE4 в таких условиях, чтобы указанный мутантный белок заменял указанное мутантное единичное основание на единичное основание дикого типа, так чтобы указанное медицинское состояние на генетической основе не развивалось.

### **Определения**

Для облегчения понимания настоящего изобретения ниже определен ряд терминов. Термины, определенные в настоящем описании, имеют значения, обычно подразумеваемые специалистом в области, к которой относится настоящее изобретение. Подразумевается, что форма единственного числа относится не только к единичному объекту, а включает общий класс, конкретный пример из которого может быть использован для иллюстрации. Терминология в настоящем описании используется для описания конкретных вариантов осуществления изобретения, однако ее применение не ограничивает изобретение за исключением того, что указано в формуле изобретения.

Как используют в рамках изобретения, термин "редактировать", "редактирование" или "отредактированный" относится к способу изменения последовательности нуклеиновой кислоты полинуклеотида (например, встречающаяся в природе нуклеиновая кислота дикого типа или мутантная встречающаяся в природе последовательность) посредством селективного удаления конкретной геномной мишени. Такая конкретная геномная мишень включает, но не ограничивается ими, хромосомную область, ген, промотор, открытую рамку считывания или любую последовательность нуклеиновой кислоты.

Как используют в рамках изобретения, термин "единичное основание" относится к одному и только одному нуклеотиду в последовательности нуклеиновой кислоты. При использовании в контексте редактирования единичных основания, он означает, что основание в конкретном положении последовательности нуклеиновой кислоты заменено на отличающееся основание. Эта замена может происходить посредством многих механизмов, включая, но не ограничиваясь ими, замену или модификацию.

Как используют в рамках изобретения, термин "мишень" или "участок-мишень" относится к предварительно идентифицированной последовательности нуклеиновой кислоты любого состава и/или длины. Такие участки-мишени включают, но не

ограничивается ими, хромосомную область, ген, промотор, открытую рамку считывания или любую последовательность нуклеиновой кислоты. В некоторых вариантах осуществления настоящее изобретение распознает эти конкретные геномные последовательности-мишени посредством комплементарных последовательностей гРНК.

Термин "последовательность связывания в мишени", как используют в рамках изобретения, относится к подпоследовательности конкретной геномной мишени, которая может быть полностью комплементарной программируемому ДНК-связывающему домену и/или единичной последовательности гидовой РНК.

Термин "последовательность связывания вне мишени", как используют в рамках изобретения, относится к подпоследовательности конкретной геномной мишени, которая может быть частично комплементарной программируемому ДНК-связывающему домену и/или единичной последовательности гидовой РНК.

Термин "эффективное количество", как используют в рамках изобретения, относится к конкретному количеству фармацевтической композиции, содержащей терапевтическое средство, которое достигает клинически благоприятного результата (т.е., например, уменьшения симптомов). Токсичность и терапевтическая эффективность таких композиций может быть определена посредством стандартных фармацевтических методик в клеточных культурах или у экспериментальных животных, например, определяющих LD<sub>50</sub> (доза, летальная для 50% популяции) и ED<sub>50</sub> (доза, терапевтически эффективная в 50% популяции). Соотношение доз между токсическими и терапевтическими эффектами является терапевтическим индексом, и он может быть выражен как соотношение LD<sub>50</sub>/ED<sub>50</sub>. Предпочтительными являются соединения, которые имеют большие терапевтические индексы. Данные, полученные в этих анализах клеточных культур и дополнительных исследованиях на животных, можно использовать при составлении ряда дозировок для применения у человека. Дозировка таких соединений предпочтительно находится в диапазоне циркулирующих концентраций, которые включают ED<sub>50</sub> с малой токсичностью или без нее. Дозировка варьируется в этом диапазоне в зависимости от используемой дозированной формы, чувствительности пациента и пути введения.

Термин "симптом", как используют в рамках изобретения, относится к любому субъективному или объективному признаку заболевания или физического нарушения, наблюдаемого у пациента. Например, субъективный признак обычно основан на сообщении непосредственно пациентом и может включать, но не ограничиваться ими, боль, головную боль, нарушение зрения, тошноту и/или рвоту. Альтернативно объективный признак обычно является результатом медицинского тестирования, включающего, но не ограничивающегося ими, температуру тела, развернутый анализ крови, липидограмму, исследование гормональной активности щитовидной железы, кровяное давление, частоту сердечбиений, электрокардиограмму, сканирование для визуализации ткани и/или организма.

Термин "заболевание" или "медицинское состояние", как используют в рамках изобретения, относится к любому нарушению нормального состояния живого животного или растения, или одной из их частей, которое прекращает или модифицирует выполнение жизненных функций. Обычно проявляющееся характерными признаками и симптомами, оно является ответом на: i) факторы внешней среды (такие как нарушение питания, производственный риск или климат); ii) конкретные инфекционные агенты (такие как черви, бактерии или вирусы); iii) собственные дефекты организма (такие как генетические нарушения); и/или iv) комбинации этих факторов.

Термины "снижать", "ингибировать", "уменьшать", "подавлять", "сокращать", "предотвращать" и их грамматические эквиваленты (включая "более низкий", "меньший" и т.д.) при указании на проявление какого-либо симптома у не подвергаемого лечению индивидууму относительно подвергаемого лечению индивидуума, означает, что количество и/или сила симптомов у подвергаемого лечению индивидуума являются более низкими, чем у не подвергаемого лечению индивидуума, на величину, которая признана клинически значимой любым лицом с медицинским образованием. В одном варианте осуществления количество и/или сила симптомов у подвергаемого лечению индивидуума по меньшей мере на 10% ниже, по меньшей мере на 25% ниже, по меньшей мере на 50% ниже, по меньшей мере на 75% ниже, и/или по меньшей мере на 90% ниже, чем количество и/или сила симптомов у не подвергаемого лечению индивидуума.

Термин "связанный", как используют в рамках изобретения, относится к любому взаимодействию между средой (или носителем) и лекарственным средством. Связывание может быть обратимым или необратимым. Такое связывание включает, но не ограничивается ими, ковалентное связывание, ионное связывание, ван-дер-ваальсовы силы или трение, и т.п. Лекарственное средство связано со средой (или носителем), если оно заполняет, включено, покрыто, находится в суспензии с, в растворе с, смешано с ней, и т.д.

Термин "лекарственное средство" или "соединение", как используют в рамках изобретения, относится к любому фармакологически активному веществу, которое может быть введено, которое достигает желаемого эффекта. Лекарственные средства или соединения могут быть синтетическими или встречающимися в природе, непептидными, белками или пептидами, олигонуклеотидами или нуклеотидами, полисахаридами или сахарами.

Термин "введенный" или "введение", как используют в рамках изобретения, относится к любому способу предоставления композиции пациенту, так чтобы композиция имела предполагаемый эффект на пациента. Иллюстративным способом введения является введение посредством прямого механизма, такого как локальное введение в ткань (т.е., например, внесосудистое размещение), пероральный прием, трансдермальный пластырь, местный путь, ингаляция, суппозиторий и т.д.

Термин "пациент" или "субъект", как используют в рамках изобретения,

представляет собой человека или животного, и он не должен быть госпитализированным. Например, "пациентами" являются амбулаторные пациенты, лица в домах престарелых. Пациент может включать человека или не являющееся человеком животное любого возраста и, таким образом, он включает как взрослых, так и малолетних (т.е. детей). Подразумевается, что термин "пациент" включает в себе потребность в медицинском лечении, таким образом, пациент может быть добровольно или непроизвольно частью эксперимента как клинического, так и фундаментального исследования.

Термин "аффинность", как используют в рамках изобретения, относится к любой силе притяжения между веществами или частицами, которая вынуждает их вступать и оставаться в химической комбинации. Например, соединение-ингибитор, которое обладает высокой аффинностью к рецептору, обеспечивает более высокую эффективность предупреждения взаимодействия рецептора с его природными лигандами, чем ингибитор с низкой аффинностью.

Термин "фармацевтически" или "фармакологически приемлемый", как используют в рамках изобретения, относятся к молекулярным структурам и композициям, которые не вызывают неблагоприятных, аллергических или других нежелательных реакций при введении животному или человеку.

Термин "фармацевтически приемлемый носитель", как используют в рамках изобретения, включает любые и все растворители или дисперсионные среды, включающие, но не ограничивающиеся ими, воду, этанол, полиол (например, глицерин, пропиленгликоль и жидкий полиэтиленгликоль, и т.п.), их подходящие смеси и растительные масла, покрытия, изотонические и замедляющие всасывание средства, липосомы, коммерчески доступные моющие средства и т.п. В такие носители также могут быть добавлены дополнительные биологически активные ингредиенты.

Термин "вирусный вектор" охватывает любую конструкцию нуклеиновой кислоты, происходящую из вирусного генома, способную вмещать гетерологичные последовательности нуклеиновых кислот, для экспрессии в организме-хозяине. Например, такие вирусные векторы могут включать, но не ограничиваться ими, векторы на основе аденоассоциированного вируса, лентивирусные векторы, векторы на основе вируса SV40, ретровирусные векторы, аденовирусные векторы. Хотя вирусные векторы редко создают из патогенных вирусов, они могут быть модифицированы таким образом, чтобы минимизировать их общий риск для здоровья. Это обычно вовлекает делецию части вирусного генома, вовлеченного в репликацию вируса. Такой вирус может инфицировать клетки, однако после инфицирования вирусу может потребоваться вирус-помощник для предоставления отсутствующих белков для продуцирования новых вирионов. Предпочтительно, вирусные векторы должны иметь минимальный эффект на физиологию клетки, которую они инфицируют, и демонстрируют свойства генетической стабильности (например, не претерпевают спонтанной перестройки генома). Большинство вирусных векторов конструируют для инфицирования настолько широкого диапазона типов клеток,

насколько это возможно. Тем не менее, вирусный рецептор может быть модифицирован для нацеливания вируса на конкретный тип клеток. Вирусы, модифицированные таким образом, называют псевдотипированными. Вирусные векторы часто конструируют так, чтобы они включали определенные гены, которые помогают идентифицировать, какие клетки приобрели вирусные гены. Эти гены называют маркерными генами. Например, часто маркерный ген сообщает резистентность к определенному антибиотику.

Как используют в рамках изобретения, "ген *ROSA26*" или "ген *Rosa26*" относится к локусу человека или мыши (соответственно), который широко используется для достижения генерализованной экспрессии у мышей. Нацеливания на локус *ROSA26* можно достигать путем встраивания желаемого гена в первый интрон локуса в уникальном участке *XbaI* приблизительно на 248 п.о. выше исходной линии генной ловушки. Конструкцию можно конструировать с использованием аденовирусного акцептора сплайсинга, за которым следует представляющий интерес ген, и участка полиаденилирования, встроенного в уникальный участок *XbaI*. Также в нацеливающий вектор может быть включена кассета резистентности к неомицину.

Как используют в рамках изобретения, "ген *PCSK9*" или "ген *Pcsk9*" относится к локусу человека или мыши (соответственно), который кодирует белок PCSK9. Ген *PCSK9* находится на хромосоме 1 в полосе 1p32.3 и включает 13 экзонов. Посредством альтернативного сплайсинга с этого гена может продуцироваться по меньшей мере две изоформы.

Термин "конвертаза пробелков субтилизин/кексинового типа 9" и "PCSK9" относится к белку, кодируемому геном, который модулирует уровни липопротеинов низкой плотности. Конвертаза пробелков субтилизин/кексинового типа 9, также известная как PCSK9, представляет собой фермент, который у человека кодируется геном *PCSK9*. Seidah et al., "The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation" *Proc. Natl. Acad. Sci. U.S.A.* 100 (3): 928-933 (2003). Сходные гены (ортологи) встречаются среди многих видов. Многие ферменты, включая PCSK9, являются неактивными, когда они первоначально синтезируются, поскольку они имеют сегмент из пептидных цепей, который блокирует их активность; конвертазы пробелков удаляют этот сегмент, активируя фермент. Полагают, что PCSK9 играет регуляторную роль в гомеостазе холестерина. Например, PCSK9 может связываться с доменом подобного эпидермальному фактору роста повтора А (EGF-A) рецептора липопротеинов низкой плотности (LDL-R), вызывая интернализацию и деградацию LDL-R. Очевидно, можно ожидать, что снижение уровней LDL-R приводит к снижению метаболизма LDL-C, которое может приводить к гиперхолестеринемии.

Термин "гиперхолестеринемия", как используют в рамках изобретения, относится к любому медицинскому состоянию, где уровни холестерина в крови повышены выше клинически рекомендованных уровней. Например, если уровень холестерина измеряют с использованием липопротеинов низкой плотности (LDL), гиперхолестеринемия может

существовать, когда измеренные уровни LDL превышают, например, приблизительно 70 мг/дл. Альтернативно, если уровень холестерина измеряют с использованием свободного холестерина плазмы, гиперхолестеринемия может существовать, когда измеренные уровни свободного холестерина превышают, например, приблизительно 200-220 мг/дл.

Как используют в рамках изобретения, термин "CRISPR" или "короткие палиндромные повторы, регулярно расположенные кластерами" относится к сокращенному названию локусов ДНК, которые содержат множество коротких прямых повторов последовательностей оснований. Каждый повтор содержит серию оснований, за которой следует 30 или около того оснований, известных как "спейсерная ДНК". Спейсеры представляют собой сегменты ДНК из вируса, и они могут служить в качестве "памяти" о предшествующих воздействиях для способствования адаптивной защите против будущих инвазий.

Как используют в рамках изобретения, термин "Cas" или "CRISPR-ассоциированный (cas)" относится к генам, часто ассоциированным с системами повтор-спейсер CRISPR.

Как используют в рамках изобретения, термин "Cas9" относится к нуклеазе из систем CRISPR типа II - ферменту, специализированному на внесении двухцепочечных разрывов в ДНК с двумя активными участками расщепления (домены HNH и RuvC), по одному для каждой цепи двойной спирали. Jinek скомбинировал tracr-РНК и спейсерную РНК в молекулу "единой гидовой РНК" (sg-РНК), которая в смеси с Cas9 может находить и расщеплять ДНК-мишени посредством образования пар Уотсона-Крика между гидовой последовательностью в sg-РНК и последовательностью ДНК, являющейся мишенью.

Термин "прилегающий к протоспейсеру мотив" (или РАМ), как используют в рамках изобретения, относится к последовательности ДНК, которая может потребоваться для Cas9/sg-РНК чтобы сформировать R-петлю для встраивания в конкретную последовательность ДНК через образование пар по принципу Уотсона-Крика их гидовой РНК с геномом. Специфичность к РАМ может быть функцией специфичности белка Cas9 в отношении связывания ДНК (например, "домен распознавания прилегающего к протоспейсеру мотива" на С-конце Cas9).

Как используют в рамках изобретения, термин "sg-РНК" относится к единичной гидовой РНК, используемой совместно с ассоциированными с CRISPR системами (Cas). sg-РНК представляют собой слитую конструкцию из cr-РНК и tracr-РНК, и они содержат нуклеотиды последовательности, комплементарной желаемому участку-мишени. Jinek et al., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity" *Science* 337(6096):816-821 (2012). Образование пар по принципу Уотсона-Крика между sg-РНК и участком-мишенью обеспечивает формирование R-петли, которая совместно с функциональным РАМ обеспечивает расщепление ДНК или в случае Cas9 с нуклеазным дефицитом позволяет связывание с ДНК в этом локусе.

Как используют в рамках изобретения, термин "флуоресцентный белок" относится

к белковому домену, который содержит по меньшей мере одну часть в виде органического соединения, которая испускает флуоресцентный свет в ответ на соответствующую длину волны. Например, флуоресцентные белки могут испускать красный, синий и/или зеленый свет. Такие белки являются коммерчески доступными, включая, но не ограничиваясь ими: i) mCherry (Clontech Laboratories): возбуждение: 556/20 нм (длина волны/ширина полосы); испускание: 630/91 нм; ii) sfGFP (Invitrogen): возбуждение: 470/28 нм; испускание: 512/23 нм; iii) TagBFP (Evrogen): возбуждение 387/11 нм; испускание 464/23 нм.

Как используют в рамках изобретения, термин "sg-РНК" относится к единичной гидовой РНК, используемой совместно с CRISPR-ассоциированными системами (Cas). sg-РНК содержат нуклеотиды последовательности, комплементарной желаемому участку-мишени. Образование пар по принципу Уотсона-Крика между sg-РНК и участком-мишенью вовлекает Cas9 с нуклеазным дефицитом в связывание с ДНК в этом локусе.

Как используют в рамках изобретения, термин "ортогональный" относится к мишеням, которые являются не перекрывающимися, некоррелирующими или независимыми. Например, если используется два ортогональных гена Cas9 с нуклеазным дефицитом, слитых с различными эффекторными доменами, то sg-РНК, кодируемые для каждой из них, не являются перекрестными или не перекрываются. Не все гены Cas9 с нуклеазным дефицитом функционируют одинаково, что позволяет использовать ортогональные гены Cas9 с нуклеазным дефицитом, слитые с различными эффекторными доменами, при условии наличия соответствующих ортогональных sg-РНК.

Как используют в рамках изобретения, термин "фенотипическое изменение" или "фенотип" относится к набору наблюдаемых характеристик или признаков организма, таких как его морфология, развитие, биохимические или физиологические свойства, фенология, поведение и результаты поведения. Фенотипы являются результатом экспрессии генов организма, а также влияния факторов внешней среды и взаимодействий между ними.

"Последовательность нуклеиновой кислоты" и "нуклеотидная последовательность", как используют в рамках изобретения, относятся к олигонуклеотиду или полинуклеотиду, и их фрагментам или частям, и к ДНК или РНК геномного или синтетического происхождения, которые могут быть одноцепочечными или двухцепочечными и соответствовать смысловой или антисмысловой цепи.

Термин "выделенная нуклеиновая кислота", как используют в рамках изобретения, относится к любой молекуле нуклеиновой кислоты, которая извлечена из ее природного окружения (например, извлечена из клетки и в предпочтительном варианте осуществления свободна от другой геномной нуклеиновой кислоты).

Термины "аминокислотная последовательность" и "полипептидная последовательность", как используют в рамках изобретения, являются взаимозаменяемыми и относятся к последовательности аминокислот.

Как используют в рамках изобретения, термин "часть" в отношении белка (как в случае "части данного белка") относится к фрагментам этого белка. Размер фрагментов может находиться в диапазоне от четырех аминокислотных остатков до целой аминокислотной последовательности минус одна аминокислота.

Термин "часть", когда его используют в отношении нуклеотидной последовательности, относится к фрагментам этой нуклеотидной последовательности. Размер фрагментов может находиться в диапазоне от 5 нуклеотидных остатков до целой нуклеотидной последовательности минус один остаток нуклеиновой кислоты.

Как используют в рамках изобретения, термины "комплементарный" или "комплементарность" используют в отношении "полинуклеотидов" и "олигонуклеотидов" (которые являются взаимозаменяемыми терминами, которые относятся к последовательности нуклеотидов), объединенных правилами образования пар оснований. Например, последовательность "С-А-Г-Т" комплементарна последовательности "G-T-C-A". Комплементарность может быть "частичной" или "полной". "Частичная" комплементарность относится к случаю, когда одно или более оснований нуклеиновой кислоты не соответствуют согласно правилам образования пар оснований. "Тотальная" или "полная" комплементарность между нуклеиновыми кислотами относится к случаю, когда каждое основание нуклеиновой кислоты соответствует другому основанию согласно правилам образования пар оснований. Степень комплементарности между цепями нуклеиновых кислот имеет значительные эффекты на эффективность и силу гибридизации между цепями нуклеиновых кислот. Это имеет особое значение в реакциях амплификации, а также в способах детекции, которые зависят от связывания между нуклеиновыми кислотами.

Термины "гомология" и "гомологичный", как используют в рамках изобретения в отношении нуклеотидных последовательностей, относятся к степени комплементарности с другими нуклеотидными последовательностями. Может существовать частичная гомология или полная гомология (т.е. идентичность). Нуклеотидная последовательность, которая является частично комплементарной, т.е. "по существу гомологичной" последовательности нуклеиновой кислоты, представляет собой последовательность, которая по меньшей мере частично ингибирует гибридизацию полностью комплементарной последовательности с последовательностью нуклеиновой кислоты-мишенью. Ингибирование гибридизации полностью комплементарной последовательности с последовательностью-мишенью можно исследовать с использованием анализа гибридизации (саузерн- или нозерн-блоттинг, гибридизация в растворе и т.п.) в условиях низкой жесткости. По существу гомологичная последовательность или зонд конкурируют за и ингибируют связывание (т.е. гибридизацию) полностью гомологичной последовательности с последовательностью-мишенью в условиях низкой жесткости. Это вовсе не означает, что условия низкой жесткости являются такими, что допускается неспецифическое связывание; условия

низкой жесткости требуют, чтобы связывание двух последовательностей друг с другом представляло собой специфическое (т.е. селективное) взаимодействие. Отсутствие неспецифического связывания можно тестировать с использованием второй последовательности-мишени, которая лишена даже частичной степени комплементарности (например, идентичность менее 30%); в отсутствие неспецифического связывания зонд не гибридизуется со второй некомплементарной мишенью.

Термины "гомология" и "гомологичный", как используют в рамках изобретения в отношении аминокислотных последовательностей, относятся к степени идентичности первичной структуры между двумя аминокислотными последовательностями. Такая степень идентичности может относиться к части каждой аминокислотной последовательности или ко всей длине аминокислотной последовательности. Две или более аминокислотных последовательности, которые являются "по существу гомологичными", могут обладать по меньшей мере 50% идентичностью, предпочтительно по меньшей мере 75% идентичностью, более предпочтительно по меньшей мере 85% идентичностью, наиболее предпочтительно по меньшей мере 95%, или 100% идентичностью.

Олигонуклеотидную последовательность, которая является "гомологом", определяют в рамках настоящего изобретения как олигонуклеотидную последовательность, которая демонстрирует 50% или более идентичность с последовательностью, когда сравнивают последовательности, имеющие длину 100 п.о. или более.

Условия низкой жесткости включают условия, эквивалентные связыванию или гибридизации при 42°C в растворе, состоящем из 5×SSPE (43,8 г/л NaCl, 6,9 г/л  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$  и 1,85 г/л EDTA, с pH 7,4, доведенным NaOH), 0,1% SDS, 5×раствор Денхардта {50× раствор Денхардта содержит на 500 мл: 5 г Ficoll (Type 400, Pharmacia), 5 г BSA (Fraction V; Sigma)} и 100 мкг/мл денатурированной ДНК спермы лосося, а затем промывание в растворе, содержащем 5×SSPE, 0,1% SDS при 42°C, когда используют зонд длиной приблизительно 500 нуклеотидов. Также можно использовать многочисленные эквивалентные условия для обеспечения условий низкой жесткости; факторы, такие как длина и природа (ДНК, РНК, состав оснований) зонда и природа мишени (ДНК, РНК, состав оснований, находящаяся в растворе или иммобилизованная, и т.д.) и концентрация солей и других компонентов (например, присутствие или отсутствие формамида, декстрансульфата, полиэтиленгликоля), а также компоненты раствора для гибридизации, можно варьировать для обеспечения гибридизации низкой жесткости, отличной от, но эквивалентной, перечисленным выше условиям. Кроме того, также можно использовать условия, которые способствуют гибридизации в условиях высокой жесткости (например, повышение температуры гибридизации и/или стадий промывания, использование формамида в растворе для гибридизации и т.д.).

Как используют в рамках изобретения, термин "гибридизация" используют в

отношении образования пар между комплементарными нуклеиновыми кислотами с использованием любого процесса, посредством которого цепь нуклеиновой кислоты связывается с комплементарной цепью посредством образования пар, формируя гибридизационный комплекс. На гибридизацию и силу гибридизации (т.е. сила ассоциации между нуклеиновыми кислотами) влияют такие факторы, как степень комплементарности между нуклеиновыми кислотами, жесткость используемых условий,  $T_m$  образовавшегося гибрида и соотношение G:C в нуклеиновых кислотах.

Как используют в рамках изобретения, термин "гибридизационный комплекс" относится к комплексу, образовавшемуся между двумя последовательностями нуклеиновых кислот вследствие образования водородных связей между комплементарными основаниями G и C и между комплементарными основаниями A и T; эти водородные связи могут быть далее стабилизированы посредством стэкинг-взаимодействий между основаниями. Две комплементарных последовательности нуклеиновых кислот образуют водородные связи в антипараллельной конфигурации. Гибридизационный комплекс может образовываться в растворе (например, анализ  $C_0 t$  или  $R_0 t$ ) или между одной последовательностью нуклеиновой кислоты, присутствующей в растворе, и другой последовательностью нуклеиновой кислоты, иммобилизованной на твердой подложке (например, нейлоновая мембрана или нитроцеллюлозный фильтр, как используют в саузерн- или нозерн-блоттинге, дот-блоттинге, или предметное стекло, как используют в гибридизации *in situ*, включая FISH (флуоресцентная гибридизация *in situ*)).

Молекулы ДНК называют имеющими "5'-концы" и "3'-концы", поскольку мононуклеотиды реагируют с образованием олигонуклеотидов так, что 5'-фосфат пентозного кольца одного мононуклеотида связывается с 3'-кислородом соседнего с ним нуклеотида в одном направлении посредством фосфодиэфирной связи. Таким образом, конец олигонуклеотида называют "5'-концом", если его 5'-фосфат не связан с 3'-кислородом пентозного кольца мононуклеотида. Конец олигонуклеотида называют "3'-концом", если его 3'-кислород не связан с 5'-фосфатом пентозного кольца другого мононуклеотида. Как используют в рамках изобретения, также может быть указано, что последовательность нуклеиновой кислоты, даже если она находится внутри более крупного олигонуклеотида, имеет 5'- и 3'-концы. Как в линейной, так и в кольцевой молекуле ДНК, отдельные элементы называют "вышележащими" или 5'- относительно "нижележащих" или 3'-элементов. Эта терминология отражает тот факт, что транскрипция протекает от 5' к 3' вдоль цепи ДНК. Промоторные и энхансерные элементы, которые обеспечивают транскрипцию связанного с ними гена, как правило находятся с 5'-стороны или выше кодирующей области. Однако энхансерные элементы могут демонстрировать их эффект, даже когда они находятся с 3'-стороны от промоторного элемента и кодирующей области. Сигналы терминации транскрипции и полиаденилирования находятся с 3'-стороны или ниже кодирующей области.

Термин "трансфекция" или "трансфицированный" относится к введению

чужеродной ДНК в клетку.

Как используют в рамках изобретения, термины "молекула нуклеиновой кислоты, кодирующая", "последовательность ДНК, кодирующая" и "ДНК, кодирующая" относятся к порядку или последовательности дезоксирибонуклеотидов вдоль цепи дезоксирибонуклеиновой кислоты. Порядок этих дезоксирибонуклеотидов определяет порядок аминокислот вдоль полипептидной (белковой) цепи. Таким образом, последовательность ДНК кодирует аминокислотную последовательность.

Как используют в рамках изобретения, термин "ген" означает последовательности дезоксирибонуклеотидов, содержащие кодирующую область структурного гена и включающие последовательности, находящиеся рядом с кодирующей областью как на 5'-, так и на 3'-концах на расстоянии приблизительно 1 т.п.н. с каждой стороны, так что ген соответствует длине полноразмерной мРНК. Последовательности, которые находятся с 5'-стороны от кодирующей области и которые присутствуют на мРНК, называют 5'-нетранслируемыми последовательностями. Последовательности, которые находятся с 3'-стороны или ниже кодирующей области и которые присутствуют на мРНК, называют 3'-нетранслируемыми последовательностями. Термин "ген" охватывает как кДНК, так и геномные формы гена. Геномная форма или клон гена содержит кодирующую область, которая прерывается некодирующими последовательностями, называемыми "интронами", или "вставочными областями", или "вставочными последовательностями". Интроны представляют собой сегменты гена, которые транскрибируются в гетерогенную ядерную РНК (hn-РНК); интроны могут содержать регуляторные элементы, такие как энхансеры. Интроны удаляются или "вырезаются" из ядерного или первичного транскрипта; таким образом, в транскрипте матричной РНК (мРНК) интроны отсутствуют. мРНК функционируют в ходе трансляции, определяя последовательность или порядок аминокислот в образующемся полипептиде.

В дополнение к наличию интронов, геномные формы гена также могут включать последовательности, находящиеся как на 5'-, так и на 3'-конце последовательностей, которые присутствуют в РНК-транскрипте. Эти последовательности называют "фланкирующими" последовательностями или областями (эти фланкирующие последовательности находятся с 5'- или 3'-стороны от нетранслируемых последовательностей, присутствующих на мРНК-транскрипте). 5'-фланкирующая область может содержать регуляторные последовательности, такие как промоторы и энхансеры, которые контролируют или влияют на транскрипцию гена. 3'-фланкирующая область может содержать последовательности, которые контролируют терминацию транскрипции, посттрансляционное расщепление и полиаденилирование.

Термины "метка" или "поддающаяся детекции метка" используют в настоящем описании для указания на любую композицию, поддающуюся детекции спектроскопическими, фотохимическими, биохимическими, иммунохимическими, электрическими, оптическими или химическими способами. Такие метки включают

биотин для окрашивания меченым конъюгатом стрептавидина, магнитные гранулы (например, Dynabeads®), флуоресцентные красители (например, флуоресцеин, тexasский красный, родамин, зеленый флуоресцентный белок и т.п.), радиоактивные метки (например,  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$  или  $^{32}\text{P}$ ), ферменты (например, пероксидаза хрена, щелочная фосфатаза и другие, часто используемые в ELISA), и калориметрические метки, такие как коллоидное золото или гранулы из окрашенного стекла или пластмассы (например, полистирол, полипропилен, латекс и т.д.). Патенты, в которых описано применение таких меток, включают, но не ограничиваются ими, патенты США № 3817837; 3850752; 3939350; 3996345; 4277437; 4275149 и 4366241 (все из которых включены в настоящее описание в качестве ссылок в полном объеме). Метки, охватываемые настоящим изобретением, можно выявлять многими способами. Например, детекцию радиоактивных меток можно проводить с использованием фотографической пленки или сцинтилляционных счетчиков, детекцию флуоресцентных маркеров можно проводить с использованием фотодетектора для детекции испускаемого света. Детекцию ферментных меток, как правило, проводят путем предоставления ферменту субстрата и детекции продукта реакции, продуцированного посредством действия фермента на субстрат, и детекцию калориметрических меток проводят, просто визуализируя окрашенную метку.

#### **Краткое описание чертежей**

Файл патента или заявки содержит по меньшей мере один чертеж, выполненный в цвете. Копии этой публикации патента или патентной заявки с цветным чертежом(ами) будут предоставлены офисом по запросу и при оплате требуемой пошлины.

На фиг. 1 проиллюстрированы иллюстративные схематические варианты осуществления редактора единичных оснований на основе слитого белка NmeCas9-дезаминаза и иллюстративные сконструированные плазмиды редакторов оснований.

На фиг. 1A представлена иллюстративная конструкция YE1-BE3-nNme2Cas9 (D16A)-UGI.

На фиг. 1B представлена иллюстративная конструкция ABE7.10 nNme2Cas9 (D16A).

На фиг. 1C представлена иллюстративная конструкция ABE7.10-nNme2Cas9 (D16A), содержащая две последовательности NLS SV40.

На фиг. 1D представлена иллюстративная конструкция nNme2Cas9-CBE4 (также называемая BE4-nNme2Cas9 (D16A)-UGI-UGI).

На фиг. 1E представлена иллюстративная оптимизированная конструкция nNme2Cas9-ABEmax.

На фиг. 2 представлены иллюстративные данные электропорации клеток HEK293T посредством ДНК-плазмид, включающих слитый белок YE1-BE3-nNme2Cas9 (D16A)-UGI, эффективно конвертирующий C на T в эндогенном участке-мишени 25 (TS25) в клетках HEK293T посредством нуклеофекции.

На фиг. 2A представлены иллюстративные последовательности для эндогенного

участка-мишени TS25 (в черном прямоугольнике). Sg-РНК GN23 образует пары с цепью ДНК-мишенью, оставляя вытесненную цепь ДНК для редактирования цитидиндезаминазой (например, новые зеленые нуклеотиды).

На фиг.2В представлены иллюстративные данные секвенирования, показывающие двойной нуклеотидный пик (7-положение с 5'-конца; стрелка), указывающий на успешное редактирование единичного основания цитидина на тимидин (например, конвертирование пар оснований С-G на пару оснований Т-А).

На фиг.2С представлено иллюстративное количественное определение данных, представленных на фиг.2В, в форме графика процентного конвертирования С → Т при редактировании единичного основания. Процент С, конвертированного в Т, составляет приблизительно 40% в образце, обработанном редактором оснований и sg-РНК (значение  $p=6,88 \times 10^{-6}$ ). Контроль "без sg-РНК" демонстрирует фоновый шум при секвенировании по методу Сэнгера. Для проведения анализа использовали EditR (Kluesner et al., 2018).

На фиг.3 представлены иллюстративные определенные участки-мишени UGI, которые соответственно интегрировали в слитые белки YE1-BE3-nNme2Cas9/мутант D16A и совместно экспрессировали с усиленным зеленым флуоресцентным белком (EGFP) в стабильной происходящей из K562 клеточной линии. Конвертированные основания выделены оранжевым цветом. Фоновые сигнал фильтровали с использованием образцов отрицательного контроля (клетки K562 с введенным посредством нуклеофекции YE1-BE3-nNme2Cas9 без конструкций sg-РНК). PAM N<sub>4</sub>CC заключены в рамку. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце.

На фиг.3А представлен иллюстративный участок 1 EGFP.

На фиг.3В представлен иллюстративный участок 2 EGFP.

На фиг.3С представлен иллюстративный участок 3 EGFP.

На фиг.3D представлен иллюстративный участок 4 EGFP.

На фиг.3Е представлен анализ посредством глубокого секвенирования, указывающий на то, где YE1-BE3-nNme2Cas9 конвертирует остатки С в остатки Т в эндогенной промоторной области *c-fos*. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце. Конвертированные основания выделены оранжевым или желтым цветом. Фоновые сигналы отфильтровывали с использованием образцов отрицательного контроля. Наиболее высокий процент редактирования составляет 32,50%.

На фиг.3F представлен иллюстративный анализ посредством глубокого секвенирования, указывающий на то, где ABE7.10-nNme2Cas9 или ABEmax (Koblan et al., 2018)-nNme2Cas9 конвертирует остатки А в остатки G в эндогенной промоторной области *c-fos*. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце. Конвертированные основания выделены оранжевым или желтым цветом. Фоновые сигналы отфильтровывали с использованием

образцов отрицательного контроля. Процент редактирования составляет 0,53% посредством ABE7.10-nNme2Cas9 или 2.33% посредством ABEmax-nNme2Cas9 (D16A).

На фиг.4 представлено иллюстративное выравнивание гена *Fah* дикого типа с мутантным геном тирозинемии *Fah*, демонстрирующее участок-мишень гена для редактирования единичного основания A-G (положение 9). Соответствующий единичный участок PAM SpyCas9 и двойные участки PAM NmeCas9 указаны для демонстрации субоптимального окна нацеливания относительно участка PAM SpyCas9.

На фиг.5 проиллюстрированы иллюстративные три близкородственных ортолога Cas9 *Neisseria meningitidis*, которые имеют различные PAM.

На фиг.5A представлена иллюстративная схема, демонстрирующая мутантные остатки (оранжевые сферы) между Nme2Cas9 (слева) и Nme3Cas9 (справа), картированные на спрогнозированной структуре Nme1Cas9, отображающей кластер мутаций в PID (черный).

На фиг.5B представлена иллюстративная экспериментальная схема процесса анализа обнаружения PAM *in vitro* с рандомизированной областью PAM из 10 п.н. После расщепления *in vitro* адаптеры лигировали с продуктами расщепления для конструирования библиотеки и секвенирования.

На фиг.5C представлены иллюстративные логотипы последовательностей, полученных в результате обнаружения PAM *in vitro*, выявившего увеличение содержания PAM N<sub>4</sub>GATT для Nme1Cas9 в соответствии с ее ранее установленной специфичностью.

На фиг.5D представлены иллюстративные логотипы последовательностей, указывающие на то, что Nme1Cas9 с заменой его PID на PID Nme2Cas9 (слева) или Nme3Cas9 (справа) требует C в положении 5 PAM. Остальные нуклеотиды не были определены с высокой достоверностью вследствие умеренной эффективности расщепления белковых химер с заменой PID (см. фиг.6C).

На фиг.5E представлен иллюстративный логотип последовательности, демонстрирующий, что полноразмерная Nme2Cas9 распознает PAM N<sub>4</sub>CC, на основе эффективного расщепления субстрата совокупности мишеней с фиксированным C в 5 положении PAM, и с рандомизированными нуклеотидами 1-4 и 6-8 PAM.

На фиг.6 представлена охарактеризация ортологов Cas9 *Neisseria meningitidis* с быстро эволюционирующими PID, связанная с фиг.5.

На фиг.6A представлено иллюстративное некорневое филогенетическое древо для ортологов NmeCas9, которые являются на >80% идентичными Nme1Cas9. Появилось три различных ветви, причем большинство мутантов кластеризуются в PID. Группы 1 (синий), 2 (оранжевый) и 3 (зеленый) имеют PID с идентичностью >98%, ~52% и ~86% с Nme1Cas9, соответственно. Указаны три репрезентативных ортолога Cas9 (по одному из каждой группы) (Nme1Cas9, Nme2Cas9 и Nme3Cas9).

На фиг.6B представлена иллюстративная схема, демонстрирующая локусы CRISPR-*cas* цепей, кодирующих три ортолога Cas9 (Nme1Cas9, Nme2Cas9 и Nme3Cas9) из

(A). Представлена процентная идентичность каждого компонента CRISPR-Cas с *N. meningitidis* 8013 (кодирующей Nme1Cas9). Синяя и красная стрелки обозначают участки инициации транскрипции пре-cr-РНК и tracr-РНК, соответственно.

На фиг.6С представлен график иллюстративного нормализованного количества чтений (% от всех чтений) расщепленных ДНК из анализов *in vitro* в отношении интактной Nme1Cas9 (серый), в отношении химер с заменой PID Nme1Cas9 на PID из Nme2Cas9 и Nme3Cas9 (смешанные цвета) и полноразмерной Nme2Cas9 (оранжевый). Уменьшенное нормализованное количество чтений указывает на более низкую эффективность расщепления в химерах.

На фиг.6D представлены иллюстративные логотипы последовательностей из анализа для обнаружения PAM *in vitro* на совокупности PAM NNNNCNNN посредством Nme1Cas9 с заменой PID на PID Nme2Cas9 (слева) или Nme3Cas9 (справа).

На фиг.7 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 использует спейсер из 22-24 нуклеотидов для редактирования участков, соседних с PAM N<sub>4</sub>CC. Все эксперименты проводили в трех повторениях, и планки погрешности соответствуют стандартной ошибке среднего значения (s.e.m.).

На фиг.7A представлена иллюстративная схематическая диаграмма, на которой изображена временная трансфекция и редактирование клеток HEK293T TLR2.0 с детекцией mCherry+ клеток посредством проточной цитометрии через 72 часа после трансфекции.

На фиг.7B представлено иллюстративное редактирование посредством Nme2Cas9 репортера TLR2.0. На участки с PAM N<sub>4</sub>CC происходило нацеливание с различной эффективностью, в то время как не наблюдали нацеливания Nme2Cas9 на PAM N<sub>4</sub>GATT или в отсутствие sg-РНК. В качестве положительных контролей использовали SpyCas9 (нацеленную на ранее подтвержденный участок с PAM NGG) и Nme1Cas9 (нацеленную на N<sub>4</sub>GATT).

На фиг.7C представлен иллюстративный эффект длины спейсера на эффективность редактирования Nme2Cas9. Нацеливание sg-РНК на единственный участок TLR2.0 с длиной спейсера, варьирующей от 24 до 20 нуклеотидов (включая 5'-концевой G, который требуется для промотора U6), показало, что наиболее высокая эффективность редактирования достигается посредством спейсеров из 22-24 нуклеотидов.

На фиг.7D показано, что иллюстративную двойную никазу Nme2Cas9 можно использовать в тандеме для обеспечения редактирования на основе NHEJ и HDR в TLR2.0. Плазмиды, экспрессирующие Nme2Cas9 и sg-РНК вместе с дцДНК-донором из 800 п.н. для гомологичной репарации вводили посредством электропорации в клетки HEK293T TLR2.0, и исходы как NHEJ (mCherry+), так и HDR (GFP+) оценивали посредством проточной цитометрии. Никаза HNH, Nme2Cas9<sup>D16A</sup>, никаза RuvC, Nme2Cas9<sup>H588A</sup>. На участки расщепления на расстоянии 32 п.н. и 64 п.н. осуществляли нацеливание с использованием каждой никазы. Никаза HNH (Nme2Cas9<sup>D16A</sup>) обеспечивала

эффективное редактирование, в частности, с участками расщепления, которые находятся на расстоянии 32 п.н., в то время как никаза RuvC (Nme2Cas9<sup>H588A</sup>) не была эффективной. В качестве контроля использовали Nme2Cas9 дикого типа.

На фиг.8 представлены иллюстративные данные, демонстрирующие требования к PAM, спейсеру и затравке для нацеливания Nme2Cas9 в клетках млекопитающих в соответствии с фиг.7. Все эксперименты проводили в трех повторениях и планки погрешности соответствуют s.e.m.

На фиг.8A представлено иллюстративное нацеливание Nme2Cas9 на участки N<sub>4</sub>CD в TLR2.0 с оценкой редактирования на основе mCherry+ клеток. Оценивали четыре участка для каждого нуклеотида, не являющегося C, в тестируемом положении (N<sub>4</sub>CA, N<sub>4</sub>CT и N<sub>4</sub>CG) и в качестве положительного контроля использовали участок N<sub>4</sub>CC.

На фиг.8B представлено иллюстративное нацеливание Nme2Cas9 на участки N<sub>4</sub>DC в TLR2.0 [аналогично (A)].

На фиг.8C представлены иллюстративные укорочения гена в участке TLR2.0 (отличном от участка на фиг.2C) с PAM N<sub>4</sub>CCA, демонстрирующие сходные требования к длине с требованиями, наблюдаемыми для другого участка.

На фиг.8D показано, что иллюстративная эффективность нацеливания Nme2Cas9 является по-разному чувствительной к однонуклеотидным несоответствиям в области затравки sg-PHK. Данные демонстрируют эффекты перемещающихся однонуклеотидных несоответствий sg-PHK вдоль спейсера из 23 нуклеотидов в участке-мишени TLR2.0.

На фиг.9 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 в эндогенных локусах в клетках млекопитающих посредством множества способов доставки. Все результаты соответствуют 3 независимым биологическим повторениям и планки погрешности соответствуют s.e.m.

На фиг.9A представлено иллюстративное редактирование генома посредством Nme2Cas9 для эндогенных участков человека в клетках HEK293T после временной трансфекции плазмид, экспрессирующих Nme2Cas9 и sg-PHK. Сначала проводили скрининг 40 участков (таблица 1); затем 14 представленных участков (выбранных для включения репрезентативных примеров различной эффективности редактирования при определении посредством TIDE) повторно анализировали в трех повторениях. В качестве отрицательного контроля использовали участок-мишень Nme1Cas9 (с PAM N<sub>4</sub>GATT).

На фиг.9B представлены иллюстративные карты данных: левая панель: временная трансфекция единичной плазмиды, экспрессирующей как Nme2Cas9, так и sg-PHK (нацеленной на локусы *Pcsk9* и *Rosa26*), позволяет редактирование в клетках мыши Нера1-6, определяемое посредством TIDE. Правая панель: введение посредством электропорации плазмид sg-PHK в клетки K562, стабильно экспрессирующие Nme2Cas9 из лентивирусного вектора, приводит к эффективному образованию инсерции-делеции.

На фиг.9C показано, что иллюстративную Nme2Cas9 можно вводить посредством электропорации в качестве RNP-комплекса для индукции редактирования генома. 40

пикомоль Cas9 вместе с 50 пикомоль транскрибированных *in vitro* sg-PHK, нацеленных на различные локусы, вводили посредством электропорации в клетки HEK293T. Инсерции-делеции определяли через 72 ч с использованием TIDE.

На фиг.10 представлены иллюстративные данные, демонстрирующие зависимость от дозы и сегментные делеции посредством Nme2Cas9, связанные с фиг.9.

На фиг.10А показано, что иллюстративное увеличение дозы введенной посредством электропорации плазмиды Nme2Cas9 (500 нг против 200 нг на фиг.3А) повышает эффективность редактирования в двух участках (TS16 и TS6). Данные, указанные желтым цветом, использовали повторно из фиг.9А.

На фиг.10В показано, что иллюстративную Nme2Cas9 можно использовать для создания точных сегментных делеций. На две мишени TLR2.0 с участками расщепления на расстоянии 32 п.н. проводили одновременное нацеливание посредством Nme2Cas9. Большинство внесенных повреждений представляли собой делеции точно 32 п.н. (синий).

На фиг.11 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 подвергается ингибированию посредством подгруппы семейств белков анти-CRISPR типа II-C *in vitro* и в клетках. Все эксперименты проводили в трех повторениях и планки погрешности соответствуют s.e.m.

На фиг.11А представлен иллюстративный анализ расщепления *in vitro* для Nme1Cas9 и Nme2Cas9 в присутствии пяти ранее охарактеризованных белков анти-CRISPR (соотношение 10:1 Acr:Cas9). Сверху: Nme1Cas9 эффективно расщепляет фрагмент, содержащий протоспейсер с PAM N<sub>4</sub>GATT в отсутствие Acr или в присутствии отрицательного контрольного Acr (AcrE2). Все пять ранее охарактеризованных семейств Acr типа II-C ингибировали Nme1Cas9, как и ожидалось. Снизу: ингибирование Nme2Cas9 отражает ингибирование Nme1Cas9, за исключением отсутствия ингибирования посредством AcrIIС5<sub>Snu</sub>.

На фиг.11В представлено иллюстративное редактирование генома в присутствии пяти ранее описанных семейств белков анти-CRISPR. Плазмиды, экспрессирующие Nme2Cas9 (200 нг), sg-PHK (100 нг) и каждый соответствующий Acr (200 нг) совместно трансфицировали в клетки HEK293T и редактирование генома определяли с использованием отслеживания инсерций-делеций по распаду (Tracking of Indels by Decomposition (TIDE)) через 72 ч после трансфекции. В соответствии с проведенным автором изобретения анализом *in vitro*, все белки анти-CRISPR типа II-C за исключением AcrIIС5<sub>Snu</sub> ингибировали редактирование генома, хотя и с различной эффективностью.

На фиг.11С показано, что иллюстративное ингибирование Nme2Cas9 посредством Acr является зависимым от дозы с двумя различными кажущимися эффективностями. Nme2Cas9 полностью ингибировалась посредством AcrIIС1<sub>Nme</sub> и AcrIIС4<sub>Hpa</sub> в соотношениях по массе 2:1 и 1:1 совместно трансфицированных плазмид Acr и Nme2Cas9, соответственно.

На фиг.12 представлены иллюстративные данные, демонстрирующие, что замена

PID Nme2Cas9 делает Nme1Cas9 нечувствительной к ингибированию AcrPC5<sub>Smu</sub>, связанные с фиг. 11. Показано расщепление *in vitro* посредством химеры Nme1Cas9-Nme2Cas9PID в присутствии ранее охарактеризованных белков Acr (10 мкМ Cas9-sgRNA+100 мкМ Acr).

На фиг.13 представлены иллюстративные данные, демонстрирующие ортогональность и относительную точность Nme2Cas9 и SpyCas9 в отношении двойных участков-мишеней, связанные с фиг.12.

На фиг.13А показано, что иллюстративные гиды Nme2Cas9 и SpyCas9 являются ортогональными. Результаты TIDE показывают частоты инсерций-делеций, внесенных обеими нуклеазами, нацеленными на DS2, либо с их собственными sg-РНК, либо с sg-РНК другого ортолога.

На фиг.13В представлены иллюстративные Nme2Cas9 и SpyCas9, демонстрирующие сравнимую эффективность редактирования в мишени при оценке посредством GUIDE-seq. Столбики указывают на количества чтений с редактированием в мишени согласно GUIDE-Seq в трех двойных участках, на которые нацелен каждый ортолог. Оранжевые столбики соответствуют Nme2Cas9 и черные столбики соответствуют SpyCas9.

На фиг.13С представлено иллюстративное количество чтений с редактированием в мишени против редактирования вне мишени для SpyCas9 в каждом участке. Оранжевые столбики соответствуют чтениям с редактированием в мишени, в то время как черные столбики соответствуют редактированию вне мишени.

На фиг.13D представлено иллюстративные чтения с редактированием в мишени против редактирования вне мишени для Nme2Cas9 в каждом участке.

На фиг.13Е представлены столбиковые диаграммы, демонстрирующие иллюстративную эффективность инсерций-делеций (при измерении посредством TIDE) в потенциальных участках вне мишени, спрогнозированных посредством CRISPRSeek. Последовательности участков в мишени и вне мишени представлены слева, причем область PAM подчеркнута и несоответствия sg-РНК и неконсенсусные нуклеотиды PAM указаны красным цветом.

На фиг.14 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 демонстрирует небольшое или не демонстрирует поддающегося обнаружению нацеливания вне мишени в клетках млекопитающих.

На фиг.14А представлена иллюстративная схема, изображающая двойные участки (DS), на которые нацеливается как SpyCas9, так и Nme2Cas9, посредством их не перекрывающихся PAM. PAM (оранжевый) Nme2Cas9 и PAM (синий) SpyCas9 выделены. Гидовая последовательность Nme2Cas9 из 24 нуклеотидов указана желтым; соответствующая гидовая последовательность для SpyCas9 на 4 нуклеотида короче на 5'-конце.

На фиг.14В представлены иллюстративные Nme2Cas9 и SpyCas9, обе из которых

индуцируют инсерции-делеции в DS. Шесть DS в *VEGFA* (с последовательностями GN<sub>3</sub>GN<sub>19</sub>NGGNCC) было отобрано для прямых сравнений редактирования посредством двух ортологов. Плазмиды, экспрессирующие каждую Cas9 (с тем же промотором, линкерами, метками и NLS) и ее собственный гид трансфицировали в клетки HEK293T. Эффективность инсерций-делеций определяли посредством TIDE через 72 часа после трансфекции. Редактирование Nme2Cas9 выявлялось во всех шести участках и было незначительно или значительно более эффективным, чем у SpyCas9 в двух участках (DS2 и DS6, соответственно). SpyCas9 редактировала четыре из шести участков (DS1, DS2, DS4 и DS6), причем два участка демонстрировали значительно более высокую эффективность редактирования, чем Nme2Cas9 (DS1 и DS4). DS2, DS4 и DS6 были отобраны для анализа GUIDE-Seq, поскольку Nme2Cas9 была в равной степени эффективной, менее эффективной и более эффективной, чем SpyCas9, соответственно, в этих участках.

На фиг.14C представлено иллюстративное редактирование генома посредством Nme2Cas9, которое является в высокой степени точным в клетках человека. Показаны количества участков вне мишени, обнаруженных посредством GUIDE-Seq для каждой нуклеазы в индивидуальных участках-мишенях. В дополнение к двойным участкам, авторы изобретения проанализировали TS6 (вследствие его высокой эффективности редактирования в мишени) и участки *Pcsk9* и *Rosa26* в клетках Нера1-6 мыши (для измерения точности в другом типе клеток).

На фиг.14D показано, что иллюстративное направленное глубокое секвенирование для детекции инсерций-делеций в отредактированных клетках подтверждает высокую точность Nme2Cas9, показанную посредством GUIDE-seq.

На фиг.14E представлена иллюстративная последовательность подтвержденного участка вне мишени для гена *Rosa26*, демонстрирующая область PAM (подчеркнута), консенсусный динуклеотид CC PAM (полужирным шрифтом) и три несоответствия оснований в дистальной от PAM части спейсера (красный).

На фиг.15 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 *in vivo* посредством доставки с использованием AAV "все в одном".

На фиг.15A представлена иллюстративная схема процесса для доставки AAV8.sgRNA.Nme2Cas9 для снижения уровней холестерина у мышей посредством нацеливания на *Pcsk9*. Сверху: схема вектора AAV "все в одном", экспрессирующего Nme2Cas9 и sg-РНК (индивидуальные геномные элементы не в масштабе). BGH, участок поли(А) бычьего гормона роста; HA, эпитопная метка; NLS, последовательность ядерной локализации; h, кодон-оптимизированный для человека. Снизу: расписание инъекций в хвостовую вену AAV8.sgRNA.Nme2Cas9 ( $4 \times 10^{11}$  GC) с последующим измерением уровня холестерина на 14 сутки и анализом инсерций-делеций, гистологии и холестерина на 28 сутки после инъекции.

На фиг.15B представлен иллюстративный анализ TIDE для определения инсерций-

делеций в ДНК, экстрагированной из печени мышей, которым инъецировали AAV8.Nme2Cas9+sg-PHK, нацеленную на локусы *Pcsk9* и *Rosa26* (контроль). Также посредством TIDE оценивали эффективность инсерций-делеций в единственном участке вне мишени, идентифицированном посредством GUIDE-seq для двух sg-PHK (*Rosa26|OT1*).

На фиг.15С представлены иллюстративные снижение сывороточных уровней холестерина у мышей, которым инъецировали нацеленный на *Pcsk9* гид, по сравнению с контролями, нацеленными на *Rosa26*. Значения *P* вычислены с использованием непарного двухстороннего t-критерия.

На фиг.16 представлены иллюстративные данные, демонстрирующие нокдаун PCSK9 и гистологию печени после доставки посредством AAV Nme2Cas9 и редактирования, связанные с фиг.15.

На фиг.16А показано, что иллюстративный вестерн-блоттинг с использованием антитело против PCSK9 демонстрирует значительно сниженные уровни PCSK9 в печени мышей, которым вводили *sgPcsk9*, по сравнению мышами, которым вводили *sgRosa26*. В качестве стандарта подвижности использовали 2 нг рекомбинантной PCSK9 (дорожка слева), и полоса перекрестного реагирования в образцах печени указана звездочкой. В качестве нагрузочного контроля использовали GAPDH (нижняя панель).

На фиг.16В представлено иллюстративное окрашивание посредством H&E образцов печени мышей, которым инъецировали векторы AAV8.Nme2Cas9+*sgRosa26* (слева) или AAV8.Nme2Cas9+*sgPcsk9* (справа). Масштабные метки, 25 мкм.

На фиг.17 представлены иллюстративные данные, демонстрирующие редактирование *Tyr ex vivo* в зиготах мышей, связанные с фиг.16.

На фиг.17А представлены иллюстративные два участка в *Tyr*, каждый с PAM N<sub>4</sub>CC, которые тестировали в отношении редактирования в клетках Нера1-6. Гид *sgTyr2* продемонстрировал более высокую эффективность редактирования и был выбран для дальнейшего тестирования.

На фиг.17В представлены иллюстративные семь мышей, которые выжили в ходе постнатального развития и которые имели фенотипы окраса шерсти, а также редактирование в мишени при анализе посредством TIDE.

На фиг.17С представлены иллюстративные спектры инсерций-делеций из ДНК хвоста каждой мыши из (В), а также мыши C57BL/6NJ без редактирования, как показано посредством анализа TIDE. Указана эффективность инсерций (положительные) и делеций (отрицательные) различных размеров.

На фиг.18 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 *ex vivo* посредством доставки с использованием AAV "все в одном".

На фиг.18А представлена иллюстративная схема процесса для редактирования Nme2Cas9 посредством одного AAV *ex vivo* для получения мышей-альбиносов C57BL/6NJ

посредством нацеливания на ген *Tyr*. Зиготы культивировали в KSOM, содержащей AAV6.Nme2Cas9:sg*Tyr*, в течение 5-6 часов, ополаскивали M2 и культивировали в течение суток перед переносом в яйцевод псевдобеременных реципиентов.

На фиг.18В представлены иллюстративные мышцы-альбиносы (слева) и шиншилловые или пятнистые (середина) мышцы, полученные посредством  $3 \times 10^9$  GC, и шиншилловые или пятнистые мышцы (справа), полученные посредством  $3 \times 10^8$  GC, для зигот с AAV6.Nme2Cas9:sg*Tyr*.

На фиг.18С представлено иллюстративное обобщение экспериментов по редактированию *Tyr* посредством одного AAV Nme2Cas9:sg*Tyr* *ex vivo* в двух дозах AAV.

На фиг.19 представлен иллюстративный репортерный анализ mCherry для nSpCas9-ABEmax и ABEmax-nNme2Cas9 с оптимизированной активностью (D16A).

На фиг.19А представлена иллюстративная информация о последовательности репортера ABE-mCherry. В кодирующей области mCherry присутствует стоп-кодон TAG. В стабильной клеточной линии с встроенным репортером отсутствует сигнал mCherry. Сигнал mCherry обнаруживается, если nSpCas9-ABEmax или оптимизированная конструкция **ABEmax-nNme2Cas9 (D16A)** может конвертировать TAG в CAG (который кодирует Gln).

На фиг.19В представлено зажигание иллюстративных сигналов mCherry, поскольку SpCas9-ABE или **ABEmax-nNme2Cas9 (D16A)** является активным в конкретной области репортера mCherry. Верхняя панель соответствует отрицательному контролю, средняя панель соответствует сигналам mCherry, загоревшимся в репортерных клетках обработанных nSpCas9-ABEmax, на нижней панели представлены сигналы mCherry, загоревшиеся в репортерных клетках, обработанных оптимизированной конструкцией **ABEmax-nNme2Cas9 (D16A)**.

На фиг.19С представлено иллюстративное количественное определение посредством FACS событий редактирования оснований в репортерных клетках mCherry, трансфицированных SpCas9-ABE или **ABEmax-nNme2Cas9 (D16A)**. N = 6; планки погрешности соответствуют S.D. Результаты получены для биологических повторений, выполненных в виде технических дублей.

На фиг.20 представлен иллюстративный репортерный анализ с GFP для активности nSpCas9-CBE4 (Addgene #100802) и **CBE4-nNme2Cas9 (D16A)-UGI-UGI** (CBE4 был клонирован из Addgene #100802).

На фиг.20А представлена иллюстративная информация о последовательности для репортера CBE-GFP. В центральной области флуорофора линии с репортером GFP существует мутация, которая конвертирует GYG в GHG. Таким образом, отсутствует сигнал GFP. Сигнал GFP появляется, если nSpCas9-CBE4 или **CBE4-nNme2Cas9 (D16A)-UGI-UGI** может конвертировать CAC в TAC/TAT (гистидин в тирозин).

На фиг.20В представлен иллюстративный сигнал GFP (зеленый), поскольку nSpCas9-CBE4 или **CBE4-nNme2Cas9(D16A)-UGI-UGI** является активной в конкретной

области репортера GFP. Верхняя панель соответствует отрицательному контролю. На средней панели показано, что сигналы mCherry загораются в репортерных клетках, обработанных **CBE4-nNme2Cas9 (D16A)-UGI-UGI**. На нижней панели показано, что сигналы GFP загораются в репортерных клетках, обработанных **CBE4-nNme2Cas9 (D16A)-UGI-UGI**.

На фиг.20С представлено иллюстративное количественное определение посредством FACS событий редактирования оснований в репортерных клетках с GFP, трансфицированных nSpCas9-CBE4 или **CBE4-nNme2Cas9(D16A)-UGI-UGI**. N = 6; планки погрешности соответствуют S.D. Результаты получены для биологических повторений, выполненных в виде технических дублей.

На фиг.21 представлено иллюстративное редактирование цитозина посредством **CBE4-nNme2Cas9(D16A)-UGI-UGI**. На верхней панели представлена информация о последовательности-мишени KANK3 (последовательности PAM подчеркнуты красным) для Nme2Cas9 и редактирование оснований в образцах отрицательного контроля. На нижней панели представлено количественное определение частоты замены каждого типа оснований в окне редактирования **CBE4-nNme2Cas9(D16A)-UGI-UGI** для последовательностей-мишеней KANK3. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования С в Т выделены красным цветом.

На фиг.22 представлено иллюстративное редактирование цитозина и аденина посредством **CBE4-nNme2Cas9 (D16A)-UGI-UGI** и оптимизированного **ABEmax-nNme2Cas9 (D16A)**, соответственно. На верхней панели представлена информация о последовательности-мишени PLXNB2 (последовательности PAM подчеркнуты красным) для Nme2Cas9 и редактирование оснований в образцах отрицательного контроля. На средней панели представлено количественное определение частоты замены каждого типа оснований в окнах редактирования оптимизированного **ABEmax-nNme2Cas9 (D16A)** для последовательностей-мишеней PLXNB2. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования А в G выделены красным цветом. На нижней панели представлено количественное определение частоты замены каждого типа оснований в окне редактирования **CBE4-nNme2Cas9 (D16A)-UGI-UGI** для последовательностей-мишеней PLXNB2. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования С в Т выделены красным цветом.

### **Подробное описание изобретения**

Настоящее изобретение относится к области редактирования генов. В частности, редактирование генов направлено на редактирование единичных нуклеотидных оснований. Например, такое редактирование единичных нуклеотидных оснований приводит к преобразованию пары оснований С-Г в пару оснований Т-А. Высокая точность и аккуратность описанного в настоящем описании однонуклеотидного редактора

достигается посредством нуклеазы NmeCas9, которая слита с белком нуклеотиддезаминазой. Компактная природа NmeCas9, сопряженной с более высоким количеством совместимых прилегающих к протоспейсеру мотивов, обеспечивает наличие у слитых конструкций Cas9, описанных в настоящем описании, возможности редактировать участки, на которые не нацеливаются другие общепринятые редактирующие основания платформы SpyCas9.

#### **А. Редактирование единичных оснований посредством NmeCas9**

Cas9 представляет собой программируемую нуклеазу, которая использует гидовую РНК для внесения двухцепочечного разрыва в любом требуемом геномном локусе. Эта программируемость используется для биомедицинских и терапевтических подходов. Однако индуцируемые Cas9 разрывы часто приводят к неточной репарации посредством клеточного аппарата, препятствуя ее терапевтическому применению для коррекций единичных оснований, а также единообразных и точных нокаутов генов. Более того, является чрезвычайно проблематичным комбинирование индуцируемых Cas9 двухцепочечных разрывов ДНК и матрицы репарации для направляемой гомологией репарации (HDR) в целях коррекции генетических мутаций в постмитотических клетках (например, нейрональных клетках).

Редактирование единичных нуклеотидных оснований представляет собой подход к редактированию генома, где Cas9 с неработающей или поврежденной нуклеазой (например, мертвая Cas9 (dCas9) или никаза Cas9 (nCas9)) слита с другим ферментом, способным редактировать нуклеотидные основания, не вызывая двухцепочечные разрывы ДНК. На сегодняшний день было разработано два широких класса редакторов оснований Cas9: i) слитый с цитидиндезаминазой (редактирует пару оснований C-G в пару оснований T-A) белок SpyCas9; и ii) SpyCas9 с аденозиндезаминазой (редактирует пару оснований A-T в пару оснований G-C). Liu et al., "Nucleobase editors and uses thereof" US 2017/0121693; и Lui et al., "Fusions of cas9 domains and nucleic acid-editing domains" US 2015/0166980; (обе из которых включены в настоящее описание в качестве ссылок).

Однако, как упоминалось выше, платформы редактирования оснований SpyCas9 не могут использоваться для нацеливания на все мутации единичных оснований вследствие их ограниченных окон редактирования. Окно редактирования ограничивается требованием PAM NGG. SpyCas9 также по своей природе ассоциирована с высокими эффектами вне мишени при редактировании генома.

В одном варианте осуществления настоящее изобретение относится к слитому белку деаминазы с компактной и сверхточной Nme2Cas9 (*Neisseria meningitidis* spp.). Эта Nme2Cas9 имеет 1082 аминокислоты по сравнению с SpyCas9, которая имеет 1368 аминокислот. Этот ортолог Nme2Cas9 функционирует эффективно в клетках млекопитающих, распознает PAM N<sub>4</sub>CC и по своей природе является сверхточной. Edraki et al., Mol Cell. (подготавливается).

Хотя не является необходимым понимание механизма изобретения, полагают, что

компактность и сверхточность редактора оснований NmeCas9 нацелена на мутации единичных оснований, на которые не могло быть осуществлено нацеливание ранее посредством других платформ Cas9, в настоящее время известных в данной области. Кроме того, полагают, что редакторы оснований NmeCas9, предусматриваемые в рамках настоящего изобретения, нацелены на патогенные мутации, нацеливание на которых не осуществимо посредством современных редактирующих платформ, и с увеличенной точностью редактирования оснований.

В одном варианте осуществления настоящее изобретение охватывает слитый белок, содержащий Nme2Cas9 и белок-дезаминазу, иллюстративные примеры которого включают ABE7.10-nNme2Cas9 (D16A); оптимизированный nNme2Cas9-ABEmax; nNme2Cas9-CBE4 (равно BE4-nNme2Cas9 (D16A)-UGI-UGI), а также ABEmax-nNme2Cas9 (D16A). См., фиг. 1A, фиг. 1B, фиг. 1C, фиг. 1D и фиг. 1E.

На фиг. 1 проиллюстрированы схематичные варианты осуществления редактора единичных оснований на основе слитого белка NmeCas9-дезаминазы и иллюстративных сконструированных плазмид редакторов оснований. На фиг. 1A представлена иллюстративная конструкция YE1-BE3-nNme2Cas9 (D16A)-UGI. На фиг. 1B представлена иллюстративная конструкция ABE7.10 nNme2Cas9 (D16A). На фиг. 1C представлена иллюстративная конструкция ABE7.10-nNme2Cas9 (D16A). На фиг. 1C представлена иллюстративная конструкция ABE7.10-nNme2Cas9 (D16A), содержащая две последовательности NLS SV40. На фиг. 1D представлена иллюстративная конструкция nNme2Cas9-CBE4 (также называемая BE4-nNme2Cas9 (D16A)-UGI-UGI). На фиг. 1E представлена иллюстративная оптимизированная конструкция nNme2Cas9-ABEmax.

В одном варианте осуществления белок дезаминаза представляет собой Arobec1 (YE1-BE3). Предполагается, что Arobec1 не ограничивается одним организмом. В одном варианте осуществления Arobec1 происходит из вида крыс. Kim et al., "Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions". *Nature Biotechnology* 35 (2017). В одном варианте осуществления Nme2Cas9 включает мутант nNme2Cas9 D16A. В одном варианте осуществления слитый белок дополнительно содержит белок-ингибитор урацилгликозилазы (UGI). В одном варианте осуществления слитый белок включает конструкцию YE1-BE3-nNme2Cas9 (D16A)-UGI. В одном варианте осуществления конструкция YE1-BE3-nNme2Cas9 (D16A)-UGI имеет последовательность:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTKETCLLYEINWGGRRHSIWRHTSQ  
NTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGECSRAITEFLSRYPHVTLFIYIAR  
LYHHADPENRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPTSNEAHWPRYPHLWVR  
LYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTS  
ESATPESMAAFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRFERAEVPKTGDS  
LAMARRLARSVRRLTRRRAHRLLRARLLKREGVLQAADFENGLIKSLPNTPWQLRAAALDR  
KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELAL

NKFEKESGHIRNQRGDYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGGLKEGIETLLMTQRP  
 ALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQGSERPLTDTERATLMD  
 EPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEKAYHAISRALEKEGLKDKK  
 SPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLALRRIVPL  
 MEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGVRRYRYS  
 ARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQ  
 QHGKCLYSGKEINLVRLENEKGYVEIDAALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYF  
 NGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDEDGFKECNLNDTRYVNRFLCQFVADHI  
 LLTGKGRVRFASNGQITNLLRGFWGLRKYRAENDRRHALDAVVACSTVAMQOKITRFVRY  
 KEMNAFDGKTIDKETGKVLHQKTHFPQWFEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRT  
 LLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFRVKNHNEKISVKRVWLTEIKLA  
 DLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKAVRVEKTQES  
 GVLLNKNAYTIADNGDMVRVDVFCVKDKKGNQYFIVPIYAWQVAENILPDIDCKGYRIDD  
 SYTFCFSLHXYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQ  
 KYQVNELGKEIRPCRLKKRPPVRSGGSTNLSDIIEKETGKQLVIOESILMLPEEVEEVIGN  
**KPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKK**  
 KRKV\*

YE1-BE3 (подчеркнут); линкер (полужирным шрифтом), nNme2Cas9 (курсивом),  
 UGI (полужирным шрифтом/подчеркнут), NLS SV40 (простой).

В одном варианте осуществления конструкция YE1-BE3-nNme2Cas9 (D16A)-UGI  
 имеет последовательность:

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKTKETCLLYEINWGGRHSIWRHTSQ  
NTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGECSRAITEFLSRYPHVTLFIYIAR  
LYHHADPENRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPTSNEAHWPRYPHLWVR  
LYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSETPGTS  
**ESATPESMAAFKPNPINYILGLAIGIASVGMAMVEIDEEENPIRLIDLGVRFERAEVPKTGDS**  
**LAMARRLARSVRLTRRRARHLLRARRLLKREGVLQAADFDEGLIKSLPNTPWQLRAALDR**  
**KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELAL**  
 NKFEKESGHIRNQRGDYSHTFSRKDLQAEILLLFEKQKEFGNPHVSGGLKEGIETLLMTQRP  
 ALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQGSERPLTDTERATLMD  
 EPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEKAYHAISRALEKEGLKDKK  
 SPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLALRRIVPL  
 MEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGVRRYRYS  
 ARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQ  
 QHGKCLYSGKEINLVRLENEKGYVEIDAALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYF  
 NGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDEDGFKECNLNDTRYVNRFLCQFVADHI  
 LLTGKGRVRFASNGQITNLLRGFWGLRKYRAENDRRHALDAVVACSTVAMQOKITRFVRY  
 KEMNAFDGKTIDKETGKVLHQKTHFPQWFEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRT  
 LLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFRVKNHNEKISVKRVWLTEIKLA

*DLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKA VRVEKTQES  
GVLNKNAYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDD  
SYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLA WHDKGSKEQQFRISTQNLVLIQ  
KYQVNELGKEIRPCRLKKRPPVRS **GGSTNLSDIIEKETGKQLVIOESILMLPEEVEEVIG  
NKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLS**GGSPK  
KKRKV\**

YE1-BE3 (подчеркнут); линкер (полужирным шрифтом), nNme2Cas9 (курсивом), UGI (полужирным шрифтом/подчеркнут), NLS SV40 (простой).

В одном варианте осуществления настоящее изобретение относится к слитому белку, содержащему белок дезаминазу NmeCas9/ABE7.10. В одном варианте осуществления белок дезаминаза представляет собой TadA. В одном варианте осуществления белок дезаминаза представляет собой TadA 7.10. В одном варианте осуществления конструкция ABE7.10-nNme2Cas9 (D16A) имеет следующую последовательность:

***MSEVEFSHEYWMRHALTLAKRAWDEREVPVGA VL VHNNRVIGEGWNRPIGRH  
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK  
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS  
GGSSGGSSGSETPGTSESATPESSGGSSGGSS** **SEVEFSHEYWMRHALTLAKRARDEREV  
PVGA VLVLNNRVIGEGWNR****RAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT  
FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA  
DECAALLCYFFRM****PROVFNAQKKAQSSTDS****GGSSGGSSGSETPGTSESATPESSGGSSG  
GSMAAFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRFERA EVPKTGDSLAMAR  
RLARSVRRLTRRRARLLRARRLLKREGVLQAADFENGLIKSLPNTPWQLRAAALDRKLTPL  
EWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKF EK  
ESGHIRNQRGDYSHTF SRKDLQAE LILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDA  
VQKMLGHCTFEP AEPKAAKN TYTAERFIWLT KLNLRILEQGSERPLTDTERATLMDEPYRKS  
KLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSS  
ELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDK FVQISLKALRRIVPLMEQGKR  
YDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVL RALSQARKVINGVVRRYGSPARIHIET  
AREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQH GKCL  
YSGKEINLVR LNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNS  
REWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGKG  
KRRVFASNGQITNLLRGFWGLRKVRAENDRH HALDAVVVACSTVAMQQKITRFVRYKEMNAF  
DGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKL  
SSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRS AKRFVKHNEKISVKRVWLTEIKLADLENM  
VNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKA VRVEKTQESGVLNKN  
KNAYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDD SYTFCFS  
LHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLA WHDKGSKEQQFRISTQNLVLIQKYQVNE  
LGKEIRPCRLKKRPPVREDKRPAATKKAGQAKKKK\****

TadA (подчеркнут), TadA 7.10 (подчеркнут/полужирным шрифтом), линкер (полужирным шрифтом), nNme2Cas9 (курсивом), нуклеоплазмин NLS (простой).

В одном варианте осуществления конструкция ABE7.10-nNme2Cas9 (D16A) имеет следующую аминокислотную последовательность:

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH  
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAK  
TGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD  
SGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREV  
PVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT  
FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA  
DECAALLCYFFRMPRQVFNQAQKKAQSSTDSGSSGGSSGSETPGTSESATPESSGGSS  
GGSMAAFKPNPINYILGLAIGIASVWAMVEIDEEENPIRLIDLGVRFERAEVPKTGDSLAMA  
RRLARSVRRLLTRRAHRLLRARLLKREGVLQAADFDEGLIKSLPNTPWQLRAAALDRKLT  
LEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFE  
KESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGD  
AVQKMLGHCTFEPAEPKAAKNYTAERFIWLTKLNNLRILEQGSRPLDTERATLMDEPYRK  
SKLYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEKAYHAISRALKEGLKDKKSPLNLS  
SELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKALRRIVPLMEQGKR  
YDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGVRRYGGSPARIHIET  
AREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRYEQQHKGKCL  
YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNS  
REWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGKG  
KRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQOKITRFVRYKEMNAF  
DGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKL  
SSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFBVKHNEKISVKRVWLTEIKLADLENM  
VNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGQVLKAVRVEKTQESGVLLNK  
KNAYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFDFS  
LHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNE  
LGKEIRPCRLKKRPPVREDKRPAATKKAGQAKKKK\*

TadA (подчеркнут), TadA 7.10 (подчеркнут/полужирным шрифтом), линкер (полужирным шрифтом), nNme2Cas9 (курсивом), нуклеоплазмин NLS (простой).

В одном варианте осуществления конструкция ABEmax-nNme2Cas9 (D16A) имеет следующую аминокислотную последовательность:

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVL  
VHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAG  
AMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFR  
MRRQEIKAQKKAQSSTDSGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYW  
MRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQ  
GGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMD

**VLHYPGMNRVEITEGILADECAALLCYFFRMPROVFNAQKKAQSSTDSGGSSGGS**  
**SGSETPGTSESATPESSGGSSGGSMAAFKPNPINYILGLAIGIASVWAMVEIDEEENPIRLID**  
 LGVRFERAIEVPTGDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFDENGLI  
 KSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANN  
 AHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHV  
 SGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILE  
 QGSRPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKA  
 YHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISF  
 DKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRAL  
 SQARKVINGVRRYGGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNF  
 VGEPKSKDILKLRLYEQQHKGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVL  
 GSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSFRPRSKKQORILLQKFDEDEGFKECNLN  
 DTRYVNRFLCQFVADHILLTGKGRRVFASNGQITNLLRGFWGLRKVRAENDRRHALDAVVV  
 ACSTVAMQOKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKP  
 DGKPEFEEADTPEKLRLLAEKLSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFVK  
 HNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYK  
 KGGQLVKAVRVEKTQESGVLLNKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQ  
 VAENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDK  
 GSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVREDKRPAATKKAGQAKKKK**FE**  
**PKKKRKV\***

TadA (подчеркнут), TadA 7.10 (подчеркнут/полужирным шрифтом), линкер (полужирным шрифтом), nNme2Cas9 (курсивом), нуклеоплазмин NLS (простой) и SV40 NLS (полужирным шрифтом).

В одном варианте осуществления конструкция **CBE4-nNme2Cas9 (D16A)-UGI-UGI** имеет следующую аминокислотную последовательность:

**PAAKRVKLDGGSGGGSGGGSGPAAKRVKLDGGSGGGSGGGSGPLEPKKKRKV**  
**PWSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCLLYEINWGRHSIWRHTSQNTN**  
**KHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYH**  
**HADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPRYPHLWVRLYV**  
**LELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGGSSGGSSGSE**  
**TPGTSESATPESSGGSSGGSIDKLA**AFKPNPINYILGLAIGIASVWAMVEIDEEENPIRLIDL  
 GVRVFERAEVPTGDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFDENGLIK  
 SLPNTPPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNA  
 HALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVS  
 GGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQ  
 GSRPLTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAY  
 HAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFD  
 KVVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALS  
 QARKVINGVRRYGGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFV

GEPKSKDILKLRLYEQQHKGKCLYSGKEINLVRLENEKGYVEIDHALPFSRTWDDSFNNKVLVLG  
 SENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEEDGFKECNLND  
 TRYVNRFLCQFVADHILLTGKGRVRFASNGQITNLLRGFWGLRKYRAENDRRHALDAVVVA  
 CSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPD  
 GKPEFEEADTPEKLRLLAEKLSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRPFVKH  
 NEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKK  
 GGQLVKAVRVEKTQESGVLNKKNAYTIADNGDMVRVDVFCVKDKKGNQYFIVPIYAWQV  
 AENILPDIDCKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDK  
 GSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVRVYPYDVPDYAGYPYDVPDYAG  
**SYPYDVPDYAGSAAPAAKKKKLDFESGEFLQPGIDLSQLGGDSGGSGGSGGSTNLSDIIE**  
**KETGKQLVIOESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK**  
**PWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDIIEKETGKQLVIOESILMLPEEVEE**  
**VIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGS**  
**PKKKRKVSRGSAAPAAKRVKLDGGSGGGSGGGSGGSGPAAKRVKLD**

гАробес1 (подчеркнут), UGI (подчеркнут/полужирным шрифтом), линкер (полужирным шрифтом, курсивом), nNme2Cas9 (D16A) (курсивом), С-мус-NLS (простой) и SV40 NLS (полужирным шрифтом).

В одном варианте осуществления оптимизированная конструкция nNme2Cas9-ABEmax относится к оптимизированной версии с улучшенным промотором, последовательностями NLS и линкерными последовательностями. В некоторых вариантах осуществления оптимизированная конструкция nNme2Cas9-ABEmax содержит, от 5' к 3', С-мус NLS, линкер из 12 а.к., линкер из 15 а.к., NLS SV40, TadA, TadA\*7.10, линкер из 48 а.к., nNme2Cas9, линкер из 73 а.к. (3×HA-метка), линкер из 15 а.к. и NLS С-мус. В некоторых вариантах осуществления оптимизированная конструкция nNme2Cas9-ABEmax дополнительно содержит по меньшей мере два чередующихся NLS С-мус и линкера из 12 а.к. на 3'-конце. В некоторых вариантах осуществления оптимизированная конструкция nNme2Cas9-ABEmax дополнительно содержит по меньшей мере два чередующихся линкера из 15 а.к. и NLS С-мус на 5'-конце. См., например, фиг.1Е.

В одном варианте осуществления оптимизированная конструкция nNme2Cas9-ABEmax имеет следующую аминокислотную последовательность:

**AKRVKLDGGSGGGSGGGSGPAAKRVKLDGGSGGGSGGGSGPLEPKKKRKVSE**  
**VEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEI**  
**MALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSL**  
**MDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSS**  
**GSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV**  
**LNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC**  
**AGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALL**  
**CYFFRMPROVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMAAF**  
**KPNPINYDIDKLAAFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKT**

GDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFDEGLIKSLPNTPWQLRAAA  
 LDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAE  
 LALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIETLLMTQ  
 RPALSGDAVQKMLGHCTFEPAEPKAAKNYTAERFIWLTCLNLRILEQGSERPLTDTERATL  
 MDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKD  
 KKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKALRRIVP  
 LMEQKGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGVRRYGS  
 PARIHIETAREVGKSFKDRKEIEKRQEEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYE  
 QQHGKCLYSKKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEY  
 FNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEEDGFKECNLNDTRYVNRFLCQFVAD  
 HILLTGKGRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVACSTVAMQOKITRFV  
 RYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKL  
 RTLLAEKLSRPEAVHEYVTPFLVSRAPNRKMSGAHKDTLRSKRFRVKNNEKISVKRVWLTEIK  
 LADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGQVLKAVRVEKTQ  
 ESGVLLNKKNAYTIADNGDMVRVDVFCVKDKKGNQYFIVPIYAWQVAENILPDIDCKGYRID  
 DSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLI  
 QKYQVNELGKEIRPCRLKKRPPVRVYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSAAPA  
**AKKKKLD**FESGEFLQPGGSTSSRGSAAAPAAKRVKLDGGSGGGSGGGSGSGPAAKRVKL  
 D

hTadA7.10 (подчеркнут), hTadA\*7.10 (подчеркнут/полужирным шрифтом), линкер (полужирным шрифтом, курсивом), nNme2Cas9 (курсивом), Cmyc-NLS (простой), SV40-NLS (полужирным шрифтом).

В некоторых вариантах осуществления плазмиду nSpCas9-ABEmax (Addgene ID:112095) использовали в качестве контролей в эксперименте и для молекулярного клонирования. В некоторых вариантах осуществления плазмиду nSpCas9-CBE4 (Addgene ID: 100802) использовали в качестве контролей в эксперименте и для молекулярного клонирования.

Введение посредством электропорации в клетки HEK293Т ДНК-плазмид, содержащий слитый белок YE1-BE3-Nme2Cas9-нуклеотиддезаминаза достигало стабильного редактирования единичного основания пары оснований С-Г на пару оснований Т-А в эндогенном участке-мишени (TS25). См., фиг.2А-С.

На фиг.2 представлены иллюстративные данные электропорации клеток HEK293Т посредством ДНК-плазмид, включающих слитый белок YE1-BE3-nNme2Cas9 (D16A)-UGI, эффективно конвертирующий С на Т в эндогенном участке-мишени 25 (TS25) в клетках HEK293Т посредством нуклеофекции. На фиг.2А представлены иллюстративные последовательности для эндогенного участка-мишени TS25 (в черном прямоугольнике). Sg-РНК GN23 образует пары с цепью ДНК-мишенью, оставляя вытесненную цепь ДНК для редактирования цитидиндезаминазой (например, новые зеленые нуклеотиды). На фиг.2В представлены иллюстративные данные секвенирования, показывающие двойной

нуклеотидный пик (7-положение с 5'-конца; стрелка), указывающий на успешное редактирование единичного основания цитидина на тимидин (например, конвертирование пар оснований С-Г на пару оснований Т-А). На фиг.2С представлено иллюстративное количественное определение данных, представленных на фиг.2В, в форме графика процентного конвертирования С → Т при редактировании единичного основания. Процент С, конвертированного в Т, составляет приблизительно 40% в образце, обработанном редактором оснований и sg-РНК (значение  $p=6,88 \times 10^{-6}$ ). Контроль "без sg-РНК" демонстрирует фоновый шум при секвенировании по методу Сэнгера. Для проведения анализа использовали EditR (Kluesner et al., 2018).

Четыре других слитых белка YE1-BE3-nNme2Cas9/мутант D16A совместно экспрессировали с усиленным зеленым флуоресцентным белком (EGFP) в стабильной происходящей из K562 клеточной линии, экспрессирующей усиленный зеленый флуоресцентный белок (EGFP). Каждый слитый белок YE1-BE3-nNme2Cas9/мутант D16A имел конкретный участок-мишень UGI. См., фиг.3А-D.

Анализ с использованием глубокого секвенирования показал, что YE1-BE3-nNme2Cas9 конвертирует остатки С в остатки Т в каждом из четырех участков-мишеней EGFP. Процент редактирования находится в диапазоне от 0,24% до 2%. Потенциальное окно редактирования оснований имеет нуклеотиды 2-8 в вытесненной цепи ДНК, считая нуклеотид на 5'-конце (дистальном относительно РАМ) за нуклеотид #1. См., фиг.3А-D.

На фиг.3 представлены иллюстративные определенные участки-мишени UGI, которые соответственно интегрировали в слитые белки YE1-BE3-nNme2Cas9/мутант D16A и совместно экспрессировали с усиленным зеленым флуоресцентным белком (EGFP) в стабильной происходящей из K562 клеточной линии. Конвертированные основания выделены оранжевым цветом. Фоновые сигналы фильтровали с использованием образцов отрицательного контроля (клетки K562 с введенным посредством нуклеофекции YE1-BE3-nNme2Cas9 без конструкций sg-РНК). РАМ N<sub>4</sub>СС заключены в рамку. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце. На фиг.3А представлен иллюстративный участок 1 EGFP. На фиг.3В представлен иллюстративный участок 2 EGFP. На фиг.3С представлен иллюстративный участок 3 EGFP. На фиг.3D представлен иллюстративный участок 4 EGFP.

Введение посредством электропорации в клетки HEK293Т ДНК-плазмид, содержащих YE1-BE3-nNme2Cas9-промотор *c-fos*, достигало стабильного редактирования единичных оснований пары С-Г на пару Т-А в эндогенных участках-мишенях в промоторе *c-fos* (фиг.3Е). На фиг.3Е представлен анализ посредством глубокого секвенирования, указывающий на то, где YE1-BE3-nNme2Cas9 конвертирует остатки С в остатки Т в эндогенной промоторной области *c-fos*. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце. Конвертированные основания выделены оранжевым или желтым

цветом. Фоновые сигналы отфильтровывали с использованием образцов отрицательного контроля. Наиболее высокий процент редактирования составляет 32,50%. На фиг.3F представлен иллюстративный анализ посредством глубокого секвенирования, указывающий на то, где ABE7.10-nNme2Cas9 или ABEmax (Koblan *et al.*, 2018)-nNme2Cas9 конвертирует остатки А в остатки G в эндогенной промоторной области *c-fos*. Процент от всех чтений, демонстрирующий мутации в участках, на которые нацелен редактор оснований, показан в правом столбце. Конвертированные основания выделены оранжевым или желтым цветом. Фоновые сигналы отфильтровывали с использованием образцов отрицательного контроля. Процент редактирования составляет 0,53% посредством ABE7.10-nNme2Cas9 или 2.33% посредством ABEmax-nNme2Cas9 (D16A).

В одном варианте осуществления настоящее изобретение охватывает экспрессию слитого белка ABE7.10-nNme2Cas9 (D16A) для редактирования оснований. Хотя не является необходимым понимание механизма изобретения, полагают, что редактирование оснований Nme2Cas9 может быть эффективным способом лечения тирозинемии посредством реверсии мутации G на А в гене *Fah* с использованием слитого белка ABE7.10-nNme2Cas9 (D16A).

Мутация G на А (красный) в последнем нуклеotide экзона 8 гена *Fah* вызывает пропускание экзона. Дефицит FАН приводит к накоплению токсинов и тяжелому повреждению печени. Положение PAM SpyCas9 (черная прямоугольная рамка) ниже мутации не является оптимальным для конструирования sg-PHK, поскольку мутация А находится вне эффективного окна редактирования оснований ABE7.10, которое представляет собой 4-7-й нуклеотид на 5'-конце (дистально от PAM) (подчеркнуто) (Gaudelli *et al.*, 2017).

Однако существует два PAM Nme2Cas9 (красная прямоугольная рамка) в нижерасположенных последовательностях, которые потенциально могут корректировать мутацию и обеспечивать реверсию последовательности ДНК в дикий тип посредством ABE7.10-nNme2Cas9 (D16A). См., фиг.4.

На фиг.4 представлено иллюстративное выравнивание гена *Fah* дикого типа с мутантным геном тирозинемии *Fah*, демонстрирующее участок-мишень гена для редактирования единичного основания А-G (положение 9). Соответствующий единичный участок PAM SpyCas9 и двойные участки PAM NmeCas9 указаны для демонстрации субоптимального окна нацеливания относительно участка PAM SpyCas9. Эта фиг. служит в качестве потенциального примера участка, где Nme2Cas9 может преодолевать ограничения существующих редакторов оснований. Кроме того, полагают, что редактор оснований NmeCas9, описанный в настоящем описании, может выполнять точное редактирование оснований, которое не может быть достигнуто посредством общепринятых происходящих из SpyCas9 редакторов оснований вследствие субоптимального окна редактирования оснований относительно доступных PAM поблизости.

Более того, авторы изобретения предусматривают расширение редактирования оснований на модель тирозинемии на мышах для реверсии точечковой мутации G на A посредством вирусных способов доставки с использованием ABEmax-nNme2Cas9 (D16A), где желаемое редактирование не может быть достигнуто посредством происходящих из SpyCas9 редакторов оснований вследствие субоптимального окна редактирования оснований относительно доступных PAM поблизости (например, фиг.4).

### **В. Конструкции NmeCas9: компактные и сверхточные**

Короткие палиндромные повторы, регулярно расположенные кластерами (CRISPR) вместе с CRISPR-ассоциированными (Cas) белками составляют адаптивные иммунные каскады бактерий и архей против фагов и других подвижных генетических элементов (MGE) (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008). В системах CRISPR типа II, РНК CRISPR (cr-РНК) связывается с транс-активирующей cr-РНК (tracr-РНК) и загружается на эффекторный белок Cas9, который расщепляет нуклеиновые кислоты MGE, комплементарные cr-РНК (Garneau et al., 2010; Deltcheva et al., 2011; Saprunauskas et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012). Гибрид cr-РНК:tracr-РНК может быть слит в единую гидовую РНК (sg-РНК) (Jinek et al., 2012). Программируемость эндонуклеаз Cas9 в отношении РНК сделала их мощной платформой редактирования генома в биотехнологии и медицине (Cho et al., 2013; Cong et al., 2013; Hwang et al., 2013; Jiang et al., 2013; Jinek et al., 2013; Mali et al., 2013b).

В дополнение к sg-РНК, распознавание мишени посредством Cas9 обычно ассоциировано с сигнатурой из 1-5 нуклеотидов ниже комплементарной последовательности ДНК, называемой прилегающим к протоспейсеру мотивом (PAM) (Deveau et al., 2008; Mojica et al., 2009). Ортологи Cas9 демонстрируют значительное разнообразие длины и последовательности PAM. Среди ортологов Cas9, которые охарактеризованы, Cas9 *Streptococcus pyogenes* (SpyCas9) используется наиболее широко, частично поскольку она распознает короткий PAM NGG (Jinek et al., 2012) (N обозначает любой нуклеотид), что позволяет высокую плотность поддающихся нацеливанию участков. Тем не менее, относительно большой размер Spy (т.е. 1368 аминокислот) затрудняет упаковку Cas9 (вместе с sg-РНК и промоторами) в один рекомбинантный аденоассоциированный вирус (rAAV). Было показано, что это является недостатком для терапевтических применений, учитывая перспективу, которую продемонстрировали векторы AAV для доставки генов *in vivo* (Keeler et al., 2017). Более того, SpyCas9 и ее РНК-гиды требуют подробной охарактеризации и модификации способами инженерии для минимизации тенденции к редактированию сходных с собственными участками вне мишени. (Bolukbasi et al., 2015b; Tsai and Joung, 2016; Tycko et al., 2016; Chen et al., 2017; Casini et al., 2018; Yin et al., 2018). На сегодняшний день дальнейшие усилия инженерии не преодолели эти ограничения по размеру.

Несколько ортологов Cas9 длиной менее 1100 аминокислот, полученных из различных видов, было валидировано для редактирования генома млекопитающих,

включая штаммы *N. meningitidis* (NmeCas9, 1082 а.к.) (Esvelt et al., 2013; Hou et al., 2013), *Staphylococcus aureus* (SauCas9, 1053 а.к.) (Ran et al., 2015), *Campylobacter jejuni* (CjeCas9, 984 а.к.) (Kim et al., 2017) и *Geobacillus stearothermophilus* (GeoCas9, 1089 а.к.) (Harrington et al., 2017b). NmeCas9, CjeCas9 и GeoCas9 являются репрезентативными примерами Cas9 типа II-C (Mir et al., 2018), большинство из которых имеют размер <1100 а.к. За исключением GeoCas9, каждая из этих более коротких ортологичных последовательностей успешно использовалась для редактирования *in vivo* посредством доставки с использованием AAV "все в одном" (когда один вектор экспрессирует как гид, так и эффектор) (Ran et al., 2015; Kim et al., 2017; Ibraheim et al., 2018, представлена для рассмотрения). Более того, было показано, что NmeCas9 и CjeCas9 естественным образом не подвержены редактированию вне мишени (Lee et al., 2016; Kim et al., 2017; Amrani et al., 2018, представлена для рассмотрения). Однако PAM, которые распознаются компактными Cas9, обычно являются более длинными, чем PAM SpyCas9, что существенно уменьшает количество поддающихся нацеливанию участков в или вблизи данного локуса; например, i) N<sub>4</sub>GAYW/N<sub>4</sub>GYTT/N<sub>4</sub>GTCT для NmeCas9 (Esvelt et al., 2013; Hou et al., 2013; Lee et al., 2016; Amrani et al., 2018); ii) N<sub>2</sub>GRRT для SauCas9 (Ran et al., 2015); iii) N<sub>4</sub>RYAC для CjeCas9 (Kim et al., 2017); и iv) N<sub>4</sub>CRAA/N<sub>4</sub>GMAA для GeoCas9 (Harrington et al., 2017b) (Y=C, T; R=A, G; M=A, C; W=A, T). Подгруппа участков-мишеней меньшего размера является преимущественной для задач в высокой степени точного и аккуратного редактирования, включая, но не ограничиваясь ими: i) редактирование мелких мишеней (например, микроРНК); ii) коррекцию мутаций посредством редактирования оснований, которое изменяет очень узкое окно оснований относительно PAM (Komor et al., 2016; Gaudelli et al., 2017); или iii) точное редактирование посредством направляемой гомологией репарации (HDR), которая является наиболее эффективной, когда изменяемые основания находятся вблизи участка расщепления (Gallagher and Haber, 2018). Вследствие ограничений PAM на многие участки редактирования не может быть осуществлено нацеливание с использованием векторов AAV "все в одном" для доставки *in vivo* даже с этими более короткими белками Cas9. Например, был разработан мутант SauCas9 (SauCas9<sup>KKH</sup>), который имеет сниженные ограничения PAM (N<sub>3</sub>RRT), хотя это повышение диапазона нацеливания часто происходит за счет снижения эффективности редактирования в мишени, и все еще наблюдается редактирование вне мишени (Kleinstiver et al., 2015).

Безопасное и эффективное терапевтическое редактирование генов на основе CRISPR значительно усиливается посредством ортологов и вариантов Cas9, которые являются в высокой степени активными в клетках человека, не подверженными нацеливанию вне мишени, достаточно компактными для доставки посредством AAV "все в одном" и способными к доступу к высокоплотным геномным участкам. В одном варианте осуществления настоящее изобретение охватывает компактную сверхточную Cas9 (Nme2Cas9) из определенного штамма *N. meningitidis*. В одном варианте осуществления

настоящее изобретение охватывает способ доставки посредством единственного AAV Nme2Cas9 и его sg-РНК для проведения эффективного редактирования генома *in vivo* и/или *ex vivo*. Хотя не является необходимым понимание механизма изобретения, полагают, что этот ортолог эффективно функционирует в клетках млекопитающих и распознает PAM N<sub>4</sub>CC, что обеспечивает плотность участков-мишеней, идентичную плотности для SpyCas9 дикого типа (например, в среднем каждые 8 п.о., когда рассматриваются обе цепи ДНК).

### **1. Взаимодействующие с PAM домены и белки анти-CRISPR**

Распознавание PAM посредством ортологов Cas9 происходит в основном посредством взаимодействий белок-ДНК между взаимодействующим с PAM доменом (PID) и нуклеотидами, соседними с протоспейсером (Jiang and Doudna, 2017). Мутации PAM часто позволяют ускользание фага от иммунитета CRISPR типа II (Paez-Espino et al., 2015), помещая эти системы под давление селекции не только для приобретения новых спейсеров CRISPR, но также для развития новой специфичности к PAM посредством мутаций PID. Кроме того, некоторые фаги и MGE экспрессируют белки анти-CRISPR (Acr), которые ингибируют Cas9 (Pawluk et al., 2016; Hynes et al., 2017; Rauch et al., 2017). Связывание PID является эффективным ингибиторным механизмом, используемым некоторыми Acr (Dong et al., 2017; Shin et al., 2017; Yang and Patel, 2017), что указывает на то, что варьирование PID также может обеспечиваться давлением селекции для ускользания от ингибирования посредством Acr. PID Cas9 могут эволюционировать таким образом, чтобы близкородственные ортологи распознавали различные PAM, как недавно было выявлено в двух видах *Geobacillus*. Cas9, кодируемая *G. stearothermophilus*, распознает PAM N<sub>4</sub>CRAA, но, когда ее PID заменен на PID Cas9 штамма LC300, необходимо, чтобы ее PAM был заменен на N<sub>4</sub>GMAA (Harrington et al., 2017b).

В одном варианте осуществления настоящее изобретение охватывает множество ортологов Cas9 *N. meningitidis* с дивергирующими PID, которые распознают различные PAM. В одном варианте осуществления настоящее изобретение охватывает белок Cas9 с высокой идентичностью последовательности (>80% вдоль всей длины) с последовательностью NmeCas9 штамма 8013 (Nme1Cas9) (Zhang et al., 2013). Nme1Cas9 также имеет небольшой размер и природную высокую точность, как обсуждалось выше. (Lee et al., 2016; Amrani et al., 2018). Выравнивание продемонстрировало три клады ортологов Cas9 менингококков, каждая из которых обладает >98% идентичностью с N-концевыми ~820 аминокислотными (а.к.) остатками, которые включают все области белка, отличные от PID. См., фиг.5А и фиг.6А.

Все из этих ортологов Cas9 имеют длину 1078-1082 а.к. Первая клада (группа 1) включает ортологов, в которых существует идентичность последовательности >98% а.к. с Nme1Cas9 в PID. Напротив, другие две группы имели PID, которые значительно дивергировали от Nme1Cas9, причем ортологи группы 2 и группы 3 имели в среднем ~52% и ~86% идентичность последовательности PID с Nme1Cas9, соответственно. Из

каждой группы было выбрано по одному штамму менингококка: i) De11444 из группы 2; и ii) 98002 из группы 3 для детального анализа, которые обозначают в настоящем описании как Nme2Cas9 (1082 а.к.) и Nme3Cas9 (1081 а.к.), соответственно. Локусы CRISPR-*cas* из этих двух штаммов имеют последовательности повторов и длины спейсеров, которые идентичны тем, что у штамма 8013. См., фиг.6В. Это строго указывает на то, что их зрелые *cr*-РНК также имеют гидовые последовательности из 24 нуклеотидов и последовательности повторов из 24 нуклеотидов (Zhang et al., 2013). Аналогично, последовательности *tracr*-РНК De11444 и 98002 были на 100% идентичными *tracr*-РНК 8013. См., фиг.6В. Эти наблюдения подразумевают, что один и тот же каркас последовательности *sg*-РНК может обеспечивать расщепление ДНК посредством всех трех Cas9.

Для определения того, имеют ли эти ортологи Cas9 различные PAM, PID Nme1Cas9 заменяли на PID либо Nme2Cas9, либо Nme3Cas9. Для идентификации соответствующих требований к PAM, эти белковые химеры экспрессировали в *Escherichia coli*, очищали и использовали для идентификации PAM *in vitro* (Karvelis et al., 2015; Ran et al., 2015; Kim et al., 2017). В кратком изложении, совокупность фрагментов ДНК, содержавших протоспейсер, за которым следовала рандомизированная последовательность из 10 нуклеотидов, расщепляли *in vitro* с использованием рекомбинантной Cas9 и собственной транскрибированной *in vitro* *sg*-РНК. См., фиг.5В. Ожидалось, что только те ДНК, которые содержат последовательность PAM Cas9, будут расщепляться. Затем продукты расщепления секвенировали для идентификации PAM. См., фиг.5С-Д.

Ожидаемая консенсусная последовательность PAM N<sub>4</sub>GATT была подтверждена в выделенной полноразмерной Nme1Cas9. См., фиг.5С. Химерные производные с обменом PID продемонстрировали строгое предпочтение в отношении остатка С в 5-м положении вместо G, распознаваемого посредством Nme1Cas9. См., фиг.5Д.

В одном варианте осуществления ABE7.10-nNme2Cas9 (D16A) используют для редактирования единичных оснований пары А-Т на пару оснований G-С. В одном варианте осуществления BEmax-nNme2Cas9 (D16A) используют для редактирования единичных оснований пары оснований А-Т на пару оснований G-С (см., фигура 3F). На фиг.5А представлена иллюстративная схема, демонстрирующая мутантные остатки (оранжевые сферы) между Nme2Cas9 (слева) и Nme3Cas9 (справа), картированные на спрогнозированной структуре Nme1Cas9, отображающей кластер мутаций в PID (черный). На фиг.5В представлена иллюстративная экспериментальная схема процесса анализа обнаружения PAM *in vitro* с рандомизированной областью PAM из 10 п.н. После расщепления *in vitro* адаптеры лигировали с продуктами расщепления для конструирования библиотеки и секвенирования. На фиг.5С представлены иллюстративные логотипы последовательностей, полученных в результате обнаружения PAM *in vitro*, выявившего увеличение содержания PAM N<sub>4</sub>GATT для Nme1Cas9 в

соответствии с ее ранее установленной специфичностью. На фиг.5D представлены иллюстративные логотипы последовательностей, указывающие на то, что Nme1Cas9 с заменой его PID на PID Nme2Cas9 (слева) или Nme3Cas9 (справа) требует С в положении 5 PAM. Остальные нуклеотиды не были определены с высокой достоверностью вследствие умеренной эффективности расщепления белковых химер с заменой PID (см. фиг.6С). На фиг.5Е представлен иллюстративный логотип последовательности, демонстрирующий, что полноразмерная Nme2Cas9 распознает PAM N<sub>4</sub>CC, на основе эффективного расщепления субстрата совокупности мишеней с фиксированным С в 5 положении PAM, и с рандомизированными нуклеотидами 1-4 и 6-8 PAM.

Никакие оставшиеся нуклеотиды PAM не могли быть достоверно распределены вследствие низкой эффективности расщепления у химерных белков в используемых условиях. См., фиг.6С. Чтобы далее распределить PAM проводили анализ *in vitro* на библиотеке с рандомизированной последовательностью из 7 нуклеотидов, имеющей инвариантный С в 5-м положении PAM (например, 5'-NNNNCNNN-3' на некомплементарной цепи sg-PHK). Эта стратегия обеспечила значительно более высокую эффективность расщепления, и результаты показали, что PID Nme2Cas9 и Nme3Cas9 распознают PAM NNNNCC(A) и NNNNCAAA, соответственно. См., фиг.6С-D. Консенсусная последовательность Nme3Cas9 является сходной с консенсусной последовательностью GeoCas9 (Harrington et al., 2017b).

Эти тесты повторяли с использованием полноразмерной Nme2Cas9 (а не химеры с обменом PID) с совокупностью ДНК NNNNCNNN, и вновь была выделена консенсусная последовательность NNNNCC(A). См., фиг.5Е. Отмечалось, что этот тест имел более эффективное расщепление. См., фиг.6С. Эти данные указывают на то, что одна или более из 15 аминокислотных замен в Nme2Cas9 (относительно Nme1Cas9) вне PID поддерживает эффективную активность расщепления ДНК. См., фиг.6С. Поскольку уникальный PAM из 2-3 нуклеотидов из Nme2Cas9 обеспечивает более высокую плотность потенциальных участков-мишеней, чем описанные ранее компактные ортологи Cas9, она была выбрана для дальнейшего анализа.

На фиг.6 представлена охарактеризация ортологов Cas9 *Neisseria meningitidis* с быстро эволюционирующими PID, связанная с фиг.5. На фиг.6А представлено иллюстративное некорневое филогенетическое древо для ортологов NmeCas9, которые являются на >80% идентичными Nme1Cas9. Появилось три различных ветви, причем большинство мутантов кластеризуются в PID. Группы 1 (синий), 2 (оранжевый) и 3 (зеленый) имеют PID с идентичностью >98%, ~52% и ~86% с Nme1Cas9, соответственно. Указаны три репрезентативных ортолога Cas9 (по одному из каждой группы) (Nme1Cas9, Nme2Cas9 и Nme3Cas9). На фиг.6В представлена иллюстративная схема, демонстрирующая локусы CRISPR-*cas* цепей, кодирующих три ортолога Cas9 (Nme1Cas9, Nme2Cas9 и Nme3Cas9) из (А). Представлена процентная идентичность каждого компонента CRISPR-Cas с *N. meningitidis* 8013 (кодирующей Nme1Cas9). Синяя и красная

стрелки обозначают участки инициации транскрипции пре-cr-PHK и tract-PHK, соответственно. На фиг.6C представлен график иллюстративного нормализованного количества чтений (% от всех чтений) расщепленных ДНК из анализов *in vitro* в отношении интактной Nme1Cas9 (серый), в отношении химер с заменой PID Nme1Cas9 на PID из Nme2Cas9 и Nme3Cas9 (смешанные цвета) и полноразмерной Nme2Cas9 (оранжевый). Уменьшенное нормализованное количество чтений указывает на более низкую эффективность расщепления в химерах. На фиг.6D представлены иллюстративные логотипы последовательностей из анализа для обнаружения PAM *in vitro* на совокупности PAM NNNNCNNN посредством Nme1Cas9 с заменой PID на PID Nme2Cas9 (слева) или Nme3Cas9 (справа).

## 2. Обеспечиваемое PAM N<sub>4</sub>CC редактирование генов

Для тестирования эффективности Nme2Cas9 при редактировании генома человека полноразмерную (например, без замены PID) оптимизированную для кодонов человека конструкцию Nme2Cas9 клонировали в экспрессирующую плазмиду млекопитающих с прилагаемыми сигналами ядерной локализации (NLS) и линкерами, подтвержденными ранее для Nme1Cas9 (Amrani et al., 2018). Для первоначальных тестов использовали модифицированный репортер на флуоресцентной основе Traffic Light Reporter (TLR2.0) (Certo et al., 2011). В кратком изложении, после разрушенного GFP следует пептид T2A не в рамке считывания и кассета mCherry. После внесения двухцепочечных разрывов (DSB) ДНК в кассету с разрушенным GFP, ряд событий репарации по принципу негомологического соединения концов (NHEJ) оставляет инсерции-делеции со сдвигом рамки считывания +1, помещая mCherry в рамку считывания и обеспечивая красную флуоресценцию, которую можно без труда количественно определить посредством проточной цитометрии. См., фиг.7A. Результаты направляемой гомологией репарации (HDR) также можно оценивать одновременно путем включения ДНК-донора, который восстанавливает функциональную последовательность GFP, обеспечивая зеленую флуоресценцию (Certo et al., 2011). Поскольку некоторые инсерции-делеции не вносят сдвиг рамки считывания +1, определение флуоресценции, как правило, обеспечивает заниженную оценку истинной эффективности редактирования. Тем не менее, скорость, простота и низкая стоимость анализа делает его пригодным в качестве первоначального полуколичественного показателя редактирования генома в клетках HEK293T, содержащих один локус TLR2.0, включенный посредством лентивирусного вектора.

Для первоначальных тестов плазмиду Nme2Cas9 временно совместно трансфицировали с одной из пятнадцати плазмид sg-PHK, содержащих спейсеры, которые нацелены на участки TLR2.0, с PAM N<sub>4</sub>CC. Донор для HDR не включали, так что оценивали только редактирование на основе NHEJ (mCherry). Большинство sg-PHK имели формат G23 (т.е. 5'-концевой G для способствования транскрипции, а затем последовательность гида из 23 нуклеотидов), как обычно используют для Nme1Cas9 (Lee et al., 2016; Pawluk et al., 2016; Amrani et al., 2018; Ibraheim et al., 2018). В качестве

отрицательных контролей использовали отсутствие sg-PHK и sg-PHK, нацеленную на PAM N<sub>4</sub>GATT, и в качестве положительных контролей включали совместные трансфекции SpyCas9+sg-PHK и Nme1Cas9+sg-PHK (нацеленные на протоспейсеры NGG и N<sub>4</sub>GATT, соответственно). Редактирование посредством SpyCas9 и Nme1Cas9 без труда поддавалось обнаружению (~28% и 10% mCherry, соответственно). См., фиг. 7В.

Для Nme2Cas9, все 15 мишеней с PAM N<sub>4</sub>CC были функциональными, хотя и в различной степени в диапазоне от 4% до 20% согласно mCherry. Эти пятнадцать участков включают примеры с каждым из четырех возможных нуклеотидов в 7-м положении PAM (например, после динуклеотида CC), что указывает на то, что небольшое предпочтение остатка А наблюдалось *in vitro* (фиг. 5Е), которое не отражает потребность в PAM для применения для редактирования в клетках человека. Контроль в виде PAM N<sub>4</sub>GATT обеспечил сигнал mCherry, сходный с контролем без sg-PHK. См., фиг. 7В.

Для определения того, вовлечены ли оба остатка С в PAM N<sub>4</sub>CC при редактировании, тестировали серию участков PAM N<sub>4</sub>DC (D=A, T, G) и N<sub>4</sub>CD в репортерных клетках TLR2.0. См., фиг. 8А и 8В. Не было обнаружено поддающегося обнаружению редактирования ни в одном из этих участков, что первоначально указало на то, что оба остатка С консенсусного PAM N<sub>4</sub>CC необходимы для эффективной активности Nme2Cas9.

Длина спейсера в sg-PHK различается среди ортологов Cas9 и может влиять на активность в мишени против активности вне мишени (Cho et al., 2014; Fu et al., 2014). Оптимальная длина спейсера SpyCas9 составляет 20 нуклеотидов с возможностью укорочения до 17 нуклеотидов (Fu et al., 2014). Напротив, Nme1Cas9 обычно имеет спейсеры из 24 нуклеотидов (Hou et al., 2013; Zhang et al., 2013) и допускает укорочения до 18-20 нуклеотидов (Lee et al., 2016; Amrani et al., 2018). Для тестирования требований к длине спейсера для Nme2Cas9, гидовые РНК-плазмиды создавали для каждого единичного участка-мишени TLR2.0, но с различной длиной спейсера. См., фиг. 7С и фиг. 8С. Сравнимую активность наблюдали для гидов G23, G22 и G21, однако значительно сниженную активность наблюдали при дальнейшем укорочении до длины G20 и G19. См., фиг. 7С. Эти результаты валидируют Nme2Cas9 в качестве платформы редактирования генома с гидовыми последовательностями из 22-24 нуклеотидов в участках PAM N<sub>4</sub>CC в культивируемых клетках человека.

На фиг. 7 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 использует спейсер из 22-24 нуклеотидов для редактирования участков, соседних с PAM N<sub>4</sub>CC. Все эксперименты проводили в трех повторениях, и планки погрешности соответствуют стандартной ошибке среднего значения (s.e.m.). На фиг. 7А представлена иллюстративная схематическая диаграмма, на которой изображена временная трансфекция и редактирование клеток НЕК293Т TLR2.0 с детекцией mCherry+ клеток посредством проточной цитометрии через 72 часа после трансфекции. На фиг. 7В представлено иллюстративное редактирование посредством Nme2Cas9 репортера TLR2.0.

На участки с PAM N<sub>4</sub>CC происходило нацеливание с различной эффективностью, в то время как не наблюдали нацеливания Nme2Cas9 на PAM N<sub>4</sub>GATT или в отсутствии sg-PHK. В качестве положительных контролей использовали SpyCas9 (нацеленную на ранее подтвержденный участок с PAM NGG) и Nme1Cas9 (нацеленную на N<sub>4</sub>GATT). На фиг. 7C представлен иллюстративный эффект длины спейсера на эффективность редактирования Nme2Cas9. Нацеливание sg-PHK на единственный участок TLR2.0 с длиной спейсера, варьирующей от 24 до 20 нуклеотидов (включая 5'-концевой G, который требуется для промотора U6), показало, что наиболее высокая эффективность редактирования достигается посредством спейсеров из 22-24 нуклеотидов. На фиг. 7D показано, что иллюстративную двойную никазу Nme2Cas9 можно использовать в тандеме для обеспечения редактирования на основе NHEJ и HDR в TLR2.0. Плазмиды, экспрессирующие Nme2Cas9 и sg-PHK вместе с дцДНК-донором из 800 п.н. для гомологичной репарации вводили посредством электропорации в клетки HEK293T TLR2.0, и исходы как NHEJ (mCherry+), так и HDR (GFP+) оценивали посредством проточной цитометрии. Никаза HNH, Nme2Cas9<sup>D16A</sup>; никаза RuvC, Nme2Cas9<sup>H588A</sup>. На участки расщепления на расстоянии 32 п.н. и 64 п.н. осуществляли нацеливание с использованием каждой никазы. Никаза HNH (Nme2Cas9<sup>D16A</sup>) обеспечивала эффективное редактирование, в частности, с участками расщепления, которые находятся на расстоянии 32 п.н., в то время как никаза RuvC (Nme2Cas9<sup>H588A</sup>) не была эффективной. В качестве контроля использовали Nme2Cas9 дикого типа.

### 3. Точное редактирование посредством HDR и никазы HNH

Ферменты Cas9 используют их домены HNH и RuvC для расщепления комплементарной гиду и некомплемтарной цепи ДНК-мишени, соответственно. Никазы SpyCas9 (nCas9s), в которых либо домен HNH, либо домен RuvC инактивирован посредством мутации, используют для индукции направляемой гомологией репарации (HDR) и для повышения специфичности редактирования генома посредством индукции DSB с использованием двойных никаз (Mali et al., 2013a; Ran et al., 2013).

Для тестирования эффективности Nme2Cas9 в качестве никазы создавали Nme2Cas9<sup>D16A</sup> (никаза HNH) и Nme2Cas9<sup>H588A</sup> (никаза RuvC), которые имеют мутации аланина в каталитических остатках доменов RuvC и HNH, соответственно (Esvelt et al., 2013; Hou et al., 2013; Zhang et al., 2013). Клетки TLR2.0, вместе с донорной дцПНК GFP, использовали для определения того, могут ли индуцированные Nme2Cas9 разрывы индуцировать точное редактирование посредством HDR. Участки-мишени в TLR2.0 использовали для тестирования функциональности каждой никазы с использованием гидов, нацеленных на участки расщепления на расстоянии 32 п.о. и 64 п.о. См., фиг. 7D. Nme2Cas9 дикого типа, нацеленная на один участок, продемонстрировала эффективное редактирование как с NHEJ, так и с HDR в качестве исходов репарации. Для никаз участки расщепления на расстоянии 32 п.о. и 64 п.о. продемонстрировали редактирование с использованием Nme2Cas9<sup>D16A</sup> (никаза HNH), но ни одна пар-мишеней не работала с

Nme2Cas9<sup>H588A</sup>. Эти результаты указывают на то, что никазу Nme2Cas9 HNH можно использовать для эффективного редактирования генома при условии, что участки находятся вблизи.

Исследования ранее охарактеризованных Cas9 идентифицировали конкретную область вблизи PAM, где активность Cas9 является в высокой степени чувствительной к несоответствиям последовательности. Эта область их 8-12 нуклеотидов известна в качестве последовательности затравки, и она обнаружена во всех Cas9, охарактеризованных на сегодняшний день (Gorski et al., 2017). Для определения того обладает ли Nme2Cas9 также последовательностью затравки, проводили серию временных трансформаций, каждая из которых направлена на тот же локус в TLR2.0, но с одним несоответствием нуклеотидов в различных положениях гена. См., фиг.8D. Значительное уменьшение количества mCherry-положительных клеток наблюдали для несоответствий оснований в первых 10-12 нуклеотидах выше PAM, что указывает на то, что Nme2Cas9 имеет последовательность затравки в этой области.

На фиг.8 представлены иллюстративные данные, демонстрирующие требования к PAM, спейсеру и затравке для нацеливания Nme2Cas9 в клетках млекопитающих в соответствии с фиг.7. Все эксперименты проводили в трех повторениях и планки погрешности соответствуют s.e.m. На фиг.8A представлено иллюстративное нацеливание Nme2Cas9 на участки N<sub>4</sub>CD в TLR2.0 с оценкой редактирования на основе mCherry+ клеток. Оценивали четыре участка для каждого нуклеотида, не являющегося C, в тестируемом положении (N<sub>4</sub>CA, N<sub>4</sub>CT и N<sub>4</sub>CG) и в качестве положительного контроля использовали участок N<sub>4</sub>CC. На фиг.8B представлено иллюстративное нацеливание Nme2Cas9 на участки N<sub>4</sub>DC в TLR2.0 [аналогично (A)]. На фиг.8C представлены иллюстративные укорочения гена в участке TLR2.0 (отличном от участка на фиг.2C) с PAM N<sub>4</sub>CCA, демонстрирующие сходные требования к длине с требованиями, наблюдаемыми для другого участка. На фиг.8D показано, что иллюстративная эффективность нацеливания Nme2Cas9 является по-разному чувствительной к однонуклеотидным несоответствиям в области затравки sg-РНК. Данные демонстрируют эффекты перемещающихся однонуклеотидных несоответствий sg-РНК вдоль спейсера из 23 нуклеотидов в участке-мишени TLR2.0.

#### **4. Способы доставки в типы клеток млекопитающих**

Способность Nme2Cas9 функционировать в различных линиях клеток млекопитающих тестировали с использованием различных способов доставки. В качестве первоначального теста исследовали сорок (40) различных участков (тестировали 29 участков с PAM N<sub>4</sub>CC и 11 участков с PAM N<sub>4</sub>CD). Было отобрано несколько локусов (AAVSI, VEGFA и т.д.), и участки-мишени с PAM N<sub>4</sub>CC случайным образом выбирали для редактирования посредством Nme2Cas9. Редактирование (%) определяли посредством временной трансфекции 150 нг Nme2Cas9 вместе с 150 нг плазмид sg-РНК с последующим анализом TIDE через 72 часов после трансфекции. Подгруппа участков,

демонстрирующих спектр эффективностей редактирования в этом первоначальном скрининге, отбирали для повторного анализа в трех повторениях. См., фиг.9А и таблицу 1.

На фиг.9 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 в эндогенных локусах в клетках млекопитающих посредством множества способов доставки. Все результаты соответствуют 3 независимым биологическим повторениям и планки погрешности соответствуют s.e.m. На фиг.9А представлено иллюстративное редактирование генома посредством Nme2Cas9 для эндогенных участков человека в клетках HEK293Т после временной трансфекции плазмид, экспрессирующих Nme2Cas9 и sg-РНК. Сначала проводили скрининг 40 участков (таблица 1); затем 14 представленных участков (выбранных для включения репрезентативных примеров различной эффективности редактирования при определении посредством TIDE) повторно анализировали в трех повторениях. В качестве отрицательного контроля использовали участок-мишень Nme1Cas9 (с PAM N<sub>4</sub>GATT). На фиг.9В представлены иллюстративные карты данных: левая панель: временная трансфекция единичной плазмиды, экспрессирующей как Nme2Cas9, так и sg-РНК (нацеленной на локусы *Pcsk9* и *Rosa26*), позволяет редактирование в клетках мыши Нера1-6, определяемое посредством TIDE. Правая панель: введение посредством электропорации плазмид sg-РНК в клетки К562, стабильно экспрессирующие Nme2Cas9 из лентивирусного вектора, приводит к эффективному образованию инсерции-делеции. На фиг.9С показано, что иллюстративную Nme2Cas9 можно вводить посредством электропорации в качестве RNP-комплекса для индукции редактирования генома. 40 пикомоль Cas9 вместе с 50 пикомоль транскрибированных *in vitro* sg-РНК, нацеленных на различные локусы, вводили посредством электропорации в клетки HEK293Т. Инсерции-делеции определяли через 72 ч с использованием TIDE.

Таблица 1. Иллюстративные эндогенные участки редактирования генома человека, на которые нацелена Nme2Cas9.

№	Название участка	Последовательность спейсера	PAM	Локус	Редактирование (%)	№	Название участка	Название праймера TIDE	Праймер FW TIDE	Праймер RV TIDE
1	TS1	GGTTCTGG GТАCTTTТА TCTGTCC	CCTCCACC	<i>AAVSI</i>	ND	1	TS1	AAVS1_TIDE1	TGGCTTAGC ACCTCTCCAT	AGAACTCAG GACCAACTТА TTCTG
2	TS4	GTCTGCCT AACAGGAG GTGGGGGT	TAGACGAA	<i>AAVSI</i>	11	2	TS4	AAVS1_TIDE1	TGGCTTAGC ACCTCTCCAT	AGAACTCAG GACCAACTТА TTCTG
3	TS5	GAATATCA GGAGACTA GGAAGGAG	GAGGCCTA	<i>AAVSI</i>	15	3	TS5	AAVS1_TIDE1	TGGCTTAGC ACCTCTCCAT	AGAACTCAG GACCAACTТА TTCTG
4	TS6	GCCTCCCT GCAGGGCT GCTCCC	CAGCCCAA	<i>LINC01588</i>	20	4	TS6	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGCA
5	TS10	GAGCTAGT CTTCTTCCT CCAACCC	GGGCCCTA	<i>AAVSI</i>	3.5	5	TS10	AAVS1_TIDE1	TGGCTTAGC ACCTCTCCAT	AGAACTCAG GACCAACTТА TTCTG
6	TS11	GATCTGTC CCCTCCAC CCCACAGT	GGGGCCAC	<i>AAVSI</i>	9	6	TS11	AAVS1_TIDE1	TGGCTTAGC ACCTCTCCAT	AGAACTCAG GACCAACTТА TTCTG

7	TS12	GGCCAAA TGAAAGGA GTGAGAGG	TGACCCGA	<i>AAVSI</i>	10	7	TS12	AAVS1_TIDE2	TCCGTCTTCC TCCACTCC	TAGGAAGGA GGAGGCCTA AG
8	TS13	GCATCCTC TTGCTTTCT TTGCCTG	GACACCCCA	<i>AVSI</i>	2	8	TS13	AAVS1_TIDE2	TCCGTCTTCC TCCACTCC	TAGGAAGGA GGAGGCCTA AG
9	TS16	GGAGTCGC CAGAGGCC GGTGGTGG	ATTCCTC	<i>LINC01588</i>	28	9	TS16	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGCA
10	TS17	GCCAGCG GCCGATA TCAGCTGC	CACGCCCG	<i>LINC01588</i>	ND	10	TS17	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGCA
11	TS18	GGAAGGGA ACATATTA CTATTGC	TTCCCTC	<i>CYBB</i>	1	11	TS18	NTS55_TIDE	TAGAGAACT GGGTAGTGT G	CCAATATTGC ATGGGATGG
12	TS19	GTGGAGTG GCCTGCTA TCAGCTAC	CTATCCAA	<i>CYBB</i>	6	12	TS19	NTS55_TIDE	TAGAGAACT GGGTAGTGT G	CCAATATTGC ATGGGATGG
13	TS20	GAGGAAGG GAACATAT TACTATTG	CTTCCCT	<i>CYBB</i>	11.2	13	TS20	NTS55_TIDE	TAGAGAACT GGGTAGTGT G	CCAATATTGC ATGGGATGG

14	TS21	GTGAATTC TCATCAGC TAAAATGC	CAAGCCTT	<i>CYBB</i>	1	14	TS21	NTS55_TIDE	TAGAGAACT GGGTAGTGT G	CCAATATTGC ATGGGATGG
15	TS25	GCTCACTC ACCCACAC AGACACAC	ACGTCCTC	<i>VEGFA</i>	15.6	15	TS25	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
16	TS26	GGAAGAAT TTCATTCTG TTCTCAG	TTTTCTG	<i>CFTR</i>	2	16	TS26	hCFTR_TIDE1	TGGTGATTA TGGGAGAAC TGGAGC	ACCATTGAG GACGTTTGTC TCAC
17	TS27	GCTCAGTT TTCCTGGA TTATGCCT	GGCACCAT	<i>CFTR</i>	4	17	TS27	hCFTR_TIDE1	TGGTGATTA TGGGAGAAC TGGAGC	ACCATTGAG GACGTTTGTC TCAC
18	TS31	GCGTTGGA GCGGGGAG AAGGCCAG	GGGTCACT	<i>VEGFA</i>	9	18	TS31	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
19	TS34	GGGCCGCG GAGATAGC TGCAGGGC	GGGGCCCC	<i>LINC01588</i>	ND	19	TS34	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGGCA
20	TS35	GCCCACCC GGCGGCGC CTCCCTGC	AGGGCTGC	<i>LINC01588</i>	ND	20	TS35	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGGCA

21	TS36	GCGTGGCA GCTGATAT CCGGCCGC	TGGGCGTC	<i>LINC01588</i>	ND	21	TS36	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGGCA
22	TS37	GCCGCGGC GCGACGTG GAGCCAGC	CCCGCAA	<i>LINC01588</i>	ND	22	TS37	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGGCA
23	TS38	GTGCTCCC CAGCCCAA ACCGCCGC	GGCGCGAC	<i>LINC01588</i>	2	23	TS38	LINC01588_TIDE	AGAGGAGCC TTCTGACTGC TGCAGA	ATGACAGAC ACAACCAGA GGGCA
24	TS41	GTCAGATT GGCTTGCT CGGAATTG	CCAGCCAA	<i>AGA</i>	3	24	TS41	AGA_TIDE1	GGCATAAGG AAATCGAAG GTC	CATGTCCTCA AGTCAAGAA CAAG
25	TS44	GCTGGGTG AATGGAGC GAGCAGCG	TCTTCGAG	<i>VEGFA</i>	3	25	TS44	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
26	TS45	GTCCTGGA GTGACCCC TGGCCTTC	TCCCCGCT	<i>VEGFA</i>	7.4	26	TS45	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
27	TS46	GATCCTGG AGTGACCC CTGGCCTT	CTCCCCGC	<i>VEGFA</i>	6	27	TS46	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC

28	TS47	GTGTGTCC CTCTCCCC ACCCGTCC	CTGTCCGG	<i>VEGFA</i>	23.1	28	TS47	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
29	TS48	GTTGGAGC GGGGAGAA GGCCAGGG	GTCACTCC	<i>VEGFA</i>	2	29	TS48	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
30	TS49	GCGTTGGA GCGGGGAG AAGGCCAG	GGGTCACT	<i>VEGFA</i>	4	30	TS49	VEGF_TIDE3	GTACATGAA GCAACTCCA GTCCCA	ATCAAATTCC AGCACCGAG CGC
31	TS50	GTACCCTC CAATAATT TGGCTGGC	AATTCCGA	<i>AGA</i>	6	31	TS50	AGA_TIDE1	GGCATAAGG AAATCGAAG GTC	CATGTCCTCA AGTCAAGAA CAAG
32	TS51	GATAATTT GGCTGGCA ATTCCGAG	CAAGCCAA	<i>AGA</i>	4.5	32	TS51	AGA_TIDE1	GGCATAAGG AAATCGAAG GTC	CATGTCCTCA AGTCAAGAA CAAG
33	TS58 (DS1)	GCAGGGGC CAGGTGTC CTTCTCTG	GGGGCCTC	<i>VEGFA</i>	5	33	TS58 (DS1)	VEGF_TIDE4	ACACGGGCA GCATGGGAA TAGTC	GCTAGGGGA GAGTCCCCT GTCCA
34	TS59 (DS2)	GAATGGCA GGCGGAGG TTGTAAGT	GGGGCCAG	<i>VEGFA</i>	11.5	34	TS59 (DS2)	VEGF_TIDE5	CCTGTGTGG CTTTGCTTTG GTC	GGTAGGGTG TGATGGGAG GCTAAGC

35	TS60 (DS3)	GAGTGAGA GAGTGAGA GAGAGACA	CGGGCCAG	<i>VEGFA</i>	3	35	TS60 (DS3)	VEGF_TIDE5	CCTGTGTGG CTTTGCTTTG GTC	GGTAGGGTG TGATGGGAG GCTAAGC
36	TS61 (DS4)	GTGAGCAG GCACCTGT GCCAACAT	GGGCCCCG	<i>VEGFA</i>	3.5	36	TS61 (DS4)	VEGF_TIDE5	CCTGTGTGG CTTTGCTTTG GTC	GGTAGGGTG TGATGGGAG GCTAAGC
37	TS62 (DS5)	GCGTGGGG GCTCCGTG CCCCACGC	GGGTCCAT	<i>VEGFA</i>	3.4	37	TS62 (DS5)	VEGF_TIDE6	GGAGGAAGA GTAGCTCGC CGAGG	AGACCGAGT GGCAGTGAC AGCAAG
38	TS63 (DS6)	GCATGGGC AGGGGCTG GGGTGCAC	AGGCCCAG	<i>VEGFA</i>	16	38	TS63 (DS6)	VEGF_TIDE7	AGGGAGAGG GAAGTGTGG GGAAGG	GTCTTCCTGC TCTGTGCGCA CGAC
39	TS64	GAAAATTG TGATTCC AGATCCAC	AAGCCCAA	<i>FANCI</i>	7	39	TS64	FancJ_TIDE5	GTTGGGGGC TCTAAGTTAT GTAT	CTTCATCTGT ATCTTCAGGA TCA
40	TS65	GAGCAGAA AAAATTGT GATTCC	AGATCCAC	<i>FANCI</i>	ND	40	TS65	FancJ_TIDE5	GTTGGGGGC TCTAAGTTAT GTAT	CTTCATCTGT ATCTTCAGGA TCA

Клетки HEK293T использовали для временной трансфекции и через 72 часа после трансфекции клетки собирали, а затем проводили экстракцию геномной ДНК и селективную амплификацию локуса-мишени. Анализ TIDE использовали для определения эффективности инсерций-делеций в каждом локусе (Brinkman et al., 2014).

Редактирование Nme2Cas9 обнаруживалось в большинстве из этих участков, даже несмотря на то, что эффективность варьировалась в зависимости от последовательности-мишени. Таблица 1. Интересно, Nme2Cas9 индуцировала инсерции-делеции в нескольких геномных участках с PAM N<sub>4</sub>CD, хотя и менее стабильно и на более низких уровнях. Таблица 1. Анализировали четырнадцать участков (14) с PAM N<sub>4</sub>CC в трех повторениях и наблюдали стабильное редактирование. См., фиг.9А. Кроме того, эффективность редактирована может быть значительно улучшена путем увеличения количества доставленной плазмиды Nme2Cas9, и эта высокая эффективность может быть расширена на точную сегментную делецию с двумя гидами. См., фиг.10А и 10В.

Способность Nme2Cas9 функционировать тестировали в клетках мыши Нера1-6 (происходящих из гепатомы). Для клеток Нера1-6 одну плазмиду, кодирующую как Nme2Cas9, так и sg-РНК (нацеленную либо на *Rosa26*, либо на *Pcsk9*), вводили посредством временной трансфекции и количественно определяли инсерции-делеции через 72 ч. Редактирование было хорошо заметным в обоих участках. См., фиг.9В, слева. Функциональность Nme2Cas9 также тестировали при стабильной экспрессии в клетках лейкоза человека K562. Для этого создавали лентивирусную конструкцию, экспрессирующую Nme2Cas9, и трансдуцировали клетки для стабильной экспрессии Nme2Cas9 под контролем промотора SFFV. Эта стабильная клеточная линия не демонстрировала никаких заметных отличий в отношении роста и морфологии по сравнению с нетрансдуцированными клетками, что указывает на то, что Nme2Cas9 является нетоксичной при стабильной экспрессии. В эти клетки временно вводили посредством электропорации плазмиды, экспрессирующие sg-РНК, и их анализировали посредством TIDE через 72 часа для определения эффективности инсерций-делеций. Эффективное (>50%) редактирование наблюдали во всех трех протестированных участках, подтверждая способность Nme2Cas9 функционировать при лентивирусной доставке в клетки K562. См., фиг.9В.

Доставка рибонуклеопротеинов (RNP) Cas9 и ее sg-РНК также является пригодной для некоторых применений для редактирования генома, и меньшее время присутствия Cas9 может минимизировать редактирование вне мишени (Kim et al., 2014; Zuris et al., 2015). Более того, некоторые типы клеток (например, определенные иммунные клетки) являются не поддающимися редактированию на основе трансфекции ДНК (Schumann et al., 2015). Для тестирования того, является ли Nme2Cas9 функциональной посредством доставки RNP, меченную 6×His Nme2Cas9 (слитую с тремя NLS) клонировали в бактериальную экспрессирующую конструкцию, и рекомбинантный белок очищали. Затем рекомбинантный белок нагружали sg-РНК, транскрибированными РНК-

полимеразой T7, нацеленными на три ранее подтвержденных участка. Электропорация комплекса Nme2Cas9:sg-PHK индуцировала успешное редактирование в каждом из трех участков-мишеней в клетках HEK293T при определении посредством TIDE. См., фиг.9С. В совокупности, эти результаты указывают на то, что Nme2Cas9 может быть эффективно доставлена посредством плазмиды или лентивируса, или в качестве комплекса RNP, в множество типов клеток.

### 5. Регуляция, направленная против CRISPR

На сегодняшний день показано, что пять семейств Aсг из различных видов бактерий ингибируют Nme1Cas9 *in vitro* и в клетках человека (Pawluk et al., 2016; Lee et al., 2018, представлена на рассмотрение). Учитывая высокую идентичность последовательностей между Nme1Cas9 и Nme2Cas9, по меньшей мере некоторые из этих семейств Aсг должны ингибировать Nme2Cas9. Для тестирования этого все пять семейств рекомбинантных Aсг экспрессировали, очищали и тестировали в отношении способности Nme2Cas9 расщеплять мишень *in vitro* в присутствии представителя каждого семейства (молярное соотношение Aсг:Cas9 10:1). Ингибитор использовали для системы CRISPR типа I-E в *E. coli* (AсгE2) в качестве отрицательного контроля, в то время как в качестве положительного контроля использовали Nme1Cas9. (Pawluk et al., 2014); (Pawluk et al., 2016). Как и ожидалось, все 5 семейств ингибировали Nme1Cas9, в то время как AсгE2 не ингибировал ее. См., фиг.11А, сверху. AсгПC1<sub>Nme</sub>, AсгПC2<sub>Nme</sub>, AсгПC3<sub>Nme</sub> и AсгПC4<sub>Hpa</sub> полностью ингибировали Nme2Cas9. Однако неожиданно AсгПC5<sub>Smu</sub>, который ранее был описан в качестве наиболее мощного из ингибиторов Nme1Cas9 (Lee et al., 2018), не ингибировал Nme2Cas9 *in vitro* даже в 10-кратном молярном избытке. Это указывает на то, что он вероятно ингибирует Nme1Cas9 посредством взаимодействия с ее PID.

На фиг.10 представлены иллюстративные данные, демонстрирующие зависимость от дозы и сегментные делеции посредством Nme2Cas9, связанные с фиг.9. На фиг.10А показано, что иллюстративное увеличение дозы введенной посредством электропорации плазмиды Nme2Cas9 (500 нг против 200 нг на фиг.3А) повышает эффективность редактирования в двух участках (TS16 и TS6). Данные, указанные желтым цветом, использовали повторно из фиг.9А. На фиг.10В показано, что иллюстративную Nme2Cas9 можно использовать для создания точных сегментных делеций. На две мишени TLR2.0 с участками расщепления на расстоянии 32 п.н. проводили одновременное нацеливание посредством Nme2Cas9. Большинство внесенных повреждений представляли собой делеции точно 32 п.н. (синий).

На фиг.11 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 подвергается ингибированию посредством подгруппы семейств белков анти-CRISPR типа II-C *in vitro* и в клетках. Все эксперименты проводили в трех повторениях и планки погрешности соответствуют s.e.m.

На фиг.11А представлен иллюстративный анализ расщепления *in vitro* для Nme1Cas9 и Nme2Cas9 в присутствии пяти ранее охарактеризованных белков анти-

CRISPR (соотношение 10:1 Асг:Cas9). Сверху: Nme1Cas9 эффективно расщепляет фрагмент, содержащий протоспейсер с PAM N<sub>4</sub>GATT в отсутствие Асг или в присутствии отрицательного контрольного Асг (АсгE2). Все пять ранее охарактеризованных семейств Асг типа II-C ингибировали Nme1Cas9, как и ожидалось. Снизу: ингибирование Nme2Cas9 отражает ингибирование Nme1Cas9, за исключением отсутствия ингибирования посредством АсгПC5<sub>Smu</sub>. На фиг.11В представлено иллюстративное редактирование генома в присутствии пяти ранее описанных семейств белков анти-CRISPR. Плазмиды, экспрессирующие Nme2Cas9 (200 нг), sg-ПНК (100 нг) и каждый соответствующий Асг (200 нг) совместно трансфицировали в клетки НЕК293Т и редактирование генома определяли с использованием отслеживания инсерций-делеций по распаду (Tracking of Indels by Decomposition (TIDE)) через 72 ч после трансфекции. В соответствии с проведенным автором изобретения анализом *in vitro*, все белки анти-CRISPR типа II-C за исключением АсгПC5<sub>Smu</sub> ингибировали редактирование генома, хотя и с различной эффективностью. На фиг.11С показано, что иллюстративное ингибирование Nme2Cas9 посредством Асг является зависимым от дозы с двумя различными кажущимися эффективностями. Nme2Cas9 полностью ингибировалась посредством АсгПC1<sub>Nme</sub> и АсгПC4<sub>Hpa</sub> в соотношениях по массе 2:1 и 1:1 совместно трансфицированных плазмид Асг и Nme2Cas9, соответственно.

Чтобы далее это протестировать, тестировали химеру Nme1Cas9/Nme2Cas9 с PID из Nme2Cas9. См., фиг.5D и фиг.6D. Вследствие сниженной активности этого гибрида в ~30 раз более высокую концентрацию Cas9 использовали для достижения сходной эффективности расщепления при поддержании молярного соотношения Cas9:Асг 10:1. Не наблюдали ингибирования этой белковой химеры посредством АсгПC5<sub>Smu</sub>. См., фиг.12. Эти данные обеспечивают дальнейшее доказательство того, что АсгПC5<sub>Smu</sub>, вероятно, взаимодействует с PID Nme1Cas9. Независимо от механистической основы отличающегося ингибирования посредством АсгПC5<sub>Smu</sub>, эти результаты указывают на то, что Nme2Cas9 ингибируется четырьмя другими семействами Асг типа II-C.

На фиг.12 представлены иллюстративные данные, демонстрирующие, что замена PID Nme2Cas9 делает Nme1Cas9 нечувствительной к ингибированию АсгПC5<sub>Smu</sub>, связанные с фиг.11. Показано расщепление *in vitro* посредством химеры Nme1Cas9-Nme2Cas9PID в присутствии ранее охарактеризованных белков Асг (10 мкМ Cas9-sgRNA+100 мкМ Асг).

Исходя из приведенных выше данных *in vitro*, было предположено, что АсгПC1<sub>Nme</sub>, АсгПC2<sub>Nme</sub>, АсгПC3<sub>Nme</sub> и АсгПC4<sub>Hpa</sub> можно использовать в качестве выключателей редактирования генома посредством Nme2Cas9. Чтобы протестировать это, проводили трансфекции плазмид Nme2Cas9/sgПНК (150 нг каждой плазмиды), нацеленных на TS16, в клетках НЕК293Т в присутствии или в отсутствие экспрессирующих Асг плазмид, поскольку было описано, что большинство Асг ингибируют Nme1Cas9 при этих соотношениях плазмид (Pawluk et al., 2016). Как и ожидалось, АсгПC1<sub>Nme</sub>, АсгПC2<sub>Nme</sub>,

АсгПС3<sub>Nme</sub> и АсгПС4<sub>Hpa</sub> ингибировали редактирование генома посредством Nme2Cas9, в то время как АсгПС5<sub>Smi</sub> не имел эффекта. См., фиг.11В. Полное ингибирование наблюдалось посредством АсгПС3<sub>Nme</sub> и АсгПС4<sub>Hpa</sub>, что указывает на то, что они обладают высокой эффективностью против Nme2Cas9 по сравнению с АсгПС1<sub>Nme</sub> и АсгПС2<sub>Nme</sub>. Чтобы далее сравнить эффективность АсгПС1<sub>Nme</sub> и АсгПС4<sub>Hpa</sub> авторы изобретения повторили эксперименты при различных соотношения плазмиды Асг и плазмиды Cas9. См., фиг.11С. Данные показывают, что плаزمида АсгПС4<sub>Hpa</sub> является особенно эффективной против Nme2Cas9. Взятые вместе, эти данные указывают на то, что несколько белков Асг можно использовать в качестве выключателей для применений на основе Nme2Cas9.

### **6. Сверхточность**

Nme1Cas9 демонстрирует примечательную точность редактирования в клетках и в моделях на мышах (Lee et al., 2016; Amrani et al., 2018; Ibraheim et al., 2018). Более того, сходство Nme2Cas9 со Nme1Cas9 по всей ее длине указывает на то, что она аналогично может быть сверхточной. Однако более высокое количество участков, пробуемых в геноме в результате динуклеотидного PAM, могут создать больше возможностей для нацеливания Nme2Cas9 вне мишени по сравнению с Nme1Cas9 и ее менее часто встречающимся 4-нуклеотидным PAM. Для оценки профиля действия вне мишени Nme2Cas9 использовали GUIDE-seq (полногеномная несмещенная идентификация двухцепочечных разрывов с использованием секвенирования) для эмпирической идентификации потенциальных участков вне мишени и без смещения (Tsai et al., 2014). Даже наилучшие алгоритмы прогнозирования действия вне мишени склонны к ложноотрицательным результатам, обуславливая необходимость в эмпирических способах определения профилей участков-мишеней (Bolukbasi et al., 2015b; Tsai and Joung, 2016; Tysko et al., 2016). GUIDE-seq основана на включении двухцепочечных олигодезоксинуклеотидов (dsODN) в участки двухцепочечного разрыва ДНК на протяжении генома. Затем проводят детекцию этих участков встраивания посредством амплификации и высокопроизводительного секвенирования.

Поскольку SpyCas9 является хорошо охарактеризованным ортологом Cas9, она является пригодной для мультиплексных применений с другими Cas9, и в качестве образца их свойств редактирования (Jiang and Doudna, 2017; Komor et al., 2017). SpyCas9 и Nme2Cas9 клонировали в идентичные плазмидные остовы с одинаковыми UTR, линкерами, NLS и промоторами, для параллельной временной трансфекции (вместе с аналогичным образом подобранными плазмидами, экспрессирующими sg-PHK) в клетки HEK293T. Сначала подтверждали, что РНК-гиды SpyCas9 и Nme2Cas9 являются ортогональными, т.е. что sg-PHK Nme2Cas9 не направляет редактирование посредством SpyCas9 и наоборот. См., фиг.13А. Это противоречило ранее описанным результатам для Nme1Cas9 (Esvelt et al., 2013; Fonfara et al., 2014).

Далее, для идентификации применения SpyCas9 в качестве образца для GUIDE-seq,

поскольку SpyCas9 и Nme2Cas9 имеют не перекрывающиеся PAM, то потенциально является возможным редактирование любого двойного участка (DS), фланкируемого последовательностью 5'-NGGNCC-3', который одновременно удовлетворяет требованиям к PAM обеих Cas9. Это позволяет непосредственное сравнение действия вне мишени с PНК-гидами, которые способствуют редактированию одного и того же участка-мишени. См., фиг.14А. Проводили нацеливание на шесть (6) DS в *VEGFA*, каждый из которых также имеет G в соответствующих положениях с 5'-стороны от PAM, так что оба гида SpyCas9 и Nme2Cas9 (контролируемых промотором U6) были на 100% комплементарными участку-мишени. Через семьдесят два (72) часа после трансфекции проводили анализ TIDE для этих участков, на которые нацелена каждая нуклеаза. Nme2Cas9 индуцировала инсерции-делеции во всех шести участках, хотя и с низкой эффективностью в двух из них, в то время как SpyCas9 индуцировала инсерции-делеции в четырех из шести участков. См., фиг.14В. В двух из четырех участков (DS1 и DS4), в которых SpyCas9 была эффективной, она индуцировала в ~7 раз больше инсерций-делеций, чем Nme2Cas9, в то время как Nme2Cas9 индуцировала в ~3 раза более высокую частоту инсерций-делеций, чем SpyCas9, в участке DS6. Оба ортолога Cas9 редактировали DS2 приблизительно с равной эффективностью.

Для GUIDE-seq DS2, DS4 и DS6 были отобраны для исследования расщепления вне мишени с гидами Nme2Cas9, которые обеспечивают редактирование в мишени настолько же эффективно, менее эффективно или более эффективно, чем соответствующие гида SpyCas9, соответственно. В дополнение к трем двойным участкам добавляли TS6, поскольку наблюдалось, что он эффективно редактировал участки-мишени Nme2Cas9, имеющие эффективность инсерций-делеций приблизительно 30-50%, в зависимости от типа клеток. См., фиг.9А и 10А. Сходные данные наблюдают для участков Nme2Cas9 *Pcsk9* и *Rosa26*. См., фиг.9В.

Трансфекции плазмид проводили для каждого аналога Cas9 с его собственной sg-PНК и dsODN. Затем получали библиотеки GUIDE-seq, как описано ранее (Amrani et al., 2018). Анализ GUIDE-seq выявил эффективное редактирование в мишени для обоих ортологов Cas9 с относительной эффективностью (отражаемой посредством количества чтений GUIDE-seq), которая сходна с эффективностью, наблюдаемой посредством TIDE. Фигура 13В и таблица 2. (Tsai et al., 2014; Zhu et al., 2017).

На фиг.13 представлены иллюстративные данные, демонстрирующие ортогональность и относительную точность Nme2Cas9 и SpyCas9 в отношении двойных участков-мишеней, связанные с фиг.12. На фиг.13А показано, что иллюстративные гида Nme2Cas9 и SpyCas9 являются ортогональными. Результаты TIDE показывают частоты инсерций-делеций, внесенных обеими нуклеазами, нацеленными на DS2, либо с их собственными sg-PНК, либо с sg-PНК другого ортолога. На фиг.13В представлены иллюстративные Nme2Cas9 и SpyCas9, демонстрирующие сравнимую эффективность редактирования в мишени при оценке посредством GUIDE-seq. Столбики указывают на

количества чтений с редактированием в мишени согласно GUIDE-Seq в трех двойных участках, на которые нацелен каждый ортолог. Оранжевые столбики соответствуют Nme2Cas9 и черные столбики соответствуют SpyCas9. На фиг.13С представлено иллюстративное количество чтений с редактированием в мишени против редактирования вне мишени для SpyCas9 в каждом участке. Оранжевые столбики соответствуют чтениям с редактированием в мишени, в то время как черные столбики соответствуют редактированию вне мишени. На фиг.13D представлено иллюстративные чтения с редактированием в мишени против редактирования вне мишени для Nme2Cas9 в каждом участке. На фиг.13Е представлены столбиковые диаграммы, демонстрирующие иллюстративную эффективность инсерций-делеций (при измерении посредством TIDE) в потенциальных участках вне мишени, спрогнозированных посредством CRISPRSeek. Последовательности участков в мишени и вне мишени представлены слева, причем область PAM подчеркнута и несоответствия sg-РНК и неконсенсусные нуклеотиды PAM указаны красным цветом.

**Таблица 2: Данные GUIDE-seq**

offTarget	peak_score	predicted_cleavage_score
chr6:-:43748587:43748609	652	100
chr1:+:82004618:82004640	304	4.1
chr1:-:31140567:31140589	275	19.6
chr16:+:30357052:30357074	226	0.6
chr5:-:33453895:33453917	217	4
chr11:+:116600352:116600374	206	0.4
chr17:-:46938649:46938671	191	0.6
chr9:-:130859778:130859800	146	5.4
chr15:+:59837681:59837703	143	2.6
chr22:-:19135541:19135563	124	0.3
chrX:+:49057600:49057622	122	0.6
chr7:-:72751388:72751410	117	2.6
chr3:-:51652045:51652067	115	0.3
chr1:-:9544334:9544356	109	0.7
chr3:-:47868006:47868028	99	2.6
chr9:+:140670069:140670091	91	0.4
chr2:-:149516035:149516057	90	0.3
chr22:-:18245713:18245735	89	0.2
chr3:+:154744438:154744460	89	2.6
chr17:-:73320669:73320691	88	0.7
chr1:-:38479457:38479479	85	2.6
chr7:+:33058792:33058814	78	0.3
chr9:+:108299833:108299855	76	1
chr1:-:23627429:23627451	74	0.5

chr2:-:63393272:63393294	74	0.5
chr16:+:71467786:71467808	70	0.6
chr1:-:111638773:111638795	67	0.3
chr1:-:213393740:213393762	67	0.5
chr7:+:38284425:38284447	67	0.3
chr7:-:134511606:134511628	66	0.2
chr7:+:152293366:152293388	66	0.7
chr17:+:60243345:60243367	63	0.5
chrX:-:48007735:48007757	60	0.6
chr1:+:52768707:52768729	58	5.4
chr19:-:38805324:38805346	58	0.3
chrX:-:41283776:41283798	58	2.6
chr11:-:14539718:14539740	57	2.6
chr6:+:32895093:32895115	57	0.7
chr7:-:138957343:138957365	56	98.6
chr3:-:63900682:63900704	52	0.4
chr5:-:79624954:79624976	52	2.6
chr7:+:76012229:76012251	52	0.7
chrX:+:39889198:39889220	52	2.6
chr4:-:99897525:99897547	51	5.4
chr1:-:25822709:25822731	50	0.7
chr5:+:17293204:17293226	50	0.7
chr13:-:66697991:66698013	49	0.1
chr5:-:80796103:80796125	49	2.6
chr16:+:49239128:49239150	45	1.9
chr3:+:69489884:69489906	43	0.5
chr8:+:113712655:113712677	42	0.3
chr2:-:24502672:24502694	39	2.6
chr7:-:65642349:65642371	39	2.6
chrX:-:135700076:135700098	37	2.6
chr1:-:99795756:99795778	36	6.2
chr19:+:1821377:1821399	36	0.2
chr4:-:75501534:75501556	36	0.3
chr18:+:74828740:74828762	34	0.3
chrX:+:133975784:133975806	34	6.2
chr14:+:55717904:55717926	33	98.6
chr13:+:49522615:49522637	32	0.3
chr3:-:77788415:77788437	32	0.7
chr11:-:48230825:48230847	31	6.2
chr1:-:1280441:1280463	30	0.3
chr7:+:44602379:44602401	30	5.4
chr12:-:108166294:108166316	29	5.4
chr7:-:111929850:111929872	29	4

chr12:-:122404237:122404259	27	0.2
chr12:-:79123453:79123475	27	0.7
chr22:-:46412541:46412563	27	6.2
chr5:+:93889070:93889092	26	0.3
chr10:-:97776548:97776570	25	0.6
chr2:-:56533335:56533357	24	98.6
chr3:+:149843401:149843423	24	0.1
chr1:-:232769157:232769179	23	2.6
chr15:-:75100050:75100072	21	2.6
chr18:+:37252965:37252987	21	0.6
chr2:-:44506208:44506230	21	2.6
chr4:+:182389352:182389374	21	0.6
chr11:+:9360929:9360951	20	98.6
chr12:+:23638452:23638474	19	0.4
chr7:-:66498753:66498775	19	1.4
chr13:+:32055862:32055884	16	6.2
chr15:-:59331986:59332008	16	6.2
chr2:+:126196868:126196890	16	0.7
chrX:-:77359566:77359588	16	0
chrX:+:24652788:24652810	16	6.2
chr17:-:17667857:17667879	15	0.4
chr21:+:34751155:34751177	15	2.6
chr2:-:48734975:48734997	14	5.4
chr1:-:69755048:69755070	13	2.6
chr16:+:90013282:90013304	13	1.1
chr18:-:630757:630779	13	5.4
chr3:-:163905630:163905652	12	0.6

SEQ ID NOS:	gRNAPlusPAM	SEQ ID NOS:	offTarget_sequence
168	GGCAGGCGGAGGTTGTACTGNNG	169	GGCAGGCGGAGGTTGTACTGGGG
168	GGCAGGCGGAGGTTGTACTGNNG	170	GGAAGGCGGAAGTTGTACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	171	GGCAGGCGGAGGTTGTAGTGGGG
168	GGCAGGCGGAGGTTGTACTGNNG	172	AGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	173	GGGAGGTGGAGGTTGTACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	174	GGCAGGGGGAAGCTGTACTGTGG
168	GGCAGGCGGAGGTTGTACTGNNG	175	AGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	176	AGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	177	GGGAGGCGGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	178	GGCAAGAGGAGGTTGGACTGGGG

168	GGCAGGCGGAGGTTGTACTGNNG	179	AGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	180	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	181	AGGAAGCGGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	182	AGGAGGCGGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	183	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	184	TCCAGGTGGAGGCTGTACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	185	AGGAGGCAGAGGTTGCACTGGGG
168	GGCAGGCGGAGGTTGTACTGNNG	186	GGGAGGCGGAGGATGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	187	CACAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	188	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	189	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	190	AGGAGGCAGAGGTTGAACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	191	GGCAAGGGGAAGTTGTACTGTGG
168	GGCAGGCGGAGGTTGTACTGNNG	192	GGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	193	GAGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	194	AGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	195	AGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	196	GGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	197	CAGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	198	AGCAGGTAGAGGTTGGACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	199	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	200	GGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	201	AGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	202	TGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	203	AGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	204	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	205	AGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	206	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	207	GGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	208	GGGAGGTGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	209	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	210	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	211	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	212	AGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	213	AGAAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	214	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	215	AGGAGGCGGAGGCTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	216	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	217	GGGAGGCGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	218	GGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	219	AGGAGGCAGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	220	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	221	GGAAGGCGGAGGTTGTAGTGAGG

168	GGCAGGCGGAGGTTGTACTGNNG	222	AGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	223	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	224	GGAAGGTGAAGGCTGTACTGCGG
168	GGCAGGCGGAGGTTGTACTGNNG	225	AGAAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	226	AGTAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	227	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	228	GGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	229	AGGAGGCAGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	230	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	231	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	232	GCCAGGCGGGTGCTGTACTGGGG
168	GGCAGGCGGAGGTTGTACTGNNG	233	AGGAGGCGGAGGTTGTACTGGGC
168	GGCAGGCGGAGGTTGTACTGNNG	234	TGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	235	GGGAGGTGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	236	AGGAGGTGGAGGTTGTAATGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	237	AGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	238	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	239	AGGAGGCAGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	240	GGGAGGCAGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	241	GGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	242	GGTAGGCAAAGGTTGTACCAGGG
168	GGCAGGCGGAGGTTGTACTGNNG	243	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	244	AGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	245	GGGAGGCAGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	246	GGGAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	247	GGGAGGCAGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	248	GGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	249	GGGAGGTGGAGGTTGCACTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	250	CAGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	251	GGGAGGCAGAGGTTGTACTGAGT
168	GGCAGGCGGAGGTTGTACTGNNG	252	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	253	AGAAGGCGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	254	CGTCTGCGAGGGTACTAGTGAGA
168	GGCAGGCGGAGGTTGTACTGNNG	255	GGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	256	GGGAGACGGAGGTTGTAGTGAGG
168	GGCAGGCGGAGGTTGTACTGNNG	257	AGGAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	258	AGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	259	AGAAGGCAGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	260	GCCAGGCTGAGGATGTACTGTGG
168	GGCAGGCGGAGGTTGTACTGNNG	261	AGGAGGCGGAGGTTGTACTGAGC
168	GGCAGGCGGAGGTTGTACTGNNG	262	GGGAGGCAGAGGTTGTAGTGAGG
<b>guideAlignment2OffTarget</b>			
<b>SEQ ID</b>		<b>offTarge</b>	<b>mismatch.distance2PAM</b>

	NOS:	tStrand	
.....		-	-
..A.....A.....		+	18,10
.....G..		-	3
A.G.....C....		+	20,18,5
..G...T.....		-	18,14
.....G..A.C.....		+	14,10,8
A.G.....C....		-	20,18,5
A.G.....		-	20,18
..G.....A..		+	18,3
....A.A.....G....		-	16,14,5
A.G.....C....		+	20,18,5
..G.....G..		-	18,3
A.G.A.....A..	<b>263</b>	-	20,18,16,3
A.G.....A..		-	20,18,3
..G.....G..		-	18,3
TC...T....C.....	<b>264</b>	+	20,19,14,8
A.G...A.....C....	<b>265</b>	-	20,18,13,5
..G.....A...A..		-	18,8,3
CA....A.....		+	20,19,13
A.G.....G..		-	20,18,3
..G.....G..		-	18,3
A.G...A.....A....	<b>266</b>	+	20,18,13,5
....A.G..A.....		+	16,14,10
..G...A.....C....		-	18,13,5
..AG.....C....		-	19,18,5
A.G.....C....		+	20,18,5
A.G...A.....C....	<b>267</b>	-	20,18,13,5
..G...A.....C....		-	18,13,5
CAG.....G..	<b>268</b>	+	20,19,18,3
A....TA.....G....	<b>269</b>	-	20,14,13,5
A.G.....G..		+	20,18,3
..G...A.....C....		+	18,13,5
A.G.....C....		-	20,18,5
T.G.....		+	20,18
A.G...A.....C....	<b>270</b>	-	20,18,13,5
..G.....G..		-	18,3
A.G...A.....		-	20,18,13
A.G.....G..		+	20,18,3
..G.....		-	18
..G...T.....C....		-	18,14,5
..G.....G..		-	18,3
A.G.....G..		+	20,18,3

..G.....G..		+	18,3
A.G.....		-	20,18
A.A.....G..		-	20,18,3
A.G.....G..		+	20,18,3
A.G.....C.C....	<b>271</b>	-	20,18,8,5
..G.....G..		-	18,3
..G.....C....		+	18,5
..G...A.....C....		+	18,13,5
A.G...A.....A..	<b>272</b>	+	20,18,13,3
..G.....G..		-	18,3
..A.....G..		-	18,3
A.G...A.....		-	20,18,13
..G...A.....		-	18,13
..A...T.A...C.....	<b>273</b>	+	18,14,12,8
A.A...A.....C....	<b>274</b>	-	20,18,13,5
A.T...A.....C....	<b>275</b>	+	20,18,13,5
..G...A.....		+	18,13
..G.....		+	18
A.G...A.....A..	<b>276</b>	+	20,18,13,3
A.G.....G..		-	20,18,3
..G...A.....		-	18,13
.C.....GT.C.....	<b>277</b>	-	19,11,10,8
A.G.....		+	20,18
T.G.....		-	20,18
..G...T.....		-	18,14
A.G...T.....A..	<b>278</b>	-	20,18,14,3
A.G.....G..		-	20,18,3
..G...A.....		-	18,13
A.G...A.....C....	<b>279</b>	+	20,18,13,5
..G...A.....G..		-	18,13,3
..G.....		-	18
..T...AA.....CA	<b>280</b>	+	18,13,12,2,1
..G.....G..		-	18,3
A.G...A.....		-	20,18,13
..G...A.....G..		+	18,13,3
..G.....G..		-	18,3
..G...A.....G..		+	18,13,3
..G.....		+	18
..G...T.....C....		+	18,14,5
CAG...A.....	<b>281</b>	-	20,19,18,13
..G...A.....		+	18,13
..G...A.....		-	18,13
A.A.....G..		+	20,18,3

C.TCT...AG...AC..G..	<b>282</b>	-	20,18,17,16,12,11,7,6,3
..G...A.....		+	18,13
..G.A.....G..		-	18,15,3
A.G...A.....		+	20,18,13
A.G.....		-	20,18
A.A...A.....		-	20,18,13
.C....T...A.....		+	19,13,8
A.G.....		-	20,18
..G...A.....G..		-	18,13,3
<b>n.PAM.mismatch</b>	<b>n.guide.mismatch</b>	<b>PAM.sequence</b>	
0	0	GGG	
0	2	AGG	
0	1	GGG	
0	3	AGG	
0	2	AGG	
0	3	TGG	
0	3	AGG	
1	2	AGC	
0	2	AGG	
0	3	GGG	
0	3	AGG	
0	2	AGG	
0	4	AGG	
0	3	AGG	
0	2	AGG	
0	4	AGG	
0	4	GGG	
0	3	AGG	
1	3	AGC	
0	3	AGG	
0	2	AGG	
0	4	AGG	
0	3	TGG	
0	3	AGG	
0	3	AGG	
0	3	AGG	
0	4	AGG	
0	3	AGG	
0	4	AGG	
0	4	AGG	
0	3	AGG	
0	3	AGG	

0	3	AGG
1	2	AGC
0	4	AGG
0	2	AGG
1	3	AGC
0	3	AGG
1	1	AGC
0	3	AGG
0	2	AGG
0	3	AGG
0	2	AGG
1	2	AGC
0	3	AGG
0	3	AGG
0	4	AGG
0	2	AGG
0	2	AGG
0	3	AGG
0	4	AGG
0	2	AGG
0	2	AGG
1	3	AGC
1	2	AGC
0	4	CGG
0	4	AGG
0	4	AGG
1	2	AGC
1	1	AGC
0	4	AGG
0	3	AGG
1	2	AGC
0	4	GGG
1	2	GGC
1	2	AGC
1	2	AGC
0	4	AGG
0	3	AGG
1	2	AGC
0	4	AGG
0	3	AGG
1	1	AGC
0	5	GGG
0	2	AGG

1	3	AGC
0	3	AGG
0	2	AGG
0	3	AGG
1	1	AGC
0	3	AGG
1	4	AGC
1	2	AGT
1	2	AGC
0	3	AGG
1	9	AGA
1	2	AGC
0	3	AGG
1	3	AGC
1	2	AGC
1	3	AGC
0	3	TGG
1	2	AGC
0	3	AGG
<b>offTarget_Start</b>	<b>offTarget_End</b>	<b>chromosome</b>
43748587	43748609	chr6
82004618	82004640	chr1
31140567	31140589	chr1
30357052	30357074	chr16
33453895	33453917	chr5
116600352	116600374	chr11
46938649	46938671	chr17
130859778	130859800	chr9
59837681	59837703	chr15
19135541	19135563	chr22
49057600	49057622	chrX
72751388	72751410	chr7
51652045	51652067	chr3
9544334	9544356	chr1
47868006	47868028	chr3
140670069	140670091	chr9
149516035	149516057	chr2
18245713	18245735	chr22
154744438	154744460	chr3
73320669	73320691	chr17
38479457	38479479	chr1
33058792	33058814	chr7

108299833	108299855	chr9
23627429	23627451	chr1
63393272	63393294	chr2
71467786	71467808	chr16
111638773	111638795	chr1
213393740	213393762	chr1
38284425	38284447	chr7
134511606	134511628	chr7
152293366	152293388	chr7
60243345	60243367	chr17
48007735	48007757	chrX
52768707	52768729	chr1
38805324	38805346	chr19
41283776	41283798	chrX
14539718	14539740	chr11
32895093	32895115	chr6
138957343	138957365	chr7
63900682	63900704	chr3
79624954	79624976	chr5
76012229	76012251	chr7
39889198	39889220	chrX
99897525	99897547	chr4
25822709	25822731	chr1
17293204	17293226	chr5
66697991	66698013	chr13
80796103	80796125	chr5
49239128	49239150	chr16
69489884	69489906	chr3
113712655	113712677	chr8
24502672	24502694	chr2
65642349	65642371	chr7
135700076	135700098	chrX
99795756	99795778	chr1
1821377	1821399	chr19
75501534	75501556	chr4
74828740	74828762	chr18
133975784	133975806	chrX
55717904	55717926	chr14
49522615	49522637	chr13
77788415	77788437	chr3
48230825	48230847	chr11
1280441	1280463	chr1
44602379	44602401	chr7

108166294	108166316	chr12
111929850	111929872	chr7
122404237	122404259	chr12
79123453	79123475	chr12
46412541	46412563	chr22
93889070	93889092	chr5
97776548	97776570	chr10
56533335	56533357	chr2
149843401	149843423	chr3
232769157	232769179	chr1
75100050	75100072	chr15
37252965	37252987	chr18
44506208	44506230	chr2
182389352	182389374	chr4
9360929	9360951	chr11
23638452	23638474	chr12
66498753	66498775	chr7
32055862	32055884	chr13
59331986	59332008	chr15
126196868	126196890	chr2
77359566	77359588	chrX
24652788	24652810	chrX
17667857	17667879	chr17
34751155	34751177	chr21
48734975	48734997	chr2
69755048	69755070	chr1
90013282	90013304	chr16
630757	630779	chr18
163905630	163905652	chr3
<b>inExon</b>	<b>entrez_id</b>	<b>symbol</b>
TRUE	7422	VEGFA
-	23266	ADGRL2
-	-	-
-	-	-
-	6897	TARS
-	-	-
-	10241	CALCOCO2
-	114789	SLC25A25
-	-	-
-	-	-
-	-	-
-	8468	FKBP6

-	23132	RAD54L2
-	-	-
-	22907	DHX30
-	79813	EHMT1
-	26122	EPC2
-	637	BID
-	4311	MME
-	2885	GRB2
-	51118	UTP11
-	51251	NT5C3A
-	83856	FSD1L
-	-	-
-	51057	WDPCP
-	-	-
-	-	-
-	26750	RPS6KC1
-	445347	TARP
-	800	CALD1
-	-	-
-	-	-
-	-	-
-	9372	ZFYVE9
-	90522	YIF1B
-	-	-
-	5682	PSMA1
-	-	-
-	254048	UBN2
-	6314	ATXN7
-	-	-
-	-	-
-	-	-
-	-	-
-	55219	MACO1
-	-	-
-	-	-
-	23635	SSBP2
-	-	-
-	23150	FRMD4B
-	114788	CSMD3
-	50618	ITSN2
-	-	-
-	-	-
-	-	-

-	57455	REXO1
-	-	-
-	4155	MBP
-	159091	FAM122C
-	-	-
-	-	-
-	-	-
-	-	-
-	1855	DVL1
-	-	-
-	-	-
-	11179	ZNF277
-	144406	WDR66
-	-	-
-	-	-
-	285600	KIAA0825
-	728558	ENTPD1-AS1
-	114800	CCDC85A
-	-	-
-	-	-
-	-	-
-	647946	MIR924HG
-	6519	SLC3A1
-	-	-
-	-	-
-	-	-
-	55253	TYW1
-	-	-
-	54778	RNF111
-	-	-
-	-	-
-	9468	PCYT1B
-	10743	RAI1
-	-	-
-	129285	PPP1R21
-	-	-
-	-	-
-	27098	CLUL1

**SpyDS4 (gRNA.name SpyDS4)**

<b>predicted_cleavage_score</b>	<b>SEQ ID NOS:</b>	<b>gRNAPlusPAM</b>
---------------------------------	--------------------	--------------------



TGG		41551021	41551043
GGC		43748564	43748586
CGG		77359654	77359676
CCC		43741999	43742021
TTT		68132445	68132467
AGG		77359345	77359367
CCC		22774978	22775000
GCC		77359596	77359618
GAG		82004622	82004644
GTG		80003891	80003913
<b>chromosome</b>			
		<b>inExon</b>	<b>entrez_id</b>
chr6		NA	7422
chr22		TRUE	2033
chr6		TRUE	7422
chrX		TRUE	5230
chr6		-	7422
chr15		-	-
chrX		-	-
chr6		-	-
chrX		-	-
chr1		-	23266
chr12		-	5074
			<b>symbol</b>
			VEGFA
			EP300
			VEGFA
			PGK1
			VEGFA
			-
			-
			-
			-
			ADGRL2
			PAWR

**SpyDS6 (gRNA.name SpyDS6)**

<b>offTarget</b>	<b>peak_score</b>	<b>predicted_cleavage_score</b>
chr6+:80816457:80816479	699	0.2
chr6:-:22774975:22774997	553	1.4
chr6:-:43742023:43742045	458	100
chr7:-:124498153:124498175	449	0.2
chr1:-:79194307:79194329	386	0.2
chr17+:77835740:77835762	383	5.2
chr19+:15313634:15313656	382	0.7
chr12+:96650610:96650632	374	3.7
chr10:-:79681895:79681917	352	1.5
chr6+:20250488:20250510	338	0.2
chr13:-:49117083:49117105	334	0.1
chr12:-:80003893:80003915	330	0.1
chr17:-:77543039:77543061	302	1.6
chr8:-:65972642:65972664	299	0.1
chr20:-:35488683:35488705	277	2.4
chr11+:100275645:100275667	271	1.6
chr22+:38338356:38338378	268	0.4

chr13:+:45356854:45356876	255	0.2
chr20:-:31061319:31061341	231	1.9
chr11:-:66051111:66051133	229	0.7
chr17:-:72637693:72637715	225	2.4
chr11:+:128772408:128772430	198	0.5
chr1:-:99257317:99257339	172	0.1
chr15:-:39243269:39243291	171	0.3
chr14:-:22258408:22258430	170	0.2
chr21:-:42506703:42506725	166	2.1
chr7:-:150036050:150036072	163	0.2
chr7:-:1140569:1140591	162	1.5
chr4:+:40239842:40239864	154	0.4
chr22:-:50743552:50743574	151	0.9
chr2:-:241904500:241904522	149	3.1
chr9:-:136776149:136776171	146	1
chr8:+:22487688:22487710	145	0.3
chr1:-:110032844:110032866	144	4.6
chr1:-:182626625:182626647	133	0.9
chr5:-:134908150:134908172	127	0.1
chr20:-:61928182:61928204	123	2.1
chr10:-:88042752:88042774	120	0.4
chr17:-:6131626:6131648	118	0.2
chr4:-:1002743:1002765	117	0.2
chr22:+:19106203:19106225	115	0.2
chr1:+:44003969:44003991	114	1.5
chr1:-:114792469:114792491	110	0.4
chr19:+:38997988:38998010	110	1.6
chr2:-:46897354:46897376	109	0.1
chr12:+:121011672:121011694	108	0.1
chr17:-:75891020:75891042	105	0.6
chr9:+:139220931:139220953	98	5.7
chr14:+:24168625:24168647	97	0.1
chr15:-:74949775:74949797	92	0.1
chr19:+:44199443:44199465	86	0.4
chr12:+:75214528:75214550	85	0.3
chr17:-:46058760:46058782	82	0.2
chr16:-:90077745:90077767	80	1.3
chr20:+:62023611:62023633	79	2.1
chr12:+:121013758:121013780	77	0.1
chrX:+:106755923:106755945	75	1
chr10:+:44417540:44417562	73	0.3
chr11:-:118193407:118193429	73	1.4
chr16:-:13411476:13411498	73	0.2
chr4:-:8206405:8206427	73	0.6
chr16:+:1517259:1517281	71	0.1
chr1:-:150849202:150849224	69	0
chr19:-:2057711:2057733	69	1.1
chr9:-:136075308:136075330	69	0.1
chr12:+:29935821:29935843	67	0.1

chr11:-:70812278:70812300	66	0.2
chr13:-:89703965:89703987	62	2.3
chr1+:110166721:110166743	60	0.6
chr11:-:114079332:114079354	58	0.2
chr10:-:71813737:71813759	57	0.4
chr19:-:17414518:17414540	56	0.3
chr3:-:184289395:184289417	56	0.3
chr14:-:94566714:94566736	55	0.2
chr5+:178665449:178665471	55	0.1
chr5+:149568491:149568513	54	0.3
chr11:-:70242709:70242731	52	0.1
chr21:-:45132035:45132057	52	0.1
chr17:-:827977:827999	47	0.2
chr18:-:35056448:35056470	44	0
chr6+:12990616:12990638	44	0.2
chr8:-:17955569:17955591	44	0.1
chr1:-:148932239:148932261	43	0.3
chr19:-:32734619:32734641	42	0.2
chr1+:228330887:228330909	41	0.1
chr3+:140221489:140221511	41	3.3
chr5:-:139938346:139938368	40	0.2
chr22+:23744878:23744900	39	1.6
chr10:-:16388635:16388657	38	0.1
chr17+:34824953:34824975	35	0.1
chr3:-:129656921:129656943	35	0.4
chr14:-:93351573:93351595	34	0
chr1+:33169255:33169277	33	0.1
chr18+:29253123:29253145	33	0
chr6+:20984444:20984466	33	0
chr10:-:77256682:77256704	32	2.5
chr15+:89196634:89196656	32	1.7
chr18:-:73391522:73391544	32	0
chr10:-:72512814:72512836	31	0
chr8+:80980935:80980957	31	0.1
chr11+:37704008:37704030	30	0.1
chr12+:52539310:52539332	29	1.7
chr14:-:56431001:56431023	27	0.3
chr15:-:66949803:66949825	26	1.6
chr7+:100879116:100879138	26	98.6
chr11:-:94784149:94784171	25	0.2
chr12+:111548733:111548755	25	0.8
chr19+:2212199:2212221	25	0.2
chr13:-:22824517:22824539	22	0.1
chr13:-:84196623:84196645	19	0.2
chr15+:96534271:96534293	19	0.1
chr21:-:21304105:21304127	17	0.2
chr17:-:39705337:39705359	16	0.2
chr20:-:56582544:56582566	15	0.9
chr20+:49479068:49479090	15	0.1

chr1+:89258185:89258207		14	0.1
chr15:-:51386687:51386709		13	0.1
chr19+:38724286:38724308		13	0.3
chr16:-:2286384:2286406		11	0.2
SEQ ID NOS.	gRNAPlusPAM	SEQ ID NOS.	offTarget_sequence
305	GGGCAGGGGCTGGGGTGCACNNG	306	CGGCAGGGGCTGAGGGGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	307	GGGTAGGAGCAGGGGTGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	308	GGGCAGGGGCTGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	309	GGGCAGGAACTGGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	310	GGGCAGGAACTGGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	311	CAGCAGGGGCTGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	312	GGGAAGGGCCTGGGGTACACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	313	GGGCCGGGGCAGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	314	AGACAGGGGCCGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	315	GGGCAGGAACTGGAGTGCACCGG
305	GGGCAGGGGCTGGGGTGCACNNG	316	AGGCAGGAACTGGAGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	317	AGGCAGGAACTGGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	318	AGGAAGGGACTGGGGTGCACACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	319	AGGCAGGAACTGGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	320	AGGTGGGGGCTGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	321	AGGCAGGAACTGGGGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	322	TGGCAGGGGCAGGGGTGAACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	323	GGGCAGGAACTGGAGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	324	GGCCAGGGGCTGGGGAGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	325	GGGCAGGGCTGGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	326	TGGGTGGGGCTGGGGTGCACACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	327	GGGAGGGGGCTGGGGAGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	328	GGGCAGGAACTGGAGTACACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	329	GAGAAGGAGCTGGGGAGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	330	GGGCAGGAACTGGAGTGCACCAG
305	GGGCAGGGGCTGGGGTGCACNNG	331	GGGCAAGGGCAGGGGTGCACCAG
305	GGGCAGGGGCTGGGGTGCACNNG	332	AAGAAGGGGCAAGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	333	GGCCAGGAGCAGGGGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	334	TGGCAGCGGCTGGGGAGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	335	GGGCGTGGGCAGGGGTGCACACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	336	GGGCAGTGGCTGGGGTGCATTGG
305	GGGCAGGGGCTGGGGTGCACNNG	337	GGCCAGGAGCTGGGGTGCTCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	338	CCTCAGGGGCTGGGGTGAACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	339	TGGCAGGGTCTGGGGTGCACAGA
305	GGGCAGGGGCTGGGGTGCACNNG	340	GAGCAGGGTCTGGGGTGCATGGG
305	GGGCAGGGGCTGGGGTGCACNNG	341	GAGCAGGGACTGAGGGGCACAGG

305	GGGCAGGGGCTGGGGTGCACNNG	342	GAGCAGGGGCTGGGGGGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	343	TGGCAGGGGTAAGGGTGCACCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	345	AGACAGAGGCTGGAGTGCACCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	346	AGGCAGGGGCTGGAGTTCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	347	AGGAAGGGACCAGGGTGCACCAG
305	GGGCAGGGGCTGGGGTGCACNNG	348	GGCCAGGAGCAGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	349	GGGCCGGGGCTGGGGTGCCAGGG
305	GGGCAGGGGCTGGGGTGCACNNG	350	GGGCGGGGGCTGGGGAGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	351	AGGCAGGAGCCAGGGTGCAGAGG
305	GGGCAGGGGCTGGGGTGCACNNG	352	GGGCAGAGGCTGGAGTGCCCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	353	GGACAGGGGCAGGGGTGCCCGGG
305	GGGCAGGGGCTGGGGTGCACNNG	354	AGGGAGGGGCTGGGGTGCACGGA
305	GGGCAGGGGCTGGGGTGCACNNG	355	GGGCAGGAACTGGAGTGCATAGG
305	GGGCAGGGGCTGGGGTGCACNNG	356	AGGCAGGAACTGGAGTGCACAAG
305	GGGCAGGGGCTGGGGTGCACNNG	357	GGGCAGAGGCTAGGGTGCAGTGG
305	GGGCAGGGGCTGGGGTGCACNNG	358	AGGTAGGGGTTGGGGGGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	359	GGGCAGAAGCAGGGGTGCTCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	360	GGGGAGGGGTGGGGGTGCACCGG
305	GGGCAGGGGCTGGGGTGCACNNG	361	GAGCAGGGGCTGGGGGGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	362	GGGCAGAGGCTGGAGTGCCCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	363	GGGTGGGGGCTGGGGTGCCCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	364	GGGCAAGGGCAGGGGTGCCCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	365	GAGAGGGAGCTGGGGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	366	AGGCAGGGACTGAGGTGCATAGG
305	GGGCAGGGGCTGGGGTGCACNNG	367	GGGCCAGGGCTGAGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	368	TGGGAGGGGCTAGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	369	CCGCAGGGGCTGGGATGCTGGGG
305	GGGCAGGGGCTGGGGTGCACNNG	371	GAGGAGGGGCTGGGGTGCCCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	372	GGGCAAAGGCCGGGGTGCCCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	373	AGGCGGGGGCTGGGGGGCTCGGG
305	GGGCAGGGGCTGGGGTGCACNNG	374	AGGCAGGGGCCAGGGTCCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	376	GGGTTGGGGTTGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	377	AGGCAGGGGCCGGGGTGCGCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	378	GGGCACAGACTGGGGTGCATTGG
305	GGGCAGGGGCTGGGGTGCACNNG	379	GGGCTGGGGCTGAGGTGCGCCGG
305	GGGCAGGGGCTGGGGTGCACNNG	380	AGGCAGGGGCTGGGGGGCAAGGG
305	GGGCAGGGGCTGGGGTGCACNNG	381	GAGCGGGAGCTGGGGGGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	382	GGGCAGGGACTGGGGTGCTTAGG
305	GGGCAGGGGCTGGGGTGCACNNG	383	GGGAAGGGGCTGGAGGGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	384	GGGCAGGGGAAGGGGTGGACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	385	AGACAGGGGCTGGAGTGCAGTGG
305	GGGCAGGGGCTGGGGTGCACNNG	386	GGGCAGAGGCTGGAGTGCAATGG
305	GGGCAGGGGCTGGGGTGCACNNG	387	GGGCTGGGGCTGGGGAGCAGGGG

305	GGGCAGGGGCTGGGGTGCACNNG	388	AGGAAAGGGCTGGAGTGCAGGGG
305	GGGCAGGGGCTGGGGTGCACNNG	389	GGGCAGGAACTGGAGTGCACCAG
305	GGGCAGGGGCTGGGGTGCACNNG	390	AGGCAGGAACTGGAGTGCACAAG
305	GGGCAGGGGCTGGGGTGCACNNG	391	AGGCAGAGCCTGGGGTGCAGGGG
305	GGGCAGGGGCTGGGGTGCACNNG	392	GGGCAGGGCCAGGGGAGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	393	AGCCAGGGGCTGGGGGGAACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	394	GGGCAGGGGATGGGGTGCAGTGG
305	GGGCAGGGGCTGGGGTGCACNNG	395	AGGCAAGGCCTGGGGTGCCAGG
305	GGGCAGGGGCTGGGGTGCACNNG	396	GGGCTGGGGCTGGGGAGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	397	GAGAAGGGGCTGGGAAGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	398	AGGCAGGAACTGGAGTGCACAAG
305	GGGCAGGGGCTGGGGTGCACNNG	399	GGGGAGGGGCTGGGGTGCCAGGG
305	GGGCAGGGGCTGGGGTGCACNNG	400	AGGAAGGGGCTGGGGAAAACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	401	CCCAGGGGCTGGGGTGCCTGGG
305	GGGCAGGGGCTGGGGTGCACNNG	402	AAGCAGAGGCTGAAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	403	AGGCAGGAACTAGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	404	GGGCAGGGGTGGGGTCCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	406	GGGGAGGGGCTGGGGAGCACGGA
305	GGGCAGGGGCTGGGGTGCACNNG	407	AGGCAGAGGCTGGAGTGGACCGG
305	GGGCAGGGGCTGGGGTGCACNNG	408	GGGTAGGGGCTGGGGGATACCGG
305	GGGCAGGGGCTGGGGTGCACNNG	409	GGGAAGGGTCTGGAGTCCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	410	GGGCAGGAACTAGAGTGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	411	GGGCAGGGACTGGGGTGCTCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	412	GAGTAGGGGCAGGGGTGCTCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	413	AGGAAGGGCCTGGGGTGCACAGA
305	GGGCAGGGGCTGGGGTGCACNNG	414	GGCCAGGGGCTGGGGTGCACGGT
305	GGGCAGGGGCTGGGGTGCACNNG	415	AGGCAGGGGCCAGGGTGCATGGG
305	GGGCAGGGGCTGGGGTGCACNNG	416	GGGCAGAGGATGGGGTGCAGGGG
305	GGGCAGGGGCTGGGGTGCACNNG	417	CGGCAGGGGCTGGAGTGCAGTGG
305	GGGCAGGGGCTGGGGTGCACNNG	418	AGGCAGGATCTGGAGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	419	AGACAGGAGCTGGAGTGCACAAG
305	GGGCAGGGGCTGGGGTGCACNNG	420	TGGCAGGGGCAGGGATGCTCTGG
305	GGGCAGGGGCTGGGGTGCACNNG	421	CCTCAGGGGTGGGATGCACTGG
305	GGGCAGGGGCTGGGGTGCACNNG	422	GAGCAGGGTCAGGGGTGCAGAGG
305	GGGCAGGGGCTGGGGTGCACNNG	423	CAGGAGTGGCTGGGGTGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	424	GGCCTGGGCTGAGATGCACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	425	ACGCAGGGGCTAGGGAGCACAAAG
305	GGGCAGGGGCTGGGGTGCACNNG	426	GGGCTGGGGCTGGGGAGGACGGG
305	GGGCAGGGGCTGGGGTGCACNNG	427	GGGCAGGGATTGGGGGGCACAGG
305	GGGCAGGGGCTGGGGTGCACNNG	428	AGGGAGGGGCCGGGCTGCACTGG
<b>SEQ ID</b>	<b>guideAlignment2OffTarget</b>	<b>offTargetStrand</b>	<b>mismatch.distance2PAM</b>

NOS.			
	C.....A.G....	+	20,8,5
	...T..A..A.....	-	17,13,10
	.....	-	-
	.....AA...A.....	-	13,12,7
	.....AA...A.....	-	13,12,7
	CA.....	+	20,19
	...A...C.....A...	+	17,12,4
	...C...A.....	+	16,10
	A.A.....C.....	-	20,18,10
	.....AA...A.....	+	13,12,7
429	A.....AA...A.....	-	20,13,12,7
430	A.....AA...A.....	-	20,13,12,7
	A..A...A.....	-	20,17,12
431	A.....AA...A.....	-	20,13,12,7
	A..TG.....	-	20,17,16
	A.....AA.....	+	20,13,12
	T.....A.....A..	+	20,10,3
	.....AA...A.....	+	13,12,7
	..C.....A....	-	18,5
	.....CTG.....	-	12,11,10
	T..GT.....	-	20,17,16
	...AG.....A....	+	17,16,5
432	.....AA...A..A..	-	13,12,7,4
433	..A.A..A.....A....	-	19,17,13,5
	.....AA...A.....	-	13,12,7
	.....A...A.....	-	15,10
434	AA.A.....AA.....	-	20,19,17,10,9
	..C...A..A.....	-	18,13,10
	T....C.....A....	+	20,14,5
	...GT...A.....	-	16,15,10
	.....T.....T	-	14,1
	..C...A.....T.	-	18,13,2
435	CCT.....A..	+	20,19,18,3
	T.....T.....	-	20,12
	..A.....T.....T	-	19,12,1
436	..A.....A..A..G....	-	19,12,8,5
	..A.....G....	-	19,5
437	T.....TAA.....	-	20,11,10,9
438	A.A...A.....A.....	-	20,18,14,7
	A.....A..T..	-	20,7,4
439	A..A...A.CA.....	+	20,17,12,10,9
	..C...A..A.....	+	18,13,10

	...C.....CA	-	16,2,1
	...G.....A...	+	16,5
<b>440</b>	A.....A..CA.....G	-	20,13,10,9,1
	.....A.....A...C.	+	14,7,2
	..A.....A.....C.	-	18,10,2
	A..G.....	+	20,17
<b>441</b>	.....AA...A....T	+	13,12,7,1
<b>442</b>	A.....AA...A.....	-	20,13,12,7
	.....A...A.....G	+	14,9,1
<b>443</b>	A..T....T....G....	+	20,17,11,5
<b>444</b>	.....AA..A.....T.	-	14,13,10,2
	...G....TG.....	-	17,11,10
	..A.....G....	+	19,5
	.....A.....A...C.	+	14,7,2
	...TG.....C.	+	17,16,2
	.....A...A.....C.	+	15,10,2
<b>445</b>	..AAG..A.....	-	19,17,16,13
<b>446</b>	A.....A..A....T	-	20,12,8,1
	...CA.....A.....	-	16,15,8
<b>447</b>	T..G.....AA.....	+	20,17,9,7
	CC.....A..TG	-	20,19,6,2,1
	..AG.....C.	-	19,17,2
<b>448</b>	.....AA...C.....C.	-	15,14,10,2
<b>449</b>	A..G.....G..T.	+	20,16,5,2
<b>450</b>	A.....CA...C...	-	20,10,9,4
	...TT...T.....	-	17,16,11
	A.....C.....G.	+	20,10,2
<b>451</b>	.....CA..A.....T	-	15,14,12,1
	...T.....A...G.	-	16,8,2
	A.....G..A	-	20,5,1
<b>452</b>	..A..G..A.....G....	-	19,16,13,5
	.....A.....TT	-	12,2,1
	...A.....A.G....	+	17,7,5
	.....AA.....G..	+	11,10,3
<b>453</b>	A.A.....A...G	-	20,18,7,1
	.....A.....A...A	-	14,7,1
	...T.....A..G	-	16,5,1
<b>454</b>	A..A.A.....A...G	-	20,17,15,7,1
	.....AA...A.....	+	13,12,7
<b>455</b>	A.....AA...A.....	-	20,13,12,7
<b>456</b>	A.....A.C.....G	-	20,14,12,1
	.....C.A...A....	-	12,10,5
<b>457</b>	A.C.....G.A..	+	20,18,5,3

	.....A.....G	+	11,1
458	A...A.C.....C.	-	20,15,12,2
	...T.....A...	+	16,5
459	.AA.....AA...	-	19,17,6,5
460	A.....AA..A.....	+	20,13,12,7
	...G.....CA	-	17,2,1
461	A..A.....AAA..	-	20,17,5,4,3
462	CCC.....CT	+	20,19,18,2,1
463	AA...A...AA.....	+	20,19,14,8,7
464	A.....AA..AA.....	+	20,13,12,9,7
	.....G.....C...	-	11,4
	...G.....A...	+	17,5
465	A...A...A..G..	-	20,14,7,3
466	...T.....GAT..	-	17,5,4,3
467	...A...T...A..C...	+	17,12,7,4
468	.....AA..AA.....	+	13,12,9,7
	.....A.....T.	+	12,2
469	.AT.....A.....T.	-	19,17,10,2
	A..A...C.....	-	20,17,12
	..C.....	+	18
470	A.....CA.....T	-	20,10,9,1
	.....A..A.....G	+	14,11,1
	C.....A...G	+	20,7,1
471	A.....AT...A.....	-	20,13,12,7
472	A.A...A...A.....	-	20,18,13,7
473	T.....A...A..T.	+	20,10,6,2
474	CCT.....T...A.....	-	20,19,18,11,6
475	.A.....TA.....G	-	19,12,10,1
476	CA.G.T.....	-	20,19,17,14
477	...CT.....AA.....	+	16,15,8,6
478	AC.....A..A...	+	20,19,9,5
	...T.....A.G..	-	16,5,3
	.....AT....G...	+	12,11,5
479	A..G.....C...C.....	-	20,17,10,6
<b>n.PAM.mismatch</b>			
<b>n.guide.mis match</b>		<b>PAM.sequence</b>	
0	3	TGG	
0	3	TGG	
0	0	AGG	
0	3	AGG	
0	3	AGG	
0	2	AGG	

0	3	GGG
0	2	AGG
0	3	AGG
0	3	CGG
0	4	GGG
0	4	AGG
0	3	TGG
0	4	AGG
0	3	AGG
0	3	GGG
0	3	TGG
0	3	GGG
0	2	AGG
0	3	AGG
0	3	TGG
0	3	AGG
0	4	GGG
0	4	TGG
1	3	CAG
1	2	CAG
0	5	AGG
0	3	GGG
0	3	TGG
0	3	TGG
0	2	TGG
0	3	AGG
0	4	AGG
1	2	AGA
0	3	GGG
0	4	AGG
0	2	TGG
0	4	TGG
0	4	TGG
0	3	AGG
1	5	CAG
0	3	AGG
0	3	GGG
0	2	AGG
0	5	AGG
0	3	AGG
0	3	GGG
1	2	GGA
0	4	AGG

1	4	AAG
0	3	TGG
0	4	AGG
0	4	AGG
0	3	CGG
0	2	TGG
0	3	AGG
0	3	AGG
0	3	TGG
0	4	GGG
0	4	AGG
0	3	AGG
0	4	AGG
0	5	GGG
0	3	TGG
0	4	AGG
0	4	GGG
0	4	AGG
0	3	AGG
0	3	AGG
0	4	TGG
0	3	CGG
0	3	GGG
0	4	AGG
0	3	AGG
0	3	AGG
0	3	TGG
0	4	TGG
0	3	TGG
0	3	GGG
0	5	GGG
1	3	CAG
1	4	AAG
0	4	GGG
0	3	AGG
0	4	AGG
0	2	TGG
0	4	AGG
0	2	GGG
0	4	AGG
1	4	AAG
0	3	GGG
0	5	AGG

0	5	GGG
0	5	AGG
0	5	AGG
0	2	AGG
1	2	GGA
0	4	CGG
0	4	CGG
0	4	TGG
0	4	GGG
0	2	TGG
0	4	TGG
1	3	AGA
1	1	GGT
0	4	GGG
0	3	GGG
0	3	TGG
0	4	AGG
1	4	AAG
0	4	TGG
0	5	TGG
0	4	AGG
0	4	AGG
0	4	GGG
1	4	AAG
0	3	GGG
0	3	AGG
0	4	TGG
<b>offTarget_Start</b>	<b>offTarget_End</b>	<b>chromosome</b>
80816457	80816479	chr6
22774975	22774997	chr6
43742023	43742045	chr6
124498153	124498175	chr7
79194307	79194329	chr1
77835740	77835762	chr17
15313634	15313656	chr19
96650610	96650632	chr12
79681895	79681917	chr10
20250488	20250510	chr6
49117083	49117105	chr13
80003893	80003915	chr12
77543039	77543061	chr17

65972642	65972664	chr8
35488683	35488705	chr20
100275645	100275667	chr11
38338356	38338378	chr22
45356854	45356876	chr13
31061319	31061341	chr20
66051111	66051133	chr11
72637693	72637715	chr17
128772408	128772430	chr11
99257317	99257339	chr1
39243269	39243291	chr15
22258408	22258430	chr14
42506703	42506725	chr21
150036050	150036072	chr7
1140569	1140591	chr7
40239842	40239864	chr4
50743552	50743574	chr22
241904500	241904522	chr2
136776149	136776171	chr9
22487688	22487710	chr8
110032844	110032866	chr1
182626625	182626647	chr1
134908150	134908172	chr5
61928182	61928204	chr20
88042752	88042774	chr10
6131626	6131648	chr17
1002743	1002765	chr4
19106203	19106225	chr22
44003969	44003991	chr1
114792469	114792491	chr1
38997988	38998010	chr19
46897354	46897376	chr2
121011672	121011694	chr12
75891020	75891042	chr17
139220931	139220953	chr9
24168625	24168647	chr14
74949775	74949797	chr15
44199443	44199465	chr19
75214528	75214550	chr12
46058760	46058782	chr17
90077745	90077767	chr16
62023611	62023633	chr20
121013758	121013780	chr12

106755923	106755945	chrX
44417540	44417562	chr10
118193407	118193429	chr11
13411476	13411498	chr16
8206405	8206427	chr4
1517259	1517281	chr16
150849202	150849224	chr1
2057711	2057733	chr19
136075308	136075330	chr9
29935821	29935843	chr12
70812278	70812300	chr11
89703965	89703987	chr13
110166721	110166743	chr1
114079332	114079354	chr11
71813737	71813759	chr10
17414518	17414540	chr19
184289395	184289417	chr3
94566714	94566736	chr14
178665449	178665471	chr5
149568491	149568513	chr5
70242709	70242731	chr11
45132035	45132057	chr21
827977	827999	chr17
35056448	35056470	chr18
12990616	12990638	chr6
17955569	17955591	chr8
148932239	148932261	chr1
32734619	32734641	chr19
228330887	228330909	chr1
140221489	140221511	chr3
139938346	139938368	chr5
23744878	23744900	chr22
16388635	16388657	chr10
34824953	34824975	chr17
129656921	129656943	chr3
93351573	93351595	chr14
33169255	33169277	chr1
29253123	29253145	chr18
20984444	20984466	chr6
77256682	77256704	chr10
89196634	89196656	chr15
73391522	73391544	chr18
72512814	72512836	chr10

80980935	80980957	chr8
37704008	37704030	chr11
52539310	52539332	chr12
56431001	56431023	chr14
66949803	66949825	chr15
100879116	100879138	chr7
94784149	94784171	chr11
111548733	111548755	chr12
2212199	2212221	chr19
22824517	22824539	chr13
84196623	84196645	chr13
96534271	96534293	chr15
21304105	21304127	chr21
39705337	39705359	chr17
56582544	56582566	chr20
49479068	49479090	chr20
89258185	89258207	chr1
51386687	51386709	chr15
38724286	38724308	chr19
2286384	2286406	chr16
TRUE	594	BCKDHB
-	-	-
-	7422	VEGFA
-	25913	POT1
-	-	-
-	-	-
-	-	-
-	2004	ELK3
-	9231	DLG5
-	-	-
-	-	-
-	5074	PAWR
-	-	-
-	-	-
-	140710	SOGA1
-	-	-
TRUE	85377	MICALL1
-	-	-
-	140688	NOL4L
TRUE	254263	CNIH2
-	-	-
TRUE	3762	KCNJ5

-	-	-
-	-	-
-	-	-
-	-	-
-	5919	RARRES2
-	84310	C7orf50
-	399	RHOH
-	23654	PLXNB2
-	200772	LOC200772
-	7410	VAV2
-	55909	BIN3
-	127002	ATXN7L2
-	85397	RGS8
-	9547	CXCL14
-	57642	COL20A1
-	2894	GRID1
-	-	-
-	-	-
-	9993	DGCR2
-	5792	PTPRF
-	-	-
-	6261	RYR1
-	-	-
-	9921	RNF10
-	-	-
TRUE	26102	DKFZP434A062
-	-	-
-	80153	EDC3
-	-	-
-	-	-
-	80279	CDK5RAP3
-	79007	DBNDD1
-	-	-
-	9921	RNF10
-	-	-
-	283033	LINC00841
-	-	-
-	-	-
-	54436	SH3TC1
-	1186	CLCN7
TRUE	405	ARNT
-	-	-
-	-	-

TRUE	83857	TMTC1
-	22941	SHANK2
-	-	-
-	271	AMPD2
-	7704	ZBTB16
-	55506	H2AFY2
-	-	-
-	2049	EPHB3
-	122509	IFI27L1
-	9509	ADAMTS2
-	-	-
-	-	-
-	-	-
-	64359	NXN
-	56853	CELF4
-	221692	PHACTR1
-	-	-
-	645166	LOC645166
-	-	-
-	2987	GUK1
-	64084	CLSTN2
TRUE	10307	APBB3
-	-	-
-	-	-
-	-	-
-	-	-
-	-	-
-	9331	B4GALT6
-	54901	CDKAL1
-	-	-
-	3669	ISG20
-	-	-
-	140766	ADAMTS14
-	7163	TPD52
-	-	-
-	-	-
-	-	-
-	24146	CLDN15
-	-	-
-	23316	CUX2
TRUE	84444	DOT1L

-	-	-
-	-	-
-	-	-
-	-	-
-	3728	JUP
-	-	-
-	55653	BCAS4
-	5586	PKN2
-	388121	TNFAIP8L3
-	-	-
-	-	-

**Nme2DS2**

<b>offTarget</b>		<b>peak score</b>	<b>predicted cleavage score</b>	
chr6:-:43748582:43748613		547	100	
chrX+:77359550:77359581		44	0	
<b>gRNA.name</b>	<b>SEQ ID NOS:</b>	<b>gRNAPlusPAM</b>	<b>SEQ ID NOS;</b>	<b>offTarget_sequence</b>
Nme2DS2	<b>480</b>	GAATGGCAGGCGGAGG TTGTA CTGNNNNCCNN	<b>481</b>	GAATGGCAGGCGGAGGT TGTACTGGGGGCCAG
Nme2DS2	<b>480</b>	GAATGGCAGGCGGAGG TTGTA CTGNNNNCCNN	<b>482</b>	AAACGGAAGC CGCACGTCTC ACTAGTACCC TC
<b>SEQ ID NO:</b>	<b>guideAlignment2OffTarget</b>	<b>offTargetStrand</b>	<b>mismatch.distance2PAM</b>	
	.....	-	-	
<b>483</b>	A..C..A..C..C..CTC...A	+	24, 21, 18, 15, 12, 10, 7, 6, 5, 1	
<b>n.PAM.mismatch</b>	<b>n.guide.mismatch</b>	<b>PAM.sequence</b>		
0	0	GGGGCCAG		
0	10	GTACCCTC		
<b>offTarget Start</b>	<b>offTarget End</b>	<b>chromosome</b>		
43748582	43748613	chr6		
77359550	77359581	chrX		
<b>inExon</b>	<b>entrez_id</b>	<b>symbol</b>		
TRUE	7422	VEGFA		
-	-	-		

**Nme2DS4**

<b>offTarget</b>	<b>peak score</b>	<b>gRNA.name</b>
chr6:-	66	c_DeCas9_human_TS1

:43748843:43748874				4
<b>gRNAPlusPAM</b>				
<b>SEQ ID NO: 486</b>		GTGAGCAGGCACCTGTGCCAACATNNNNCCNN		
<b>offTarget_sequence</b>				<b>guideAlignment2Off Target</b>
<b>SEQ ID NO: 487</b>	GTGAGCAGGCACCTGTGCCAACATGGGCCCGC		.....	
<b>offTargetStrand</b>		<b>predicted_cleavage_score</b>	<b>mismatch.distance2PAM</b>	
-		100	-	
<b>n.PAM.mismatch</b>		<b>n.guide.mismatch</b>	<b>PAM.sequence</b>	
0		0	<b>SEQ ID NO: 488</b>	GGGC CCGC
<b>offTarget_Start</b>		<b>offTarget_End</b>	<b>chromosome</b>	
43748843		43748874	chr6	
<b>inExon</b>		<b>entrez_id</b>	<b>symbol</b>	
-		7422	VEGFA	

**Nme2DS6**

<b>offTarget</b>		<b>peak_score</b>		<b>predicted_cleavage_score</b>
chr6:-:43742018:43742049		483		100
chrX:-:77359465:77359496		12		0
<b>gRNA.name</b>		<b>gRNAPlusPAM</b>		
d_DeCas9_human_TS16		<b>SEQ ID NO: 489</b>	GCATGGGCAGGGGCTGGGGTGCAC NNNNCCNN	
d_DeCas9_human_TS16		<b>SEQ ID NO: 489</b>	GCATGGGCAGGGGCTGGGGTGCAC NNNNCCNN	
<b>offTarget_sequence</b>		<b>guideAlignment2Off Target</b>		<b>offTargetStrand</b>
<b>SEQ ID NO: 490</b>	GCATGGGCA GGGGCTGGG GTGCACAGG CCCAG	.....		-
<b>SEQ ID NO: 491</b>	GCAGGAAGC GTCGCCGGG GGGCCACA AGGGT	<b>SEQ ID NO: 492</b>	...G.AAGC.TC ..C....G..C.	21,19,18,17,16,14,13,10,5,2
<b>n.PAM.mismatch</b>		<b>PAM.sequence</b>		<b>PAM.sequence</b>
0		<b>SEQ ID</b>	AGGCCAG	<b>SEQ ID</b>   AATCCCTT

	<b>NO: 493</b>		<b>NO: 499</b>	
10	<b>SEQ ID NO: 494</b>	ACAAGGGT	<b>SEQ ID NO: 500</b>	ACTCCCTC
<b>offTarget_Start</b>				
<b>offTarget_End</b>				
<b>chromosome</b>				
43742018				
43742049				
77359465				
77359496				
chr6				
chrX				
<b>inExon</b>				
<b>entrez_id</b>				
<b>symbol</b>				
-				
7422				
-				
-				
VEGFA				
VEGFA				

**Rosa26**

<b>offTarget</b>		<b>peak_score</b>		<b>predicted_cleavage_score</b>	
chr6:- :113076072:113076103		1175		100	
chr11:- :73171296:73171327		24		1.4	
<b>gRNA.name</b>		<b>gRNAPlusPAM</b>			
Nme2Rosa		<b>SEQ ID NO: 495</b>	TGAGGACCGCCCTGGGCCTGGGAGNNNNCC NN		
Nme2Rosa		<b>SEQ ID NO: 495</b>	TGAGGACCGCCCTGGGCCTGGGAGNNNNCC NN		
<b>offTarget_sequence</b>			<b>guideAlignment2OffTarget</b>		<b>offTargetStr and</b>
<b>SEQ ID NO: 496</b>	TGAGGACCGCCCTGGG CCTGGGAGAATCCCTT			.....	-
<b>SEQ ID NO: 497</b>	GAAGGACCACCCTAGG CCTGGGAGACTCCCT		<b>SEQ ID NO: 498</b>	GA.....A...A..... .	-
<b>mismatch.distance2PAM</b>		<b>n.PAM.mismatch</b>		<b>n.guide.mismatch</b>	
-		0		0	
24, 23, 16, 11		0		4	
<b>PAM.sequence</b>		<b>offTarget_Start</b>		<b>offTarget_End</b>	
<b>SEQ ID NO: 499</b>	AATCCCTT		113076072		113076103
<b>SEQ ID NO: 500</b>	ACTCCCTC		73171296		73171327
<b>chromosome</b>		<b>inExon</b>		<b>entrez_id</b>	
chr6		-		14910	
chr11		-		94045	

**PCSK9**

<b>offTarget</b>	<b>peak_score</b>	<b>gRNA.name</b>
------------------	-------------------	------------------

chr4:-:106463720:106463751		266		Nme2PCSK9	
<b>gRNAPlusPAM</b>			<b>offTarget sequence</b>		
<b>SEQ ID NO: 501</b>	GGCCTGGCTGATGAGGCCG CACATNNNNCCNN	<b>SEQ ID NO: 502</b>	GGCCTGGCTGATGAGGCCG ACATGTGGCCAC		
<b>guideAlignment2OffTarget</b>	<b>offTargetStrand</b>	<b>predicted_cleavage_score</b>			
.....	-	100			
<b>mismatch.distance2PAM</b>	<b>n.PAM.mismatch</b>	<b>n.guide.mismatch</b>			
-	0	0			
<b>PAM.sequence</b>		<b>offTarget_Start</b>	<b>offTarget_End</b>		
<b>SEQ ID NO: 503</b>	GTGGCCAC	106463720	106463751		
<b>chromosome</b>		<b>inExon</b>	<b>entrez_id</b>		
chr4		TRUE	100102		

В отношении идентификации действия вне мишени, анализ показал, что для DS2, DS4 и DS6 sg-PHK SpyCas9 обуславливает редактирование в 93, 10 и 118 участках вне мишени, являющихся кандидатами, соответственно, с нормальным диапазоном участков вне мишени, когда редактирование на основе плазмиды SpyCas9 анализировали посредством GUIDE-seq (Fu et al., 2014; Tsai et al., 2014). Совершенно наоборот, оказалось, что для DS2, DS4 и DS6 sg-PHK Nme2Cas9 обуславливают редактирование в 1, 0 и 1 участков вне мишени, соответственно. Фигура 14С и таблица 2. По сравнению с количеством чтений посредством GUIDE-seq для участков SpyCas9 вне мишени, количество чтений для Nme2Cas9 было очень низким, далее указывая на то, что Nme2Cas9 является в высокой степени специфической. Фиг.13С против фиг.13D. Анализ GUIDE-seq для Nme2Cas9 с TS6, *Pcsk9* и *Rosa26* дал сходные результаты (0, 0 и 1 участков вне мишени, соответственно, с умеренным количеством чтений для участка вне мишени *Rosa26*-OT1). Фиг.13С, фиг.14D и таблица 2.

На фиг.14 представлены иллюстративные данные, демонстрирующие, что Nme2Cas9 демонстрирует небольшое или не демонстрирует поддающегося обнаружению нацеливания вне мишени в клетках млекопитающих.

На фиг.14А представлена иллюстративная схема, изображающая двойные участки (DS), на которые нацеливается как SpyCas9, так и Nme2Cas9, посредством их не

перекрывающихся PAM. PAM (оранжевый) Nme2Cas9 и PAM (синий) SpyCas9 выделены. Гидовая последовательность Nme2Cas9 из 24 нуклеотидов указана желтым; соответствующая гидовая последовательность для SpyCas9 на 4 нуклеотида короче на 5'-конце.

На фиг.14B представлены иллюстративные Nme2Cas9 и SpyCas9, обе из которых индуцируют инсерции-делеции в DS. Шесть DS в *VEGFA* (с последовательностями GN<sub>3</sub>GN<sub>19</sub>NGGNCC) было отобрано для прямых сравнений редактирования посредством двух ортологов. Плазмиды, экспрессирующие каждую Cas9 (с тем же промотором, линкерами, метками и NLS) и ее собственный гид трансфицировали в клетки HEK293T. Эффективность инсерций-делеций определяли посредством TIDE через 72 часа после трансфекции. Редактирование Nme2Cas9 выявлялось во всех шести участках и было незначительно или значительно более эффективным, чем у SpyCas9 в двух участках (DS2 и DS6, соответственно). SpyCas9 редактировала четыре из шести участков (DS1, DS2, DS4 и DS6), причем два участка демонстрировали значительно более высокую эффективность редактирования, чем Nme2Cas9 (DS1 и DS4). DS2, DS4 и DS6 были отобраны для анализа GUIDE-Seq, поскольку Nme2Cas9 была в равной степени эффективной, менее эффективной и более эффективной, чем SpyCas9, соответственно, в этих участках. На фиг.14C представлено иллюстративное редактирование генома посредством Nme2Cas9, которое является в высокой степени точным в клетках человека. Показаны количества участков вне мишени, обнаруженных посредством GUIDE-Seq для каждой нуклеазы в индивидуальных участках-мишенях. В дополнение к двойным участкам, авторы изобретения проанализировали TS6 (вследствие его высокой эффективности редактирования в мишени) и участки *Pcsk9* и *Rosa26* в клетках Нера1-6 мыши (для измерения точности в другом типе клеток). На фиг.14D показано, что иллюстративное направленное глубокое секвенирование для детекции инсерций-делеций в отредактированных клетках подтверждает высокую точность Nme2Cas9, показанную посредством GUIDE-seq. На фиг.14E представлена иллюстративная последовательность подтвержденного участка вне мишени для гена *Rosa26*, демонстрирующая область PAM (подчеркнута), консенсусный динуклеотид CC PAM (полужирным шрифтом) и три несоответствия оснований в дистальной от PAM части спейсера (красный).

Для валидации участков вне мишени, выявленных посредством GUIDE-seq, проводили направленное глубокое секвенирование для определения образования инсерций-делеций для основных локусов вне мишени после независимого от GUIDE-seq редактирования (т.е. без совместной трансфекции dsODN). В то время как SpyCas9 продемонстрировала выраженное редактирование в большинстве протестированных участков вне мишени и в некоторых случаях была более эффективной, чем в соответствующем участке-мишени, Nme2Cas9 продемонстрировала отсутствие поддающихся детекции инсерций-делеций в единственных участках DS2 и DS6 вне мишени, являющихся кандидатами. См., фиг.14D. В случае sg-PHK *Rosa26*, Nme2Cas9

индуцировала редактирование ~1% в участке *Rosa26-OT1* в клетках Нера1-6 по сравнению с ~30% редактированием в мишени. См., фиг.14D. Следует отметить, что этот участок вне мишени имеет консенсусный PAM Nme2Cas9 (ACTCCCT) только с 3 несоответствиями на дистальном относительно PAM конце комплементарной гиды области (т.е. вне затравки). См., фиг.14E. Эти данные подтверждают и подкрепляют результаты GUIDE-seq, полученные авторами изобретения, указывающие на высокую степень точности для редактирования генома Nme2Cas9 в клетках млекопитающих.

Для дальнейшего подтверждения полученных выше результатов GUIDE-Seq использовали CRISPRseek для компьютерного прогнозирования потенциальных участков вне мишени для двух активных sg-PHK Nme2Cas9, которые были нацелены на TS25 и TS47, оба из которых также находятся в *VEGFA*. См., фиг.9A; (Zhu et al., 2014). Каждый из трех (TS25) или четырех (TS47) из наиболее близко совпавших спрогнозированных участков, пять с PAM N<sub>4</sub>CC и два с PAM N<sub>4</sub>CA; имел 2-5 несоответствия, в основном в их дистальных относительно PAM участках, не находящихся в затравках. См., фиг.13E. Редактирование в мишени против редактирования вне мишени сравнивали после трансфекций плазмид Nme2Cas9+sg-PHK в клетки HEK293T посредством направленной амплификации каждого локуса с последующим анализом TIDE. Равным образом, нельзя было обнаружить инсерций-делеций в этих участках вне мишени для любой из sg-PHK посредством TIDE, в то время как эффективное редактирование в мишени легко обнаруживалось в ДНК из тех же популяций клеток. Взятые вместе, данные, полученные авторами изобретения, указывают на то, что Nme2Cas9 является естественным образом сверхточной платформой редактирования генома в клетках млекопитающих.

### 7. Доставка посредством аденоассоциированного вируса

Компактный размер, малый PAM и высокая точность Nme2Cas9 обеспечивают значительные преимущества редактирования генома *in vivo* с использованием доставки посредством аденоассоциированного вируса (AAV). Для тестирования того, может ли быть достигнуто эффективное редактирование генома Nme2Cas9 посредством доставки с использованием одного AAV, Nme2Cas9 клонировали с его sg-PHK и их промоторами (U1a и U6, соответственно) в остов вектора AAV. См., фиг.15A. AAV "все в одном" создавали с конструкцией sgPHK-Nme2Cas9, упакованной в гепатотропный капсид AAV8, для нацеливания на два гена в печени мыши: i) *Rosa26* (часто используемый локус "безопасной гавани" для встраивания трансгена) (Friedrich and Soriano, 1991) в качестве отрицательного контроля; и ii) *Pcsk9*, основной регулятор гомеостаза циркулирующего холестерина (Rashid et al., 2005), в качестве фенотипической мишени.

Индукцированные SauCas9 или Nme1Cas9 инсерции-делеции в *Pcsk9* в печени мыши приводят к снижению уровней холестерина, обеспечивая полезный и легко оцениваемый образец *in vivo* для новых платформ редактирования (Ran et al., 2015; Ibraheim et al., 2018). PHK-гиды Nme2Cas9 были такими же как те, которые использовались выше. См., фиг.9B, фиг.13D и фиг.14. Поскольку *Rosa26-OT1* был

единственным участком вне мишени для Nme2Cas9, который был валидирован в культивированных клетках млекопитающих, гид *Rosa26* также обеспечил автора изобретения возможность оценить редактирование в мишени против редактирования вне мишени *in vivo*. см., фиг.14D-E). В хвостовые вены двух групп мышей ( $n=5$ ) инъецировали  $4 \times 10^{11}$  копий генома AAV8.sgRNA.Nme2Cas9 (GC), нацеленных либо на *Pcsk9*, либо на *Rosa26*. Взятие сыворотки проводили через 0, 14 и 28 суток после инъекции для определения уровня холестерина. Мышей умерщвляли через 28 суток после инъекции и проводили сбор тканей печени. См., фиг.15A. Направленное глубокое секвенирование каждого локуса выявило индукцию инсерций-делеций ~38% и ~46% в локусах редактирования *Pcsk9* и *Rosa26*, соответственно, в печени. См., фиг.15B. Поскольку гепатоциты составляют только 65-70% от общего состава клеток во взрослой печени, эффективность индуцируемого AAV Nme2Cas9 редактирования в гепатоцитах с *sgPcsk9* и *sgRosa* составляла приблизительно 54-58% и 66-71%, соответственно (Racanelli and Rehmann, 2006).

Только 2,25% от всех инсерций-делеций в печени (~3-3,5% в гепатоцитах) было выявлено в участке вне мишени *Rosa26*-OT1, по сравнению с редактированием 1%, которое наблюдалось в этом участке в трансфицированных клетках Hepa1-6. Фиг.15B против фиг.14D. Как через 14, так и через 28 суток после инъекции редактирование *Pcsk9* сопровождалось снижением на ~44% уровней сывороточного холестерина, в то время как мыши, которым вводили AAV, экспрессирующий *sgRosa26*, сохраняли нормальный уровень холестерина на протяжении исследования. См., фиг.15C. Снижение на ~44% сывороточного холестерина у мышей, которым вводили AAV Nme2Cas9/*sgPcsk9* хорошо сравнимо со снижением на ~40%, описанным для AAV "все в одном" с SauCas9 при нацеливании на тот же ген (Ran et al., 2015).

На фиг.15 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 *in vivo* посредством доставки с использованием AAV "все в одном". На фиг.15A представлена иллюстративная схема процесса для доставки AAV8.sgRNA.Nme2Cas9 для снижения уровней холестерина у мышей посредством нацеливания на *Pcsk9*. Сверху: схема вектора AAV "все в одном", экспрессирующего Nme2Cas9 и sg-РНК (индивидуальные геномные элементы не в масштабе). ВGH, участок поли(А) бычьего гормона роста; HA, эпитопная метка; NLS, последовательность ядерной локализации; h, кодон-оптимизированный для человека. Снизу: расписание инъекций в хвостовую вену AAV8.sgRNA.Nme2Cas9 ( $4 \times 10^{11}$  GC) с последующим измерением уровня холестерина на 14 сутки и анализом инсерций-делеций, гистологии и холестерина на 28 сутки после инъекции. На фиг.15B представлен иллюстративный анализ TIDE для определения инсерций-делеций в ДНК, экстрагированной из печени мышей, которым инъецировали AAV8.Nme2Cas9+sg-РНК, нацеленную на локусы *Pcsk9* и *Rosa26* (контроль). Также посредством TIDE оценивали эффективность инсерций-делеций в единственном участке вне мишени, идентифицированном посредством GUIDE-seq для

двух sg-РНК (*Rosa26|OT1*). На фиг. 15С представлены иллюстративное снижение сывороточных уровней холестерина у мышей, которым инъецировали нацеленный на *Pcsk9* гид, по сравнению с контролями, нацеленными на *Rosa26*. Значения *P* вычислены с использованием непарного двухстороннего *t*-критерия.

Проводили вестерн-блоттинг с использованием антитела против PCSK9 для оценки уровней белка PCSK9 в печени мышей, которым вводили *sgPcsk9* и *sgRosa26*. Уровень PCSK9 печени были ниже предела обнаружения у мышей, которым вводили *sgPcsk9*, в то время как мыши, которым вводили *sgRosa26*, имели нормальные уровни PCSK9. См., фиг. 16А. Окрашивание гематоксилином и эозином (H&E) и гистология не выявили признаков токсичности или повреждения тканей в любой из групп после экспрессии *Nme2Cas9*. См., фиг. 16В. Эти данные подтверждают *Nme2Cas9* в качестве высокоэффективной системы редактирования генома *in vivo*, в том числе при доставке посредством единичных векторов AAV.

Недавно векторы на основе AAV использовали для получения мышей с отредактированным геномом без необходимости в микроинъекции или электропорации, просто посредством погружения зигот в культуральную среду, содержащую вектор(ы) AAV, с последующей реимплантацией псевдобеременным самкам (Yoon et al., 2018). Редактирование было осуществлено ранее с использованием системы с двойным AAV, в которой *SpyCas9* и ее sg-РНК доставляли в отдельных векторах (Yoon et al., 2018). Для тестирования того, может ли *Nme2Cas9* проводить точное и эффективное редактирование в зиготах мышей с использованием систем доставки AAV "все в одном" авторы осуществили нацеливание на тирозиназу (*Tyr*). Биаллельная инактивация *Tyr* нарушает продуцирование меланина, что приводит к фенотипу альбиносов (Yokoyama et al., 1990).

Была подтверждена эффективная sg-РНК к *Tyr*, которая расщепляет локус *Tyr* только на расстоянии семнадцати (17) п.н. от участка классической мутации *albino* в клетках Нера1-6 посредством временных трансфекций. См., фиг. 17А. Далее зиготы C57BL/6NJ инкубировали в течение 5-6 часов в культуральной среде, содержащей  $3 \times 10^9$  или  $3 \times 10^8$  GC вектора AAV6 "все в одном", экспрессирующего *Nme2Cas9* вместе с sg-РНК *Tyr*. После культивирования в течение ночи в свежей среде зиготы, которые достигли стадии двух клеток, переносили в яйцевод псевдобеременных реципиентов, и им позволяли развиваться до полного срока. См., фиг. 18А. Анализ цвета шерсти детенышей выявил мышей, которые были альбиносами, шиншиловыми (что указывает на гипоморфный аллель тирозиназы) или которые имели пятнистую шерсть, состоящую из альбиносных и шиншилловых пятен, но лишенную черной пигментации. См., фиг. 18В-С. Эти результаты указывают на высокую частоту биаллельных мутаций, поскольку присутствие аллеля тирозиназы дикого типа должно обеспечить черную пигментацию. Всего пять детенышей (10%) было рождено в эксперименте с  $3 \times 10^9$  GC. Все из них имели инсерции-делции, фенотипически двое были альбиносами, один был шиншилловым и два имели пятнистую пигментацию, указывающую на мозаицизм.

В эксперименте с  $3 \times 10^8$  GC было получено четыре (4) детеныша (14%), оба из которых умерли при рождении, что препятствовало анализу цвета шерсти или геномному анализу. Анализ цвета шерсти оставшихся двух детенышей продемонстрировал одного шиншиллового и одного мозаичного детеныша. Эти результаты указывают на то, что доставку посредством единого AAV Nme2Cas9 и ее гида можно использовать для получения мутаций в зиготах мышей без микроинъекции или электропорации.

Для измерения образования инсерции-делеции в мишени в гене *Tyr*, ДНК выделяли из хвоста каждой мыши, локус амплифицировали и после этого проводили анализ TIDE. Все мыши имели высокие уровни редактирования в мишени посредством Nme2Cas9, варьирующиеся от 84% до 100%. См., фиг.17B-C. Большинство очагов повреждения у мышей-альбиносов 9-1 представляли собой делецию либо 1, либо 4 п.н., предполагая либо мозаицизм, либо трансгетерозиготность, однако мышь-альбинос 9-2 имела единообразную делецию 2 п.н. См., фиг.17C. На фиг.17 представлены иллюстративные данные, демонстрирующие редактирование *Tyr ex vivo* в зиготах мышей, связанные с фиг.16. На фиг.17A представлены иллюстративные два участка в *Tyr*, каждый с PAM N<sub>4</sub>CC, которые тестировали в отношении редактирования в клетках Нера1-6. Гид *sgTyr2* демонстрировал более высокую эффективность редактирования и был выбран для дальнейшего тестирования. На фиг.17B представлены иллюстративные семь мышей, которые выжили в ходе постнатального развития и которые имели фенотипы окраса шерсти, а также редактирование в мишени при анализе посредством TIDE. На фиг.17C представлены иллюстративные спектры инсерций-делеций из ДНК хвоста каждой мыши из (B), а также мыши C57BL/6NJ без редактирования, как показано посредством анализа TIDE. Указана эффективность инсерций (положительные) и делеций (отрицательные) различных размеров.

На фиг.18 представлены иллюстративные данные, демонстрирующие редактирование генома Nme2Cas9 *ex vivo* посредством доставки с использованием AAV "все в одном". На фиг.18A представлена иллюстративная схема процесса для редактирования Nme2Cas9 посредством одного AAV *ex vivo* для получения мышей-альбиносов C57BL/6NJ посредством нацеливания на ген *Tyr*. Зиготы культивировали в KSOM, содержащей AAV6.Nme2Cas9:sg*Tyr*, в течение 5-6 часов, ополаскивали M2 и культивировали в течение суток перед переносом в яйцевод псевдобеременных реципиентов. На фиг.18B представлены иллюстративные мыши-альбиносы (слева) и шиншилловые или пятнистые (середина) мыши, полученные посредством  $3 \times 10^9$  GC, и шиншилловые или пятнистые мыши (справа), полученные посредством  $3 \times 10^8$  GC, для зигот с AAV6.Nme2Cas9:sg*Tyr*. На фиг.18C представлено иллюстративное обобщение экспериментов по редактированию *Tyr* посредством одного AAV Nme2Cas9:sg*Tyr ex vivo* в двух дозах AAV.

Данные о том, что у мыши 9-2 не было мозаицизма или что у мыши 9-1 отсутствовали дополнительные аллели являются неокончательными, поскольку

секвенировали только образцы из хвоста, и другие ткани могли иметь отличающиеся участки повреждения. Анализ ДНК хоста шиншилловых мышей продемонстрировал присутствие мутаций в рамке считывания, которые потенциально вызывают шиншиловый окрас шерсти. Ограниченная мутационная комплексность указывает на то, что редактирование у этих мышей произошло на раннем этапе эмбрионального развития. Эти результаты обеспечивают упрощенный путь мутагенеза у млекопитающих с использованием одного вектора AAV, в этом случае доставляющего как Nme2Cas9, так и ее sg-РНК.

На фиг.19 представлен иллюстративный репортерный анализ mCherry для nSpCas9-ABEmax и ABEmax-nNme2Cas9 с оптимизированной активностью (D16A). На фиг.19А представлена иллюстративная информация о последовательности репортера ABE-mCherry. В кодирующей области mCherry присутствует стоп-кодон TAG. В стабильной клеточной линии с встроенным репортером отсутствует сигнал mCherry. Сигнал mCherry обнаруживается, если nSpCas9-ABEmax или оптимизированная конструкция ABEmax-nNme2Cas9 (D16A) может конвертировать TAG в CAG (который кодирует Gln). На фиг.19В представлено зажигание иллюстративных сигналов mCherry, поскольку SpCas9-ABE или ABEmax-nNme2Cas9 (D16A) является активным в конкретной области репортера mCherry. Верхняя панель соответствует отрицательному контролю, средняя панель соответствует сигналам mCherry, загоревшимся в репортерных клетках обработанных nSpCas9-ABEmax, на нижней панели представлены сигналы mCherry, загоревшиеся в репортерных клетках, обработанных оптимизированной конструкцией ABEmax-nNme2Cas9 (D16A). На фиг.19С представлено иллюстративное количественное определение посредством FACS событий редактирования оснований в репортерных клетках mCherry, трансфицированных SpCas9-ABE или ABEmax-nNme2Cas9 (D16A). N = 6; планки погрешности соответствуют S.D. Результаты получены для биологических повторений, выполненных в виде технических дублей.

На фиг.20 представлен иллюстративный репортерный анализ с GFP для активности nSpCas9-CBE4 (Addgene #100802) и CBE4-nNme2Cas9 (D16A)-UGI-UGI (CBE4 был клонирован из Addgene #100802). На фиг.20А представлена иллюстративная информация о последовательности для репортера CBE-GFP. В центральной области флуорофора линии с репортером GFP существует мутация, которая конвертирует GYG в GHG. Таким образом, отсутствует сигнал GFP. Сигнал GFP появляется, если nSpCas9-CBE4 или CBE4-nNme2Cas9 (D16A)-UGI-UGI может конвертировать CAC в TAC/TAT (гистидин в тирозин). На фиг.20В представлен иллюстративный сигнал GFP (зеленый), поскольку nSpCas9-CBE4 или CBE4-nNme2Cas9(D16A)-UGI-UGI является активной в конкретной области репортера GFP. Верхняя панель соответствует отрицательному контролю. На средней панели показано, что сигналы mCherry загораются в репортерных клетках, обработанных CBE4-nNme2Cas9 (D16A)-UGI-UGI. На нижней панели показано, что сигналы GFP загораются в репортерных клетках, обработанных CBE4-nNme2Cas9

(D16A)-UGI-UGI). На фиг.20С представлено иллюстративное количественное определение посредством FACS событий редактирования оснований в репортерных клетках с GFP, трансфицированных nSpCas9-CBE4 или CBE4-nNme2Cas9(D16A)-UGI-UGI. N = 6; планки погрешности соответствуют S.D. Результаты получены для биологических повторений, выполненных в виде технических дублей.

На фиг.21 представлено иллюстративное редактирование цитозина посредством CBE4-nNme2Cas9(D16A)-UGI-UGI. На верхней панели представлена информация о последовательности-мишени KANK3 (последовательности PAM подчеркнуты красным) для Nme2Cas9 и редактирование оснований в образцах отрицательного контроля. На нижней панели представлено количественное определение частоты замены каждого типа оснований в окне редактирования CBE4-nNme2Cas9(D16A)-UGI-UGI для последовательностей-мишеней KANK3. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования С в Т выделены красным цветом.

На фиг.22 представлено иллюстративное редактирование цитозина и аденина посредством CBE4-nNme2Cas9 (D16A)-UGI-UGI и оптимизированного ABEmax-nNme2Cas9 (D16A), соответственно. На верхней панели представлена информация о последовательности-мишени PLXNB2 (последовательности PAM подчеркнуты красным) для Nme2Cas9 и редактирование оснований в образцах отрицательного контроля. На средней панели представлено количественное определение частоты замены каждого типа оснований в окнах редактирования оптимизированного ABEmax-nNme2Cas9 (D16A) для последовательностей-мишеней PLXNB2. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования А в G выделены красным цветом. На нижней панели представлено количественное определение частоты замены каждого типа оснований в окне редактирования CBE4-nNme2Cas9 (D16A)-UGI-UGI для последовательностей-мишеней PLXNB2. В таблицах последовательностей представлены частоты нуклеотидов в каждом положении. Частоты ожидаемого конвертирования С в Т выделены красным цветом.

## 8. Последовательности

### Выравнивание Nme1Cas9 и Nme2Cas9

Отличия а.к. не в PID (сине-зеленый - подчеркнуты); отличия а.к. в PID (желтый - подчеркнуты, полужирным шрифтом); остатки активного центра (красный - полужирным шрифтом).

Nme1Cas9 (1-60)

MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAM

Nme2Cas9 (1-60)

MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAM

Nme1Cas9 (61-120)

ARRLARSVRRLTRRRAHRLRLRTRRLKREGVLQAANFDENGLIKSLPNTPWQLRAAAL

DR

Nme2Cas9 (61-120)

ARRLARSVRRRLTRRRRAHRLLRARRLKREGVLQAADFDENGLIKSLPNTPWQLRAAAL

DR

Nme1Cas9 (121-180)

KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVAGNAHALQTGDFRTPA

EL

Nme2Cas9 (121-180)

KLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPA

EL

Nme1Cas9 (181-240)

ALNKFEKESGHIRNQRSDYSHTF SRKDLQAE LILLFEKQKEFGNPHVSGGLKEGIETLLM

Nme2Cas9 (181-240)

ALNKFEKESGHIRNQRGDYSHTF SRKDLQAE LILLFEKQKEFGNPHVSGGLKEGIETLLM

Nme1Cas9 (241-300)

TQRPALSGDAVQKMLGHCTFEPAPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTD

T

Nme2Cas9 (241-300)

TQRPALSGDAVQKMLGHCTFEPAPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTD

T

Nme1Cas9 (301-360)

ERATLMDEPYRKS<sup>K</sup>SLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISR

AL

Nme2Cas9 (301-360)

ERATLMDEPYRKS<sup>K</sup>SLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISR

AL

Nme1Cas9 (361-420)

EKEGLKDKKSPNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKF

Nme2Cas9 (361-420)

EKEGLKDKKSPNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKF

Nme1Cas9 (421-480)

VQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRA

Nme2Cas9 (421-480)

VQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRA

Nme1Cas9 (481-540)

LSQARKVINGVRRYGGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFRE

Y

Nme2Cas9 (481-540)

LSQARKVINGVRRYGGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFRE

Y

Nme1Cas9 (541-600)

FPNFVGE**P**KSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDH**A**LPFSRTWDDSF

Nme2Cas9 (541-600)

FPNFVGE**P**KSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDH**A**LPFSRTWDDSF

Nme1Cas9 (601-660)

NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD  
ED

Nme2Cas9 (601-660)

NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD  
ED

Nme1Cas9 (661-720)

GFKERNLNDRYVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGLRKVRA  
END

Nme2Cas9 (661-720)

GFKECNLNDRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAE  
ND

Nme1Cas9 (721-780)

RHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWE  
FFA

Nme2Cas9 (721-780)

RHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWE  
FFA

Nme1Cas9 (781-840)

QEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSRPEAVHEYVTPLFVSRAPNRKMS  
G

Nme2Cas9 (781-840)

QEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSRPEAVHEYVTPLFVSRAPNRKMS  
G

Nme1Cas9 (841-895)

**QGHMETVKSAK---RLDEGVSVLRVPLTQLKLKDLEKMVNR--  
ERE**P**KLYEALKARLEAH**

Nme2Cas9 (841-899)

**AHK-  
DTLRS**A**K**R****F****V****K****H****N****E****K****I****S****V****K****R****V****W****L****T****E****I****K****L****A****D****L****E****N****M****V****N****Y****K****N****G****R****E****I****E****L****Y****E****A****L****K****A****R****L****E****A****Y****

Nme1Cas9 (896-950)

**KDDPA**K****A****F****A****E**---PFY**K****Y****D****K****A****G****N****R****T****Q****Q****V****K****A****V****R****V****E****Q****V****K****T****G****V****W****V****R****N****H**--  
**NGIAD**N****A****T****M****V****R****V******

Nme2Cas9 (900 -954)

GGNAKQAFDPKDNPFYKK---G--

GQLVKA VRVEKTQESGVLLNKKNAYTIADNGDMVRV

Nme1Cas9 (951-1005)

DVFEKG-----

DKYYLVPIYSWQVAKGI LPDRAVVQ GKDEEDWQLIDDSFNFKFSLHPND

Nme2Cas9 (955-1007)

DVFCVKVDKKGKNQYFIVPIYA WQVAENILPDIDCKG-----

YRIDDSYTF CFSLHKYD

Nme1Cas9 (1006-1063)

LVEVIT--

KKARMF GYFASCHRGTGNINIRI HDLDHKIGKNGILEGIGVKTALS FQKYQI

Nme2Cas9 (1008-1063)

LIAFQKDEKSKVEFA YYINC DSSNGRFYLA WHDKGSKEQ----

QFRISTQNLVLIQKYQV

Nme1Cas9 (1064-1082)

DELGKEIRPCRLKKRPPVR

Nme2Cas9 (1064-1082)

NELGKEIRPCRLKKRPPVR

*Выравнивание Nme1Cas9 и Nme3Cas9*

Отличия а.к. не в PID (сине-зеленый - подчеркнуты); отличия а.к. в PID (желтый - подчеркнуты, полужирным шрифтом); остатки активного центра (красный - полужирным шрифтом).

Nme1Cas9 1

MAAFKPNSINYLGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAE 50

Nme3Cas9 1

MAAFKPNPINYLGLDIGIASVGWAMVEIDEEENPIRLIDLGVRFERAE 50

Nme1Cas9 51

VPKTGDSLAMARRLARSVRRLTRRRAHRLLRTRRLKREGVLQAANFDEN 100

Nme3Cas9 51

VPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLKREGVLQAADFDEN 100

Nme1Cas9 101

GLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGET 150

Nme3Cas9 101

GLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGET 150

Nme1Cas9 151

ADKELGALLKGVAGNAHALQTGDFRTPAELALNKFEKESGHIRNQRSDYS 200

Nme3Cas9 151

ADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKECGHIRNQRGDYS 200

Nme1Cas9 201

HTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDA 250  
 Nme3Cas9 201

HTFSRKDLQAENLLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDA 250  
 Nme1Cas9 251

VQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSRPLTDT 300  
 Nme3Cas9 251

VQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSRPLTDT 300  
 Nme1Cas9 301

ERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEM 350  
 Nme3Cas9 301

ERATLMDEPYRKSKLTYAQARKLLSLEDTAFFKGLRYGKDNAEASTLMEM 350  
 Nme1Cas9 351

KAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLK 400  
 Nme3Cas9 351

KAYHTISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLK 400  
 Nme1Cas9 401

DRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYG 450  
 Nme3Cas9 401

DRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYG 450  
 Nme1Cas9 451

DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPAR 500  
 Nme3Cas9 451

DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPAR 500  
 Nme1Cas9 501

IHIETAREVGKSFKDRKEIEKRQEENRDKREKAAAKFREYFPNFVGEPKS 550  
 Nme3Cas9 501

IHIETAREVGKSFKDRKEIEKRQEENRDKREKAAAKFREYFPNFVGEPKS 550  
 Nme1Cas9 551

KDILKLRLYEQQHKGCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSF 600  
 Nme3Cas9 551

KDILKLRLYEQQHKGCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSF 600  
 Nme1Cas9 601

NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQ 650  
 Nme3Cas9 601

NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQ 650  
 Nme1Cas9 651

RILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNG 700  
 Nme3Cas9 651

RILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKKRVFASNG 700

Nme1Cas9 701  
 QITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEM 750  
 Nme3Cas9 701  
 QITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEM 750  
 Nme1Cas9 751  
 NAFDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEA 800  
 Nme3Cas9 751  
 NAFDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEA 800  
 Nme1Cas9 801  
 DTLEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSA 850  
 Nme3Cas9 801  
 DTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSA 850  
 Nme1Cas9 851  
 KRLDEGVSVLRVPLTQLKLDLEKMNRREREPKLYEALKARLEAHKDDPA 900  
 Nme3Cas9 851  
 KRLDEGVSVLRVPLTQLKLDLEKMNRREREPKLYEALKARLEAHKDDPA 900  
 Nme1Cas9  
 901  
 KAFAEPFYKYDKAGNRTQQVKA VRVEQVQKTGVWVRNHNGIADNATMVRV 950  
 Nme3Cas9  
 901  
 KAFAEPFYKYDKAGNRTQQVKA VRVEQVQKTGVWVRNHNGIADNATMVRV 950  
 Nme1Cas9 951  
 DVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFS 1000  
 Nme3Cas9 951  
 DVFEKGDKYYLVPIYSWQVAKGILPDRAVVAYADEEDWTVIDESFRFKFV 1000  
 Nme1Cas9 1001  
 LHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGI 1050  
 Nme3Cas9 1001  
 LYSNDLIKVQLKKDSFLGYFSGLDRATGAISLREHDLEKSKGKDG-MHRI 1049  
 Nme1Cas9 1051 GVKTALSFKYQIDELGKEIRPCRLKKRPPVR 1082  
 Nme3Cas9 1050 GVKTALSFKYQIDEMGKEIRPCRLKKRPPVR 1081  
Экспрессируемая плазмидой Nme2Cas9  
 NLS SV40 (желтый - полужирным шрифтом); 3X-НА-метка (зеленый -  
 (подчеркиванием/полужирным шрифтом); сМус-подобный NLS (сине-зеленый -  
 простым текстом); линкер (фуксиновый - *полужирным шрифтом, курсивом*) и  
 Nme2Cas9 (*курсивом*).  
 MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSL  
 MARRLARSVRRLLTRRRRAHRLLRARLLKREGVLQAADFDEGLIKSLPNTPWQLRAAALDRKL

TPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNK  
 FEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALS  
 GDAVQKMLGHCTFEPAPKAAKNTTYAERFIWLTCLNLRILEQGSERPLDTERATLMDEP  
 YRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPL  
 NLSSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLALRRIVPLMEQ  
 GKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGVRRYGSAPARI  
 HIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPEKSKDILKLRLYEQQH  
 GKCLYSGKEINLVRLENEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNG  
 KDNSREWQEFKARVETSRFPRSKKQRILLQKFDDEDGFKECNLNDTRYVNRFLCQFVADHILL  
 TGKGRRVFASNGQITNLLRGFWGLRQVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKE  
 MNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRLL  
 AEKLSRPEAVHEYVTPFVSRAPNRKMSGAHKDTLRSKRFRVKNHNEKISVKRVWLTEIKLAD  
 LENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGQLVKAVRVEKTQESG  
 VLLNKKNAYTIADNGDMVRVDVFCCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSY  
 TFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQK  
 YQVNELGKEIRPCRLKKRPPVRGTGGPKKKRKVYPYDVPDYAGYPYDVPDYAGSYPYD  
VPDYAGSAAPAAKKKKLDFESG\*

*Nme2Cas9, экспрессируемая с AAV*

NLS SV40 (желтый - полужирным шрифтом); 3X-НА-метка (зеленый - **подчеркиванием/полужирным шрифтом**); нуклеоплазмин-подобный NLS (красный - **подчеркиванием**); сМус NLS (сине-зеленый - простым текстом); линкер (фуксиновый - ***полужирным шрифтом, курсивом***) и *Nme2Cas9* (***курсивом***).

***MVPKKKRKVEDKRPAAATKKAGQAKKKKMAAFKPNPINYILGLDIGIASVGMAMV***  
 EIDEEENPIRLIDLGVRFERAEVPKTGDSLAMARRLARSVRLTRRRARHLLRARLLKREGV  
 LQAADFENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADK  
 ELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAEILLF  
 EKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAPKAAKNTTYAERFI  
 WLTCLNLRILEQGSERPLDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDN  
 AEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSSELQDEIGTAFSLFKTDEDITGRLKDRVQ  
 PEILEALLKHISFDKQVQISLALRRIVPLMEQGGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPA  
 DEIRNPVLRALSQARKVINGVRRYGSAPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKA  
 AAKFREYFPNFVGEPEKSKDILKLRLYEQQH GKCLYSGKEINLVRLENEKGYVEIDHALPFSRTW  
 DDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD  
 EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGRRVFASNGQITNLLRGFWGLRQVRAEND  
 RHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQ  
 EVMIRVFGKPDGKPEFEEADTPEKLRLLAEKLSRPEAVHEYVTPFVSRAPNRKMSGAHK  
 DTLRSKRFRVKNHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAF  
 FDPKDNPFYKKGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCCKVDKKGKN  
 QYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSN

GRFYLA WHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVREDKRPAAATKK  
AGQAKKKKYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAААРААККККLD\*

Рекомбинантная Nme2Cas9

NLS SV40 (желтый - полужирным шрифтом); нуклеоплазмин-подобный NLS (красный-подчеркиванием); линкер (фуксиновый - *полужирным шрифтом, курсивом*) и Nme2Cas9 (*курсивом*).

**PKKKRKVN**АМАААFKPNPINYILGLDIGIASV**GWAMVEIDEEENPIRLIDL**GVRVFERA  
EVPKTGDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFDENGLIKSLPNTPWQ  
LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDF  
RTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIET  
LLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNYTAERFIWLTCLNNLRILEQGSERPLTDT  
ERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEK  
EGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKA  
LRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGV  
VRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDIL  
KLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGN  
QTPYEYFNGKDNSREWQEFKARVETSFRPFRSKKQRIILLQKFEDEDFKECNLNDTRYVNRFLC  
QFVADHILLTGKGRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQ  
KITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEA  
DTPEKLRLLAEKLSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSKRFRVKHNEKISVCR  
VWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKGGQVLKAV  
RVEKTQESGVLLNKNAYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPIYAWQVAENILPDID  
CKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLA WHDKGSKEQQFRIS  
TQNLVLIQKYQVNELGKEIRPCRLKKRPPV**RGGGGSGGGGSGGGGSPA**AKKKKLDGGG  
**SKRPAATKKAGQAKKKK\***

Рекомбинантная Nme2Cas9 для применения для доставки RNP в клетки  
млекопитающих:

NLS SV40 (желтый - полужирным шрифтом); нуклеоплазмин-подобный NLS (красный-подчеркиванием); линкер (фуксиновый - *полужирным шрифтом, курсивом*) и Nme2Cas9 (*курсивом*).

**PKKKRKVN**АМАААFKPNPINYILGLDIGIASV**GWAMVEIDEEENPIRLIDL**GVRVFERA  
EVPKTGDSLAMARRLARSVRRLTRRRRAHRLLRARLLKREGVLQAADFDENGLIKSLPNTPWQ  
LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDF  
RTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIET  
LLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNYTAERFIWLTCLNNLRILEQGSERPLTDT  
ERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEK  
EGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKA  
LRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVLRALSQARKVINGV  
VRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDIL

KLRLYEQQHKGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENGQNKGN  
 QTPYEYFNGKDNSREWQEFKARVETSFRFPRSKKQRILLQKFEDEDGFKECNLNDTRYVNRFLC  
 QFVADHILLTGKGGKRRVFAVNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQ  
 KITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEA  
 DTPEKLRITLLAEKLSRPEAVHEVYVTPLFVSRAPNRKMSGAHKDTLRSKRFRVVKHNEKISVKR  
 VWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGLVKAV  
 RVEKTQESGVLLNKKNAVYTIADNGDMVRVDVFCVKVDKKGKNQYFIVPIYAWQVAENILPDD  
 CKGYRIDDSYTFCSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRI  
 STQNLVLIQKYQVNELGKEIRPCRLKKRPPVRRGGGGSGGGGSGGGGSPAAKKKKLDGGG  
 SKRPAATKKAGQAKKKK\*

### 9. Терапевтические применения

Хотя компактные ортологи Cas9 ранее были валидированы в отношении редактирования генома, в том числе доставки посредством одного AAV, их более длинные PAM имеют ограниченную возможность терапевтической разработки вследствие частот участков-мишеней, которые являются более низким, чем у более широко признанных Spycas9. Кроме того, Saucas9 и ее вариант KKH со сниженными требованиями к PAM (Kleinstiver et al., 2015) склонны к редактированию вне мишени с некоторыми из sg-PHK (Friedland et al., 2015; Kleinstiver et al., 2015). Эти ограничения усугубляются локусами-мишенями, которые требуют редактирования в пределах узкого окна последовательности или которые требуют точной сегментной делеции. Авторы изобретения идентифицировали Nme2Cas9 в качестве компактной и высокоточной Cas9 с менее ограничивающим динуклеотидным PAM для редактирования генома посредством доставки AAV *in vivo*. Разработка Nme2Cas9 значительно расширяет геномный охват редактирования *in vivo*, особенно посредством вирусной доставки вектора. Платформа для доставки Nme2Cas9 на основе AAV "все в одном", полученная в этом исследовании, в принципе может использоваться для нацеливания на настолько же широкий диапазон участков, что и у Spycas9 (вследствие идентичной плотности оптимальных PAM N<sub>4</sub>CC и NGG), но без необходимости в доставке двух отдельных векторов в одни и те же клетки-мишени. Доступность каталитически мертвой версии Nme2Cas9 (dNme2Cas9) также является перспективной для расширения объема применений, таких как CRISPRi, CRISPRa, редактирование оснований и сходные подходы (Dominguez et al., 2016; Komor et al., 2017). Более того, сверхточность Nme2Cas9 позволяет точное редактирование генов-мишеней, потенциально улучшая проблемы безопасности в результате активности вне мишени. Возможно, вопреки очевидному, более высокая плотность участков-мишеней Nme2Cas9 (по сравнению с Nme1Cas9) не приводит к относительному повышению редактирования вне мишени у первой из них. Сходные результаты были описаны недавно для вариантов Spycas9, которые эволюционировали для наличия более коротких PAM (Hu et al., 2018). Ортологи Cas9 типа II-C, как правило, являются более медленными нуклеазами *in vitro*, чем Spycas9 (Ma et al., 2015; Mir et al., 2018); интересно, что

принципы энзимологии указывают на то, что снижение кажущейся  $k_{cat}$  (в рамках границ) может повысить специфичность в мишени против специфичности вне мишени для РНК-направляемых (Bisaria et al., 2017).

Открытие Nme2Cas9 и Nme3Cas9 было основано на неисследованных Cas9, которые являются в высокой степени родственными (вне PID) ортологу, который ранее был валидирован для редактирования генома человека (Esvelt et al., 2013; Hou et al., 2013; Lee et al., 2016; Amrani et al., 2018). Родство Nme2Cas9 и Nme3Cas9 с Nme1Cas9 добавило дополнительную пользу, а именно, то, что они могут использовать в точности тот же каркас sg-РНК, обходя необходимость в идентификации и валидации функциональных последовательностей tracr-РНК для каждой из них. В контексте естественного иммунитета CRISPR, ускоренная эволюция новой специфичности в отношении PAM может отражать давление селекции для восстановления нацеливания на фаги и MGE, которые ускользнули от вмешательства посредством мутаций PAM (Deveau et al., 2008; Paez-Espino et al., 2015). Наблюдение авторов изобретения, что AcrIIIC5<sub>Smi</sub> ингибирует Nme1Cas9, но не Nme2Cas9, указывает на вторую не взаимоисключающую основу для ускоренного варьирования PID, а именно, ускользание от ингибирования посредством белков анти-CRISPR. Также авторы изобретения полагают, что ускоренная варибельность может не ограничиваться PID, возможно в результате давления селекции для ускользания от белков анти-CRISPR, которые связывают другие домены Cas9. Ингибиторы Cas9, такие как AcrIIIC1, которые связывают более консервативные области Cas9, вероятно, имеют меньше путей для мутационного ускользания и, таким образом, демонстрируют более широкий спектр ингибирования (Harrington et al., 2017a). Независимо от источников давления селекции, движущего совместной эволюцией Acr и Cas9, доступность подтвержденных ингибиторов Nme2Cas9 (например, AcrIIIC1-4) обеспечивает возможность для дополнительных уровней контроля над ее активностью.

Подход, использованный в этом испытании (т.е. поиск быстро эволюционирующих доменов в Cas9) может быть применен в других областях, особенно для видов бактерий, которые хорошо изучены на уровне геномной последовательности. Этот подход также может быть применен для других эффекторных белков CRISPR-Cas, таких как Cas12 и Cas13, которые также внедрены в инженерию генома или транскриптома и другие применения. Эта стратегия может быть особенно основательной в случае белков Cas, которые являются близкородственными ортологами с доказанной эффективностью в гетерологичных контекстах (например, в эукариотических клетках), как было в случае Nme1Cas9. Применение этого подхода к менингококковым ортологам Cas9 обеспечило новую платформу редактирования генома, Nme2Cas9, с уникальной комбинацией характеристик (компактный размер, динуклеотидный PAM, сверхточность, возможность доставки в одном AAV и чувствительность к Acr), которая предвещает ускорение разработки инструментов для редактирования генома как для общих, так и для терапевтических применений.

**Таблица 3. Ниже представлены иллюстративные последовательности для плазмид и олигонуклеотидов, как описано в настоящем описании**

<b>Иллюстративные плазмиды</b>					
<b>Плазмида №</b>	<b>Название</b>	<b>Описание вставки</b>	<b>Остов</b>	<b>Назначение</b>	<b>Последовательность вставки</b>
1	pAE70	PID Nme3Cas9 на Nme1Cas9	pMCSG7	Бактериальная экспрессия Nme1Cas9 с PID Nme3Cas9	См. примеры настоящего описания.
2	pAE71	PID Nme2Cas9 на Nme1Cas9	pMCSG7	Бактериальная экспрессия Nme1Cas9 с PID Nme2Cas9	См. примеры настоящего описания.
3	pAE113	Nme2TLR1	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GTCACCTGCCTCGTGG AATACGG
4	pAE114	Nme2TLR2	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCACCTGCCTCGTGG ATACGGT
5	pAE115	Nme2TLR5	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	G TTCAGCGTGTCCGGC TTTGGC
6	pAE116	Nme2TLR11	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GTGGTGAGCAAGGGC GAGGAGCTG
7	pAE117	Nme2TLR12	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GGGCGAGGAGCTGTT CACCGGGGT
8	pAE118	Nme2TLR13	pLKO	Нацеливание на	GTGAACTTGTGGCCGT

				TLR2.0 посредством Nme2Cas9	TTACGTCG
9	pAE119	Nme2TLR14	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCGTCCAGCTCGACC AGGATGGGC
10	pAE120	Nme2TLR15	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCGGTGAACAGCTCC TCGCCCTTG
11	pAE121	Nme2TLR16	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GGGCACCACCCCGGT GAACAGCTC
12	pAE122	Nme2TLR17	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GGCACCACCCCGGTG AACAGCTCC
13	pAE123	Nme2TLR18	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GGGATGGGCACCACC CCGGTGAAC
14	pAE124	Nme2TLR19	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCGTGTCCGGCTTTGG CGAGACAA
15	pAE125	Nme2TLR20	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GTCCGGCTTTGGCGAG ACAAATCA
16	pAE126	Nme2TLR21	pLKO	Нацеливание на TLR2.0	GATCACCTGCCTCGTG GAATACGG

				посредством Nme2Cas9	
17	pAE149	Nme2TLR22	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GACGCTGAACTTGTG GCCGTTTAC
18	pAE150	Nme2TLR23	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCCAAAGCCGGACAC GCTGAACTT
19	pAE193	Nme2TLR13 со спейсером из 23 нуклеотидов	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GGAACTTGTGGCCGTT TACGTCG
20	pAE194	Nme2TLR13 со спейсером из 22 нуклеотидов	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GAACTTGTGGCCGTTT ACGTCG
21	pAE195	Nme2TLR13 со спейсером из 21 нуклеотида	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GACTTGTGGCCGTTTA CGTCG
22	pAE196	Nme2TLR13 со спейсером из 20 нуклеотидов	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCTTGTGGCCGTTTAC GTCG
23	pAE197	Nme2TLR13 со спейсером из 19 нуклеотидов	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GTTGTGGCCGTTTACG TCG
24	pAE213	Nme2TLR21 со спейсером G22	pLKO	Нацеливание на TLR2.0 посредством	GTCACCTGCCTCGTGG AATACGG

				Nme2Cas9	
25	pAE214	Nme2TLR21 со спейсером G21	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCACCTGCCTCGTGGA ATACGG
26	pAE215	Nme2TLR21 со спейсером G20	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GACCTGCCTCGTGGA ATACGG
27	pAE216	Nme2TLR21 со спейсером G19	pLKO	Нацеливание на TLR2.0 посредством Nme2Cas9	GCCTGCCTCGTGGAAT ACGG
28	pAE90	Nme2TS1	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GGTTCTGGGTACTTTT ATCTGTCC
29	pAE93	Nme2TS4	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GTCTGCCTAACAGGA GGTGGGGGT
30	pAE94	Nme2TS5	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GAATATCAGGAGACT AGGAAGGAG
31	pAE129	Nme2TS6	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GCCTCCCTGCAGGGCT GCTCCC
32	pAE130	Nme2TS10	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GAGCTAGTCTTCTTCC TCCAACCC
33	pAE131	Nme2TS11	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GATCTGTCCCCTCCAC CCCACAGT
34	pAE132	Nme2TS12	pLKO	Нацеливание на	GGCCCAAATGAAAGG

				<i>AAVS1</i> посредством Nme2Cas9	AGTGAGAGG
35	pAE133	Nme2TS13	pLKO	Нацеливание на <i>AAVS1</i> посредством Nme2Cas9	GCATCCTCTTGCTTTC TTTGCCTG
36	pAE136	Nme2TS16	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GGAGTCGCCAGAGGC CGGTGGTGG
37	pAE137	Nme2TS17	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GCCCAGCGGCCGGAT ATCAGCTGC
38	pAE138	Nme2TS18	pLKO	Нацеливание на <i>CYBB</i> посредством Nme2Cas9	GGAAGGGAACATATT ACTATTGC
39	pAE139	Nme2TS19	pLKO	Нацеливание на <i>CYBB</i> посредством Nme2Cas9	GTGGAGTGGCCTGCT ATCAGCTAC
40	pAE140	Nme2TS20	pLKO	Нацеливание на <i>CYBB</i> посредством Nme2Cas9	GAGGAAGGGAACATA TТАCTATTG
41	pAE141	Nme2TS21	pLKO	Нацеливание на <i>CYBB</i> посредством Nme2Cas9	GTGAATTCTCATCAGC TAAAATGC
42	pAE144	Nme2TS25	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GCTCACTCACCCACAC AGACACAC
43	pAE145	Nme2TS26	pLKO	Нацеливание на <i>CFTR</i> посредством Nme2Cas9	GGAAGAATTTTATTCT GTTCTCAG
44	pAE146	Nme2TS27	pLKO	Нацеливание на	GCTCAGTTTTCTGGA

				<i>CFTR</i> посредством Nme2Cas9	TTATGCCT
45	pAE152	Nme2TS31	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GCGTTGGAGCGGGGA GAAGGCCAG
46	pAE153	Nme2TS34	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GGGCCGCGGAGATAG CTGCAGGGC
47	pAE154	Nme2TS35	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GCCACCCGGCGGCG CCTCCCTGC
48	pAE155	Nme2TS36	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GCGTGGCAGCTGATA TCCGGCCGC
49	pAE156	Nme2TS37	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GCCGCGGCGCGACGT GGAGCCAGC
50	pAE157	Nme2TS38	pLKO	Нацеливание на <i>LINC01588</i> посредством Nme2Cas9	GTGCTCCCCAGCCCAA ACCGCCGC
51	pAE159	Nme2TS41	pLKO	Нацеливание на <i>AGA</i> посредством Nme2Cas9	GTCAGATTGGCTTGCT CGGAATTG
52	pAE185	Nme2TS44	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GCTGGGTGAATGGAG CGAGCAGCG

53	pAE186	Nme2TS45	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GTCCTGGAGTGACCCC TGGCCTTC
54	pAE187	Nme2TS46	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GATCCTGGAGTGACC CCTGGCCTT
55	pAE188	Nme2TS47	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GTGTGTCCCTCTCCCC ACCCGTCC
56	pAE189	Nme2TS48	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GTTGGAGCGGGGAGA AGGCCAGGG
57	pAE190	Nme2TS49	pLKO	Нацеливание на <i>VEGFA</i> посредством Nme2Cas9	GCGTTGGAGCGGGGA GAAGGCCAG
58	pAE191	Nme2TS50	pLKO	Нацеливание на <i>AGA</i> посредством Nme2Cas9	GTACCCTCCAATAATT TGGCTGGC
59	pAE192	Nme2TS51	pLKO	Нацеливание на <i>AGA</i> посредством Nme2Cas9	GATAATTTGGCTGGCA ATTCCGAG
60	pAE232	TS64_FancJ1	pLKO	Нацеливание на <i>FANCJ</i> посредством Nme2Cas9	GAAAATTGTGATTTC AGATCCAC
61	pAE233	TS65_FancJ2	pLKO	Нацеливание на <i>FANCJ</i> посредством Nme2Cas9	GAGCAGAAAAAATTG TGATTTC
62	pAE200	Nme2TS58	pLKO	Нацеливание на DS	GCAGGGGCCAGGTGT

		(Nme2DS1)		в <i>VEGFA</i> посредством Nme2Cas9	CCTTCTCTG
63	pAE201	Nme2TS59 (Nme2DS2)	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством Nme2Cas9	GAATGGCAGGCGGAG GTTGTA CTG
64	pAE202	Nme2TS60 (Nme2DS3)	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством Nme2Cas9	GAGTGAGAGAGTGAG AGAGAGACA
65	pAE203	Nme2TS61 (Nme2DS4)	pLKO	Нацеливание на DS в <i>VEGFA</i> with Nme2Cas9	GTGAGCAGGCACCTG TGCCAACAT
66	pAE204	Nme2TS62 (Nme2DS5)	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством Nme2Cas9	GCGTGGGGGCTCCGT GCCCCACGC
67	pAE205	Nme2TS63 (Nme2DS6)	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством Nme2Cas9	GCATGGGCAGGGGCT GGGGTGAC
68	pAE207	SpyDS1	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством SpyCas9	GGGCCAGGTGTCCTTC TCTG
69	pAE208	SpyDS2	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством SpyCas9	GGCAGGCGGAGGTTG TACTG
70	pAE209	SpyDS3	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством	GAGAGAGTGAGAGAG AGACA

				SpyCas9	
71	pAE210	SpyDS4	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством SpyCas9	GCAGGCACCTGTGCC AACAT
72	pAE211	SpyDS5	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством SpyCas9	GGGGGCTCCGTGCCC CACGC
73	pAE212	SpyDS6	pLKO	Нацеливание на DS в <i>VEGFA</i> посредством SpyCas9	GGGCAGGGGCTGGGG TGCAC
74	pAE169	hDeCas9 Wt в остоле AAV	AAV	Экспрессия Nme2Cas9 посредством AAV "все в одном" с кассетой sg-РНК	См. примеры настоящего описания.
75	pAE217	hDeCas9 wt в остоле pMSCG7	pMSCG7	Nme2Cas9 дикого типа для бактериальной экспрессии	См. примеры настоящего описания.
76	pAE107	2xNLS Nme2Cas9 с HA	pCdest	CMV-контролируемая плазмида для экспрессии Nme2Cas9	См. примеры настоящего описания.
77	pAE127	hDemonCas9 3X NLS в pMSCG7	pMSCG7	Нацеливание на эндогенные локусы посредством Nme2Cas9	См. примеры настоящего описания.
78	pAM172	hNme2Cas9 4X NLS с 3XHA	pCVL	Лентивирусный вектор, содержащий	См. примеры настоящего описания.

				UCOE, SFFV- контролируемую Nme2Cas9 и Puro	
79	pAM174	Никаза hNme2Cas9 D16A 4X NLS с 3XHA	pCVL	Лентивирусный вектор, содержащий UCOE, SFFV- контролируемую Nme2Cas9 и Puro	См. примеры настоящего описания.
80	pAM175	Никаза hNme2Cas9 H588A 4X NLS с 3XHA	pCVL	Лентивирусный вектор, содержащий UCOE, SFFV- контролируемую Nme2Cas9 и Puro	См. примеры настоящего описания.
81	pAM177	Мертвая hNme2Cas9 4X NLS с 3XHA	pCVL	Лентивирусный вектор, содержащий UCOE, SFFV- контролируемую Nme2Cas9 и Puro	См. примеры настоящего описания.

#### Иллюстративные олигонуклеотиды

Номер	Название	Последовательность	Назначение
1	AAVS1_TIDE1_FW	TGGCTTAGCACCTCTCCAT	Анализ TIDE
2	LINC01588_TIDE_FW	AGAGGAGCCTTCTGACTGCTG CAGA	Анализ TIDE
3	AAVS1_TIDE2_FW	TCCGTCTTCCTCCACTCC	Анализ TIDE
4	NTS55_TIDE_FW	TAGAGAACTGGGTAGTGTG	Анализ TIDE
5	VEGF_TIDE3_FW	GTACATGAAGCAACTCCAGT CCCA	Анализ TIDE
6	hCFTR_TIDE1_FW	TGGTGATTATGGGAGAACTG GAGC	Анализ TIDE
7	AGA_TIDE1_FW	GGCATAAGGAAATCGAAGGT C	Анализ TIDE
8	VEGF_TIDE4_FW	ACACGGGCAGCATGGGAATA GTC	Анализ TIDE

9	VEGF_TIDE5_FW	CCTGTGTGGCTTTGCTTTGGT CG	Анализ TIDE
10	VEGF_TIDE6_FW	GGAGGAAGAGTAGCTCGCCG AGG	Анализ TIDE
11	VEGF_TIDE7_FW	AGGGAGAGGGAAGTGTGGGG AAGG	Анализ TIDE
12	AAVS1_TIDE1_RV	AGAACTCAGGACCAACTTATT CTG	Анализ TIDE
13	LINC01588_TIDE_RV	ATGACAGACACAACCCAGAGG GCA	Анализ TIDE
14	AAVS1_TIDE2_RV	TAGGAAGGAGGAGGCCTAAG	Анализ TIDE
15	NTS55_TIDE_RV	CCAATATTGCATGGGATGG	Анализ TIDE
16	VEGF_TIDE3_RV	ATCAAATTCCAGCACCGAGC GC	Анализ TIDE
17	hCFTR_TIDE1_RV	ACCATTGAGGACGTTTGTCTC AC	Анализ TIDE
18	AGA_TIDE1_RV	CATGTCCTCAAGTCAAGAAC AAG	Анализ TIDE
19	VEGF_TIDE4_RV	GCTAGGGGAGAGTCCCCTG TCCA	Анализ TIDE
20	VEGF_TIDE5_RV	GTAGGGTGTGATGGGAGGCT AAGC	Анализ TIDE
21	VEGF_TIDE6_RV	AGACCGAGTGGCAGTGACAG CAAG	Анализ TIDE
22	VEGF_TIDE7_RV	GTCTTCCTGCTCTGTGCGCAC GAC	Анализ TIDE
23	RandomPAM_FW	TAGCGGCCGCTCATGCGCGG CGCATTACCTTTA CNNNNNNNNNNGGATCCTCT AGAGTCG	Протоспейсер с рандомизированным PAM
24	RandomPAM_RV	ACAGGAAACAGCTATGACCA TGAAAGCTTGCATG CCTGCAGGTCGACTCTAGAG	Протоспейсер с рандомизированным PAM

		GATC	
25	DS2_ON_FW1	ctacacgacgctcttccgatctCCTGGAG CGTGTACGTTGG	Направленное глубокое секвенирование
26	SpyDS2_OT1_FW1	ctacacgacgctcttccgatctCCTGTGGT CCCAGCTACTTG	Направленное глубокое секвенирование
27	SpyDS2_OT2_FW1	ctacacgacgctcttccgatctATCTGCG ATGTCCTCGAGG	Направленное глубокое секвенирование
28	SpyDS2_OT3_FW1	ctacacgacgctcttccgatctTGGTGTGC GCCTCTAACG	Направленное глубокое секвенирование
29	SpyDS2_OT4_FW1	ctacacgacgctcttccgatctGGAGTCTT GCTTTGTCACTCAGA	Направленное глубокое секвенирование
30	SpyDS2_OT5_FW1	ctacacgacgctcttccgatctAGCCTAG ACCCAGTCCCAT	Направленное глубокое секвенирование
31	SpyDS2_OT6_FW1	ctacacgacgctcttccgatctGCTGGGC ATAGTAGTGGACT	Направленное глубокое секвенирование
32	SpyDS2_OT7_FW1	ctacacgacgctcttccgatctTGGGGAG GCTGAGACACGA	Направленное глубокое секвенирование
33	SpyDS2_OT8_FW1	ctacacgacgctcttccgatctCTTGGGA GGCTGAGGCAAG	Направленное глубокое секвенирование
34	DS2_ON_RV1	agacgtgtgctcttccgatctCAGGAGGA TGAGAGCCAGG	Направленное глубокое секвенирование
35	SpyDS2_OT1_RV1	agacgtgtgctcttccgatctCAGGGTCT CACTCTATCACCCA	Направленное глубокое секвенирование
36	SpyDS2_OT2_RV1	agacgtgtgctcttccgatctACTGAATG GGTTGAACTTGGC	Направленное глубокое секвенирование
37	SpyDS2_OT3_RV1	agacgtgtgctcttccgatctGAGACAGA ATCTTGCTCTGTCTCC	Направленное глубокое секвенирование
38	SpyDS2_OT4_RV1	agacgtgtgctcttccgatctTCCCAGCT ACTTGGGAGGC	Направленное глубокое секвенирование
39	SpyDS2_OT5_RV1	agacgtgtgctcttccgatctCCTGCCCA AATAGGGAAGCAG	Направленное глубокое секвенирование
40	SpyDS2_OT6_RV1	agacgtgtgctcttccgatctTGGCGCCT	Направленное глубокое секвенирование

		TAGTCTCTGCTAC	секвенирование
41	SpyDS2_OT7_RV1	agacgtgtgctcttccgatctGCATGAGA CACAGTTTCACTCTG	Направленное глубокое секвенирование
42	SpyDS2_OT8_RV1	agacgtgtgctcttccgatctGAGAGAGT CTCACTGCGTTGC	Направленное глубокое секвенирование
43	DS4_ON_FW3	ctacacgacgctcttccgatctTCTCTCAC CCACTGGGCAC	Направленное глубокое секвенирование
44	DS4_ON_RV3	agacgtgtgctcttccgatctGCTTCCAG ACGAGTGCAGA	Направленное глубокое секвенирование
45	SpyDS4_OT1_FW1	ctacacgacgctcttccgatctAAGTTTTC AAACCAGAAGAАCTACGAC	Направленное глубокое секвенирование
46	SpyDS4_OT2_FW1	ctacacgacgctcttccgatctCCGGTAT AAGTCCTGGAGCG	Направленное глубокое секвенирование
47	SpyDS4_OT3_FW1	ctacacgacgctcttccgatctGCCAGGG AGCAATGGCAG	Направленное глубокое секвенирование
48	SpyDS4_OT6_FW1	ctacacgacgctcttccgatctCCTCGAAT TCCACGGGGTT	Направленное глубокое секвенирование
49	DS16_ON_FW1	ctacacgacgctcttccgatctGTTGGTG GGAGGGAAGTGAG	Направленное глубокое секвенирование
50	SpyDS6_OT1_FW1	ctacacgacgctcttccgatctGATGGCG GTTGTAGCGGC	Направленное глубокое секвенирование
51	SpyDS6_OT2_FW1	ctacacgacgctcttccgatctCACATAA ACCTATGTTTCAGCAGA	Направленное глубокое секвенирование
52	SpyDS6_OT3_FW1	ctacacgacgctcttccgatctGCTAGTTG GATTGAAGCAGGGT	Направленное глубокое секвенирование
53	SpyDS6_OT4_FW1	ctacacgacgctcttccgatctTTGAGTGC GGCAGCTTCC	Направленное глубокое секвенирование
54	SpyDS6_OT6_FW1	ctacacgacgctcttccgatctATAACCCT CCCAGGCAAAGTC	Направленное глубокое секвенирование
55	SpyDS6_OT7_FW1	ctacacgacgctcttccgatctAGCCTGC ACATCTGAGCTC	Направленное глубокое секвенирование
56	SpyDS6_OT8_FW1	ctacacgacgctcttccgatctGGAGCAT TGAAGTGCCTGG	Направленное глубокое секвенирование

57	DeDS6_ON_RV1	agacgtgtgctcttccgatctCAGCCTGG GACCACTGA	Направленное глубокое секвенирование
58	SpyDS6_OT1_RV1	agacgtgtgctcttccgatctCATCCTCG ACAGTCGCGG	Направленное глубокое секвенирование
59	SpyDS6_OT2_RV1	agacgtgtgctcttccgatctGACTGATC AAGTAGAATACTCATGGG	Направленное глубокое секвенирование
60	SpyDS6_OT3_RV1	agacgtgtgctcttccgatctCCCTGCCA GCACTGAAGC	Направленное глубокое секвенирование
61	SpyDS6_OT4_Rv1	agacgtgtgctcttccgatctGGTTCCTA TCTTTCTAGACCAGGAGT	Направленное глубокое секвенирование
62	SpyDS6_OT6_RV1	agacgtgtgctcttccgatctAGTGTGGA GGGCTCAGGG	Направленное глубокое секвенирование
63	SpyDS6_OT7_RV1	agacgtgtgctcttccgatctGATGGGCA GAGGAAGGCAA	Направленное глубокое секвенирование
64	SpyDS6_OT8_RV1	agacgtgtgctcttccgatctTCACTCTC ATGAGCGTCCCA	Направленное глубокое секвенирование
65	Nme2DS2_OT1_FW1	ctacacgacgctcttccgatctAAGGTTC CTTGCGGTTCGC	Направленное глубокое секвенирование
66	Nme2DS2_OT1_RV1	agacgtgtgctcttccgatctCGCTGCCA TTGCTCCCT	Направленное глубокое секвенирование
67	Nme2DS6_OT1_FW1	ctacacgacgctcttccgatctTCTCGCAC ATTCTTCACGTCC	Направленное глубокое секвенирование
68	Nme2DS6_OT1_RV1	agacgtgtgctcttccgatctAGGAACCT TCCCGACTTAGGG	Направленное глубокое секвенирование
69	Rosa26_ON_FW1	ctacacgacgctcttccgatctCCCGCCC ATCTTCTAGAAAGAC	Направленное глубокое секвенирование
70	Rosa26_OT1_FW1	ctacacgacgctcttccgatctTGCCAGG TGAGGGACTGG	Направленное глубокое секвенирование
71	Rosa26_ON_RV1	agacgtgtgctcttccgatctTCTGGGAG TTCTCTGCTGCC	Направленное глубокое секвенирование
72	Rosa26_OT1_RV1	agacgtgtgctcttccgatctTGCCCAAC CTTAGCAAGGAG	Направленное глубокое секвенирование
73	pCSK9_ON_FW2	ctacacgacgctcttccgatcttaccttgagca	Направленное глубокое

		acggcg	секвенирование
74	PCSK9_ON_RV2	agacgtgtgctctccgatctcccaggacgagg atggag	Направленное глубокое секвенирование
75	Tyr_500_FW3	GATAGTCACTCCAGGGGTTG	Анализ TIDE
76	Tyr_500_RV3	GTGGTGAACCAATCAGTCCT	Анализ TIDE

#### Доставка RNP для редактирования генома млекопитающих

Для экспериментов с RNP использовали систему электропорации Neon в точности как описано (Amrani et al., 2018). В кратком изложении, 40 пикомоль  $3 \times \text{NLS-Nme2Cas9}$  вместе с 50 пикомоль транскрибированной посредством T7 sg-РНК объединяли в буфере R и вводили посредством электропорации с использованием 10-мкл наконечников Neon. После электропорации клетки высевали в предварительно нагретые 24-ячеечные планшеты, содержавшие соответствующую культуральную среду без антибиотиков. Параметры электропорации (напряжение, ширина, количество импульсов) представляли собой 1150 В, 20 мс, 2 импульса для клеток HEK293T; 1000 В, 50 мс, 1 импульс для клеток K562.

#### Доставка AAV8.Nme2Cas9+sgRNA *in vivo* и процессинг тканей печени

Для инъекций вектора AAV8 самкам мышей C57BL/6NJ в возрасте 8 недель инъецировали  $4 \times 10^{11}$  копий генома на мышью через хвостовую вену с sg-РНК, нацеленной на подтвержденный участок либо в *Pcsk9*, либо в *Rosa26*. Мышей умерщвляли через 28 суток после введения вектора и ткани печени собирали для анализа. Ткани печени фиксировали в 4% формалине в течение ночи, заливали парафином, нарезают на срезы и окрашивали гематоксилином и эозином (H&E). Проводили взятие крови из лицевой вены через 0, 14 и 28 суток после инъекции и сыворотку выделяли с использованием сепаратора сыворотки (BD, каталожный номер № 365967) и хранили при  $-80^\circ\text{C}$  до анализа. Сывороточный уровень холестерина измеряли с использованием колориметрического анализа Infinity™ (Thermo-Scientific) в соответствии с протоколом изготовителя и как описано ранее (Ibraheim et al., 2018). Для вестерн-блоттинга PCSK9, 40 мкг белка из ткани или 2 нг рекомбинантного белка PCSK9 мыши (R&D Systems, 9258-SE-020) наносили на готовый гель MiniPROTEAN® TGX™ Precast Gel (Bio-Rad). Отделенные полосы переносили на мембрану PVDF и блокировали раствором 5% Blocking-Grade Blocker solution (Bio-Rad) в течение 2 часов при комнатной температуре. Далее мембрану инкубировали с антителами кролика против GAPDH (Abcam ab9485, 1:2000) или антителами козы против PCSK9 (R&D Systems AF3985, 1:400) в течение ночи. Мембраны промывали TBST и инкубировали с конъюгированными с пероксидазой хрена (HRP) вторичными антителами козы против антител кролика (Bio-Rad 1706515, 1:4000) и вторичными антителами осла против антител козы (R&D Systems HAF109, 1:2000) в течение 2 часов при комнатной температуре. Мембраны вновь промывали TBST и визуализировали с использованием субстрата Clarity™ western ECL substrate (Bio-Rad) с

использованием M35A XOMAT Processor (Kodak).

#### Доставка AAV6.Nme2Cas9 ex vivo в мышинные зиготы

Зиготы инкубировали в 15-мкл каплях KSOM (дополненная калием симплексная оптимизированная среда, Millipore, каталожный номер № MR-106-D), содержащей  $3 \times 10^9$  или  $3 \times 10^8$  GC вектора AAV6.Nme2Cas9.sgTyr, в течение 5-6 ч (4 зиготы в каждой капле). После инкубации зиготы ополаскивали в M2 и переносили в свежую KSOM для культивирования в течение ночи. На следующие сутки эмбрионы, которые достигли стадии 2-клеток, переносили в яйцеводы псевдобеременных реципиентов, и им позволяли развиваться до полного срока.

#### **Экспериментальная часть**

##### *Пример I*

#### Обнаружение ортологов Cas9 с дифференциально дивергировавшими PID

Последовательность пептида из Nme1Cas9 использовали в качестве последовательности запроса при поиске в BLAST для поиска всех ортологов Cas9 в виде *Neisseria meningitidis*. Ортологи с >80% идентичностью с Nme1Cas9 были отображены для следующей части исследования. Затем PID выравнивали с PID Nme1Cas9 (остатки 820-1082) с использованием ClustalW2, и те из них, в которых были кластеры мутаций в PID, отбирали для дальнейшего анализа. Некорневое филогенетическое древо ортологов NmeCas9 конструировали с использованием FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

##### *Пример II*

#### Клонирование, экспрессия и очистка ортологов Cas9 и Acr

Примеры плазмид и олигонуклеотидов, использованных в этом исследовании, приведены в таблице 3. PID Nme2Cas9 и Nme3Cas9 заказывали в качестве gBlocks (IDT) для замены PID Nme1Cas9 с использованием Gibson Assembly (NEB) в бактериальной экспрессирующей плазмиде pMSCG7 (Zhang et al., 2015), которая кодирует Nme1Cas9 с 6×His-меткой. Конструкцию трансформировали в *E. coli*, экспрессировали и очищали, как описано ранее (Pawluk et al., 2016). В кратком изложении, клетки Rosetta (DE3), содержавшие соответствующие плазмиды Cas9, выращивали при 37°C до OD<sub>600</sub> 0,6 и экспрессию белка индуцировали 1 mM IPTG в течение 16 ч при 18°C. Клетки собирали и лизировали посредством обработки ультразвуком в лизирующем буфере [50 mM Tris-HCl (pH 7,5), 500 mM NaCl, 5 mM имидазол, 1 mM DTT], дополненном 1 мг/мл лизоцима и коктейлем ингибиторов протеаз (Sigma). Затем лизат разделяли через колонку с агарозой Ni<sup>2+</sup>-NTA (Qiagen), и связанный белок элюировали посредством 300 mM имидазола и подвергали диализу в буфер для хранения [20 mM HEPES-NaOH (pH 7,5), 250 mM NaCl, 1 mM DTT]. Для белков Acr, 6×His-меченные белки экспрессировали в *E. coli* штамма BL21 Rosetta (DE3). Клетки выращивали при 37°C до оптической плотности (OD<sub>600</sub>) 0,6 в качающемся инкубаторе. Бактериальные культуры охлаждали до 18°C и экспрессию белка индуцировали добавлением 1 mM IPTG для экспрессии в течение ночи. На следующие

сутки клетки собирали и ресуспендировали в лизирующем буфере, дополненном 1 мг/мл лизоцима и коктейля ингибиторов протеаз (Sigma) и белок очищали с использованием того же протокола, что и для Cas9. 6×His-метку удаляли посредством инкубации связанного со смолой белка с протеазой вируса гравировки табака (TEV) в течение ночи при 4°C для выделения немеченых Асг.

### *Пример III*

#### Анализ для выявления РАМ *in vitro*

Библиотеку дцДНК-мишеней с рандомизированными последовательностями РАМ получали посредством перекрывающейся ПЦР с прямым праймером, содержащим рандомизированную область РАМ из 10 нуклеотидов. Библиотеку очищали из геля и подвергали реакции расщепления *in vitro* посредством очищенной Cas9 вместе с транскрибированными T7 sg-РНК. 300 нМ комплекс Cas9:sg-РНК использовали для расщепления 300 нМ фрагмента-мишени в 1× NEBuffer 3.1 (NEB) при 37°C в течение 1 ч. Затем реакционную смесь обрабатывали протеазой К при 50°C в течение 10 минут и разделяли на 4% геле агароза/1×ТАЕ. Продукт расщепления вырезали, элюировали и клонировали с использованием ранее описанного протокола (Zhang et al., 2012) с модификациями. В кратком изложении концы ДНК репарировали, добавляли нематричные хвостовые части из 2'-дезоксиденозина и лигировали Y-образные адаптеры. После ПЦР продукт количественно определяли посредством набора КАРА Library Quantification Kit и секвенировали с использованием NextSeq 500 (Illumina) с получением чтений с парными концами из 75 нуклеотидов. Последовательности анализировали с использованием специализированных скриптов и R.

### *Пример IV*

#### Трансфекции и редактирование генома млекопитающих

Кодон-оптимизированную Nme2Cas9 человека клонировали в Gibson Assembly в плазмидный остов pCDest2, ранее использованный для экспрессии Nme1Cas9 и SpyCas9 (Pawluk et al., 2016; Amrani et al., 2018). Трансфекцию клеток НЕК293Т и НЕК293Т-TLR2.0 проводили, как описано ранее (Amrani et al., 2018). Для трансфекций Нера1-6 липофектамин LTX использовали для трансфицирования 500 нг плазмиды AAV.sgRNA.Nme2Cas9 "все в одном" в 24-ячеечных планшетах (~10<sup>5</sup> клеток/лунка), с использованием клеток, культивированных в течение 24 часов перед трансфекцией. Для клеток K562, стабильно экспрессирующих Nme2Cas9, доставленную посредством лентивирусного вектора (см. ниже), 50000-150000 клеток вводили посредством электропорации с 500 нг плазмиды sg-РНК с использованием 10-мкл наконечников Neon. Для количественного определения инсерций-делеций во всех клетках через 72 ч после трансфекций клетки собирали и геномную ДНК экстрагировали с использованием набора DNAesy Blood and Tissue kit (Qiagen). Локус-мишень амплифицировали посредством ПЦР, секвенировали по методу Сэнгера (Genewiz) и анализировали посредством TIDE (Brinkman et al., 2014) с использованием веб-интерфейса Desktop Genetics

(<http://tide.deskgen.com>).

*Пример V*

Лентивирусная трансдукция клеток K562 для стабильной экспрессии Nme2Cas9

Клетки K562, стабильно экспрессирующие Nme2Cas9, получали, как описано ранее для Nme1Cas9 (Amrani et al., 2018). Для получения лентивируса лентивирусный вектор котрансфицировали в клетки HEK293T вместе с упаковывающими плазмидами (Addgene 12260 & 12259) в 6-луночных планшетах с использованием реагента для трансфекции TransIT-LT1 (Mirus Bio). Через 24 часа культуральную среду аспирировали из трансфицированных клеток и заменяли на 1 мл свежей DMEM. На следующие сутки супернатант, содержащий вирус, собирали и фильтровали через 0,45-мкм фильтр. 10 мкл неразбавленного супернатанта вместе с 2,5 мкг Polybrene использовали для трансдукции  $\sim 10^6$  клеток K562 в 6-ячеечных планшетах. Селекцию трансдуцированных клеток проводили с использованием среды, дополненной 2,5 мкг/мл пурамицина.

*Пример VI*

Доставка RNP для редактирования генома млекопитающих

Для экспериментов с RNP систему электропорации Neon использовали в точности, как описано (Amrani et al., 2018). В кратком изложении, 40 пикомоль  $3 \times \text{NLS-Nme2Cas9}$  вместе с 50 пикомоль транскрибированной T7 sg-РНК собирали в буфере R и вводили посредством электропорации с использованием 10-мкл наконечников Neon. После электропорации клетки высевали в предварительно нагретые 24-ячеечные планшеты, содержавшие соответствующую культуральную среду без антибиотиков. Параметры электропорации (напряжение, ширина, количество импульсов) составляли 1150 В, 20 мс, 2 импульса для клеток HEK293T; 1000 В, 50 мс, 1 импульс для клеток K562.

*Пример VII*

GUIDE-seq

Эксперименты GUIDE-seq проводили, как описано ранее (Tsai et al., 2014), с незначительными модификациями (Bolukbasi et al., 2015a). В кратком изложении, клетки HEK293T трансфицировали 200 нг плазмиды Cas9, 200 нг плазмиды sg-РНК, и 7,5 пмоль гибридных олигонуклеотидов GUIDE-seq с использованием Polyfect (Qiagen). Альтернативно клетки Hepa1-6 трансфицировали, как описано выше. Геномную ДНК экстрагировали с использованием набора DNeasy Blood and Tissue kit (Qiagen) через 72 ч после трансфекции в соответствии с протоколом изготовителя. Получение и секвенирование библиотеки проводили в точности, как описано ранее (Bolukbasi et al., 2015a). Для анализа все последовательности с вплоть до десяти несоответствиями с участком-мишенью, а также С в пятом положении PAM ( $\text{N}_4\text{CN}$ ), считались потенциальными участками вне мишени. Данные анализировали с использованием пакета Bioconductor GUIDESeq версии 1.1.17 (Zhu et al., 2017).

*Пример VIII*

Направленное глубокое секвенирование и анализ

Авторы изобретения использовали направленное глубокое секвенирование для подтверждения результатов GUIDE-seq и для измерения частот инсерций-делеций с максимальной точностью. Авторы изобретения использовали двухстадийную амплификацию способом ПЦР для получения фрагментов ДНК для каждого участка в мишени и участка вне мишени. Для редактирования посредством SpyCas9 в DS2 и DS6, авторы изобретения отобрали основные участки вне мишени на основе количеств чтений посредством GUIDE-seq. Для редактирования посредством SpyCas9 в DS4 посредством GUIDE-seq было идентифицировано меньшее количество участков вне мишени, являющихся кандидатами, и только участки с PAM NGG (DS4|OT1, DS4|OT3, DS4|OT6) или NGC (DS4|OT2) исследовали посредством секвенирования. На первой стадии авторы изобретения использовали локус-специфические праймеры, содержащие универсальные выступающие концы, окончания которых бил комплементарными адаптерам. На первой стадии 2× основную смесь для ПЦР (NEB) использовали для получения фрагментов, имеющих выступающие концы. На второй стадии очищенные продукты ПЦР амплифицировали с универсальным прямым праймером и индексированными обратными праймерами. Полноразмерные продукты (~250 п.н.) очищали из геля и секвенировали на Illumina MiSeq в режиме спаренных концов. Анализ данных MiSeq проводили, как описано ранее (Pinello et al., 2016; Ibraheim et al., 2018).

#### *Пример IX*

##### Анализ действия вне мишени с использованием CRISPRseek

Глобальное прогнозирование связывания вне мишени для TS25 и TS47 проводили с использованием пакета CRISPRseek от Bioconductor. Незначительные изменения вносили для учета характеристик Nme2Cas9, не являющихся общими с SpyCas9. В частности, авторы изобретения использовали следующие изменения: gRNA.size=24, PAM = "NNNNCC", PAM.size=6, RNA.PAM.pattern = "NNNNCN", и получали участки вне мишени, являющиеся кандидатами, с менее чем 6 несоответствиями. Основные потенциальные участки вне мишени отбирали на основе числа и положений несоответствий. Геномную ДНК из клеток, на которую нацелена каждая соответствующая sg-РНК, использовали для амплификации каждого локуса вне мишени, являющегося кандидатом, а затем анализировали посредством TIDE.

#### *Пример X*

##### Линии мышей и сбор эмбрионов

Все эксперименты проводили под руководством Комитета по содержанию и использованию лабораторных животных (IACUC) University of Massachusetts Medical School. C57BL/6NJ (Stock No. 005304). Мышей получали от The Jackson Laboratory. Всех животных содержали с циклом с 12 ч на свету. Середину светового дня цикла, когда наблюдали копулятивную пробку, считали сутками 0,5 эмбрионального развития (E0,5). Зиготы извлекали на E0,5 посредством разрыва ампулы щипцами и инкубации в среде M2, содержащей гиалуронидазу, для удаления скоплений клеток.

*Пример XI*Доставка AAV8.Nme2Cas9+sgPНК *in vivo* и обработка ткани печени

Для инъекций вектора AAV8 самкам мышей C57BL/6NJ в возрасте 8 недель инъецировали  $4 \times 10^{11}$  копий генома на мышью через хвостовую вену с sg-PНК, нацеленной на подтвержденный участок либо в *Pcsk9*, либо в *Rosa26*. Мышей умерщвляли через 28 суток после введения вектора и ткани печени собирали для анализа. Ткани печени фиксировали в 4% формалине в течение ночи, заливали парафином, нарезали на срезы и окрашивали гематоксилином и эозином (H&E). Проводили взятие крови из лицевой вены через 0, 14 и 28 суток после инъекции и сыворотку выделяли с использованием сепаратора сыворотки (BD, каталожный номер № 365967) и хранили при  $-80^{\circ}\text{C}$  до анализа. Сывороточный уровень холестерина измеряли с использованием колориметрического анализа Infinity™ (Thermo-Scientific) в соответствии с протоколом изготовителя и как описано ранее (Ibraheim et al., 2018). Для вестерн-блоттинга PCSK9, 40 мкг белка из ткани или 2 нг рекомбинантного белка PCSK9 мыши (R&D Systems, 9258-SE-020) наносили на готовый гель MiniPROTEAN® TGX™ Precast Gel (Bio-Rad). Отделенные полосы переносили на мембрану PVDF и блокировали раствором 5% Blocking-Grade Blocker solution (Bio-Rad) в течение 2 часов при комнатной температуре. Далее мембрану инкубировали с антителами кролика против GAPDH (Abcam ab9485, 1:2000) или антителами козы против PCSK9 (R&D Systems AF3985, 1:400) в течение ночи. Мембраны промывали TBST и инкубировали с конъюгированными с пероксидазой хрена (HRP) вторичными антителами козы против антител кролика (Bio-Rad 1706515, 1:4000) и вторичными антителами осла против антител козы (R&D Systems HAF109, 1:2000) в течение 2 часов при комнатной температуре. Мембраны вновь промывали TBST и визуализировали с использованием субстрата Clarity™ western ECL substrate (Bio-Rad) с использованием M35A XOMAT Processor (Kodak).

*Пример XII*Доставка AAV6.Nme2Cas9 в зиготы мышей *ex vivo*

Зиготы инкубировали в 15-мкл каплях KSOM (дополненная калием симплексная оптимизированная среда, Millipore, каталожный номер № MR-106-D), содержащей  $3 \times 10^9$  или  $3 \times 10^8$  GC вектора AAV6.Nme2Cas9.sgTyr, в течение 5-6 ч (4 зиготы в каждой капле). После инкубации зиготы ополаскивали в M2 и переносили в свежую KSOM для культивирования в течение ночи. На следующие сутки эмбрионы, которые достигли стадии 2-клеток, переносили в яйцеводы псевдобеременных реципиентов, и им позволяли развиваться до полного срока.

**ССЫЛКИ**, каждая из которых включена в качестве ссылки в полном объеме:

Amrani, N., Gao, X.D., Liu, P., Edraki, A., Mir, A., Ibraheim, R., Gupta, A., Sasaki, K.E., Wu, T., Donohoue, P.D., et al. (2018). NmeCas9 is an intrinsically high-fidelity genome editing platform. BioRxiv, <https://doi.org/10.1101/172650>.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S.,

Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709-1712.

Bisaria, N., Jarmoskaite, I., and Herschlag, D. (2017). Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. *Cell Syst.* 4, 21-29.

Bolukbasi, M.F., Gupta, A., Oikemus, S., Derr, A.G., Garber, M., Brodsky, M.H., Zhu, L.J., and Wolfe, S.A. (2015a). DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat. Methods* 12, 1150-1156.

Bolukbasi, M.F., Gupta, A., and Wolfe, S.A. (2015b). Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. *Nat. Methods* 13, 41-50.

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 42, e168.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964.

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., et al. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36, 265-271.

Certo, M.T., Ryu, B.Y., Annis, J.E., Garibov, M., Jarjour, J., Rawlings, D.J., and Scharenberg, A.M. (2011). Tracking genome engineering outcome at individual DNA breakpoints. *Nat. Methods* 8, 671-676.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407-410.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 230-232.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132-141.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607.

Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390-1400.

Dominguez, A.A., Lim, W.A., and Qi, L.S. (2016). Beyond editing: repurposing

CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* 17, 5-15.

Dong, Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., and Huang, Z. (2017). Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* 546, 436-439.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116-1121.

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lecrivain, A.L., Bzdrenga, J., Koonin, E.V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 2577-2590.

Friedland, A.E., Baral, R., Singhal, P., Loveluck, K., Shen, S., Sanchez, M., Marco, E., Gotta, G.M., Maeder, M.L., Kennedy, E.M., et al. (2015). Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol.* 16, 257.

Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* 5, 1513-1523.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279-284.

Gallagher, D.N., and Haber, J.E. (2018). Repair of a Site-Specific DNA Cleavage: Old-School Lessons for Cas9-Mediated Gene Editing. *ACS Chem. Biol.* 13, 397-405.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67-71.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* 109, E2579-2586.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A\*T to G\*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471.

Ghanta, K., Dokshin, G., Mir, A., Krishnamurthy, P., Gneid, H., Edraki, A., Watts, J., Sontheimer, E., and Mello, C. (2018). 5' Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. *Biorxiv* 354480.

Gorski, S.A., Vogel, J., and Doudna, J.A. (2017). RNA-based recognition and targeting: sowing the seeds of specificity. *Nat. Rev. Mol. Cell Biol.* 18, 215-228.

Harrington, L.B., Doxzen, K.W., Ma, E., Liu, J.J., Knott, G.J., Edraki, A., Garcia, B., Amrani, N., Chen, J.S., Cofsky, J.C., et al. (2017a). A Broad-Spectrum Inhibitor of CRISPR-Cas9. *Cell* 170, 1224-1233.

- Harrington, L.B., Paez-Espino, D., Staahl, B.T., Chen, J.S., Ma, E., Kyrpides, N.C., and Doudna, J.A. (2017b). A thermostable Cas9 with increased lifetime in human plasma. *Nat. Commun.* 8, 1424.
- Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA* 110, 15644-15649.
- Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57-63.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227-229.
- Hynes, A.P., Rousseau, G.M., Lemay, M.-L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. *Nat. Microbiol.* 2, 1374-1380.
- Ibraheim, R., Song, C.-Q., Mir, A., Amrani, N., Xue, W., and Sontheimer, E.J. (2018). All-in-One Adeno-associated Virus Delivery and Genome Editing by *Neisseria meningitidis* Cas9 in vivo. *BioRxiv*, <https://doi.org/10.1101/295055>.
- Jiang, F., and Doudna, J.A. (2017). CRISPR-Cas9 Structures and Mechanisms. *Annu. Rev. Biophys.* 46, 505-529.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233-239.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821.
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* 2, e00471.
- Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol.* 16, 253.
- Keeler, A.M., ElMallah, M.K., and Flotte, T.R. (2017). Gene Therapy 2017: Progress and Future Directions. *Clin. Transl. Sci.* 10, 242-248.
- Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K.-E., Kim, K., Cho, H.-Y., Song, D.W., Lee, K.J., Jung, M.H., et al. (2017). In vivo genome editing with a small Cas9 ortholog derived from *Campylobacter jejuni*. *Nat. Commun.* 8, 14500.
- Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012-1019.
- Kim, B., Komor, A., Levy, J., Packer, M., Zhao, K., and Liu, D. (2017). Increasing the

genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nature Biotechnology* 35.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Topkar, V.V., Zheng, Z., and Joung, J.K. (2015). Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* 33, 1293-1298.

Kluesner, M., Nedveck, D., Lahr, W., Garbe, J., Abrahante, J., Webber, B., and Moriarity, B. (2018). EditR: A Method to Quantify Base Editing from Sanger Sequencing. *The CRISPR Journal* 1, 239-250.

Koblan, L., Doman, J., Wilson, C., Levy, J., Tay, T., Newby, G., Maianti, J., Raguram, A., and Liu, D. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* 36, 843.

Komor, A.C., Badran, A.H., and Liu, D.R. (2017). CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* 168, 20-36.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424.

Lee, C.M., Cradick, T.J., and Bao, G. (2016). The *Neisseria meningitidis* CRISPR-Cas9 system enables specific genome editing in mammalian cells. *Mol. Ther.* 24, 645-654.

Lee, J., Mir, A., Edraki, A., Garcia, B., Amrani, N., Lou, H.E., Gainetdinov, I., Pawluk, A., Ibraheim, R., Gao, X.D., et al. (2018). Potent Cas9 inhibition in bacterial and human cells by new anti-CRISPR protein families. *BioRxiv*, <https://www.biorxiv.org/content/early/2018/2006/2020/350504>.

Ma, E., Harrington, L.B., O'Connell, M.R., Zhou, K., and Doudna, J.A. (2015). Single-Stranded DNA Cleavage by Divergent CRISPR-Cas9 Enzymes. *Mol. Cell* 60, 398-407.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833-838.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. *Science* 339, 823-826.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843-1845.

Mir, A., Edraki, A., Lee, J., and Sontheimer, E.J. (2018). Type II-C CRISPR-Cas9 biology, mechanism and application. *ACS Chem. Biol.* 13, 357-365.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733-740.

Paez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B.C., Barrangou, R., and Banfield, J.F. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus*

*thermophilus*. *mBio* 6.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., et al. (2016). Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167, 1829-1838 e1829.

Pawluk, A., Bondy-Denomy, J., Cheung, V.H., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* 5, e00896.

Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., and Yuan, G.C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat. Biotechnol.* 34, 695-697.

Racanelli, V., and Rehermann, B. (2006). The liver as an immunological organ. *Hepatology* 43, S54-62.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186-191.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389.

Rashid, S., Curtis, D.E., Garuti, R., Anderson, N.N., Bashmakov, Y., Ho, Y.K., Hammer, R.E., Moon, Y.A., and Horton, J.D. (2005). Decreased plasma cholesterol and hypersensitivity to statins in mice lacking *Pcsk9*. *Proc. Natl. Acad. Sci. USA* 102, 5374-5379.

Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and Bondy-Denomy, J. (2017). Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* 168, 150-158 e110.

Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275-9282.

Schumann, K., Lin, S., Boyer, E., Simeonov, D.R., Subramaniam, M., Gate, R.E., Haliburton, G.E., Ye, C.J., Bluestone, J.A., Doudna, J.A., et al. (2015). Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci. USA* 112, 10437-10442.

Shin, J., Jiang, F., Liu, J.J., Bray, N.L., Rauch, B.J., Baik, S.H., Nogales, E., Bondy-Denomy, J., Corn, J.E., and Doudna, J.A. (2017). Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* 3, e1701620.

Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet.* 17, 300-312.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2014). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187-197.

Tycko, J., Myer, V.E., and Hsu, P.D. (2016). Methods for optimizing CRISPR-Cas9 genome editing specificity. *Mol. Cell* 63, 355-370.

Yang, H., and Patel, D.J. (2017). Inhibition Mechanism of an Anti-CRISPR Suppressor AcrIIA4 Targeting SpyCas9. *Mol Cell* 67, 117-127 e115.

Yin, H., Song, C.Q., Suresh, S., Kwan, S.Y., Wu, Q., Walsh, S., Ding, J., Bogorad, R.L., Zhu, L.J., Wolfe, S.A., et al. (2018). Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat. Chem. Biol.* 14, 311-316.

Yokoyama, T., Silversides, D.W., Waymire, K.G., Kwon, B.S., Takeuchi, T., and Overbeek, P.A. (1990). Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucleic Acids Res.* 18, 7293-7298.

Yoon, Y., Wang, D., Tai, P.W.L., Riley, J., Gao, G., and Rivera-Perez, J.A. (2018). Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses. *Nat. Commun.* 9, 412.

Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., Vogel, J., and Sontheimer, E.J. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* 50, 488-503.

Zhang, Y., Rajan, R., Seifert, H.S., Mondragón, A., and Sontheimer, E.J. (2015). DNase H activity of *Neisseria meningitidis* Cas9. *Mol. Cell* 60, 242-255.

Zhang, Z., Theurkauf, W.E., Weng, Z., and Zamore, P.D. (2012). Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. *Silence* 3, 9.

Zhu, L.J., Holmes, B.R., Aronin, N., and Brodsky, M.H. (2014). CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. *PLoS One* 9, e108424.

Zhu, L.J., Lawrence, M., Gupta, A., Pagés, H., Kucukural, A., Garber, M., and Wolfe, S.A. (2017). GUIDEseq: a bioconductor package to analyze GUIDE-Seq datasets for CRISPR-Cas nucleases. *BMC Genomics* 18, 379.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 33, 73-80.

Все публикации и патенты, упомянутые в приведенном выше описании, включены в настоящее описание в качестве ссылок. Различные модификации и отклонения от описанных способов и систем по изобретению будут понятными специалистам в данной области без отклонения от объема и сущности изобретения. Хотя изобретение описано применительно к конкретным предпочтительным вариантам осуществления, следует понимать, что заявленное изобретение не должно быть чрезмерно ограничено такими конкретными вариантами осуществления. Действительно, подразумевается, что различные модификации описанных способов осуществления изобретения, которые

очевидны специалистам в области биологического контроля, биохимии, молекулярной биологии, энтомологии, планктонных организмов, систем рыболовства и экологии пресной воды, или родственных областях, входят в объем прилагаемой ниже формулы изобретения.

**ФОРМУЛА ИЗОБРЕТЕНИЯ**

1. Мутантный белок NmeCas9, содержащий слитую с ним нуклеотиддезаминазу и связывающую область для нуклеотидной последовательности N<sub>4</sub>CC.
2. Белок по п. 1, где указанный белок представляет собой Nme2Cas9.
3. Белок по п. 1, дополнительно содержащий белок сигнала ядерной локализации.
4. Белок по п. 1, где указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу.
5. Белок по п. 1, где указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу.
6. Белок по п. 1, дополнительно содержащий ингибитор урацилгликозилазы.
7. Белок по п. 1, где указанный белок сигнала ядерной локализации выбран из нуклеоплазмина и SV40.
8. Белок по п. 1, где указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера.
9. Белок по п. 8, где указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию.
10. Белок по п. 9, где указанная мутация представляет собой мутацию D16A.
11. Аденоассоциированный вирус, содержащий мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для нуклеотидной последовательности N<sub>4</sub>CC.
12. Вирус по п. 11, где указанный вирус представляет собой аденоассоциированный вирус 8.
13. Вирус по п. 11, где указанный вирус представляет собой аденоассоциированный вирус 6.
14. Вирус по п. 11, где указанный белок представляет собой Nme2Cas9.
15. Вирус по п. 11, где указанный белок дополнительно содержит белок сигнала ядерной локализации.
16. Вирус по п. 11, где указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу.
17. Вирус по п. 11, где указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу.
18. Вирус по п. 11, где указанный белок дополнительно содержит ингибитор урацилгликозилазы.
19. Вирус по п. 11, где указанный белок сигнала ядерной локализации выбран из нуклеоплазмина и SV40.
20. Вирус по п. 11, где указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера.
21. Вирус по п. 20, где указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию.
22. Вирус по п. 21, где указанная мутация представляет собой мутацию D16A.

23. Способ, включающий:
- a) предоставление;
    - i) нуклеотидной последовательности, содержащей ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC;
    - ii) мутантный белок NmeCas9, содержащий слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC;
  - b) приведение указанной нуклеотидной последовательности в контакт с указанным мутантным белком NmeCas9 в таких условиях, чтобы указанная связывающая область связывалась с указанной нуклеотидной последовательностью N<sub>4</sub>CC; и
  - c) замену указанного единичного основания на основание дикого типа посредством указанного мутантного белка NmeCas9.
24. Способ по п.23, где указанный белок представляет собой Nme2Cas9.
25. Способ по п.23, где указанный белок дополнительно содержит белок сигнала ядерной локализации.
26. Способ по п.23, где указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу.
27. Способ по п.23, где указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу.
28. Способ по п.23, где указанный белок дополнительно содержит ингибитор урацилгликозилазы.
29. Способ по п.23, где указанный белок сигнала ядерной локализации выбран из группы, состоящей из нуклеоплазмина и SV40.
30. Способ по п.23, где указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера.
31. Способ по п.30, где указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию белка Cas9.
32. Способ по п.31, где указанная мутация белка Cas9 представляет собой мутацию D16A.
33. Способ, включающий:
- a) предоставление;
    - i) пациента, имеющего нуклеотидную последовательность, содержащую ген с мутантным единичным основанием, где указанный ген фланкирован нуклеотидной последовательностью N<sub>4</sub>CC, где указанный мутантный ген вызывает медицинское состояние на генетической основе;
    - ii) аденоассоциированного вируса, включающего мутантный белок NmeCas9, причем указанный мутантный белок NmeCas9 содержит слитую с ним нуклеотиддезаминазу и связывающую область для указанной нуклеотидной последовательности N<sub>4</sub>CC;
  - b) лечение указанного пациента указанным аденоассоциированным вирусом в таких условиях, чтобы указанный мутантный белок NmeCas9 заменял указанное

мутантное единичное основание на единичное основание дикого типа, так чтобы указанное медицинское состояние на генетической основе не развивалось.

34. Способ по п.33, где указанный ген кодирует белок тирозиназу.

35. Способ по п.33, где указанное медицинское состояние на генетической основе представляет собой тирозинемия.

36. Способ по п.33, где указанный вирус представляет собой аденоассоциированный вирус 8.

37. Способ по п.33, где указанный вирус представляет собой аденоассоциированный вирус 6.

38. Способ по п.33, где указанный белок представляет собой Nme2Cas9.

39. Способ по п.33, где указанный белок дополнительно содержит белок сигнала ядерной локализации.

40. Способ по п.33, где указанная нуклеотиддезаминаза представляет собой цитидиндезаминазу.

41. Способ по п.33, где указанная нуклеотиддезаминаза представляет собой аденозиндезаминазу.

42. Способ по п.33, где указанный белок дополнительно содержит ингибитор урацилгликозилазы.

43. Способ по п.33, где указанный белок сигнала ядерной локализации выбран из группы, состоящей из нуклеоплазмина и SV40.

44. Способ по п.33, где указанная связывающая область представляет собой домен, взаимодействующий с вспомогательным мотивом протоспейсера.

45. Способ по п.44, где указанный домен, взаимодействующий с вспомогательным мотивом протоспейсера, содержит указанную мутацию.

46. Способ по п.45, где указанная мутация представляет собой мутацию D16A.

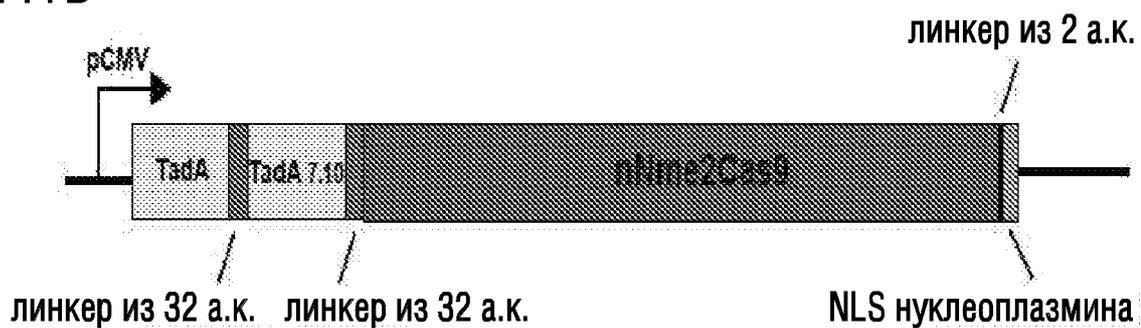
По доверенности

## ФИГ.1

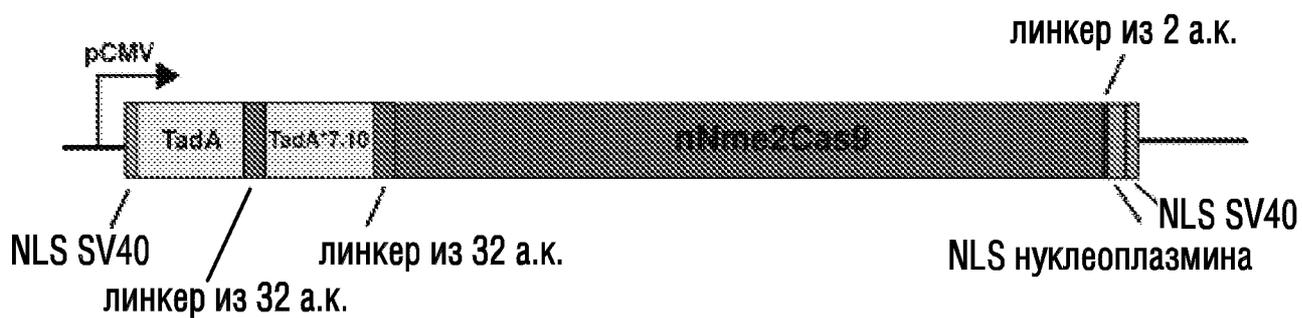
## ФИГ.1А



## ФИГ.1В



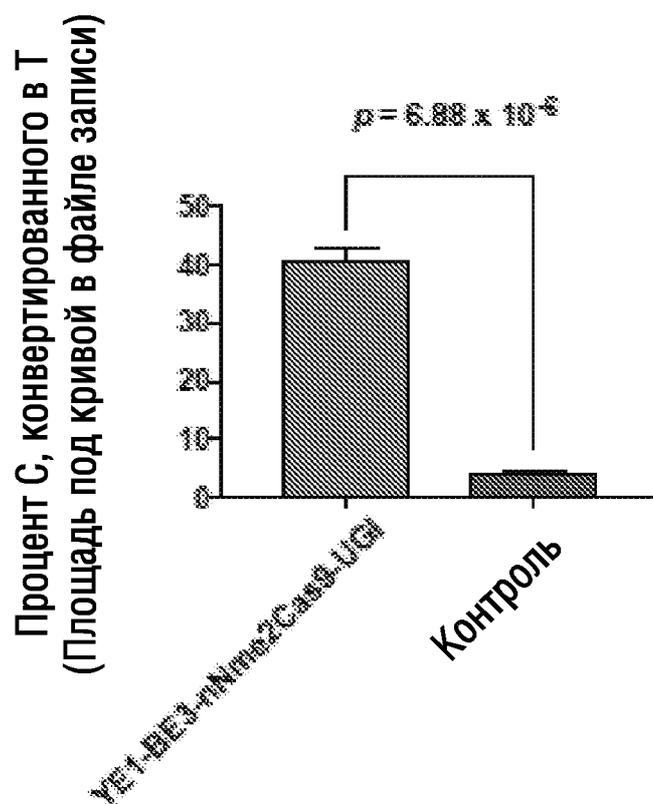
## ФИГ.1С





## ФИГ.2 (продолжение)

## ФИГ.2С



## ФИГ.3

## ФИГ.3А

EGFP - участок 1

																							PAM						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23							
5'	T	C	C	A	G	C	T	C	G	A	C	C	A	G	G	A	T	G	G	G	C	A	C	C	C	C	G	3'	эталон
	T	C	C	A	G	C	T	C	G	A	C	C	A	ФИГ.6D															94.83%
	T	C	C	A	G	C	T	C	G	A	C	C	A	G	G	A	T	G	G	G	C	A	C	C	C	C	G		2.00%
	T	C	C	A	G	C	T	C	G	A	C	C	A	G	G	A	T	G	G	G	C	A	C	C	C	C	G		1.76%
	T	C	C	A	G	C	T	C	G	A	C	C	A	G	G	A	T	G	G	G	C	A	C	C	C	C	G		0.24%

## ФИГ.3В

EGFP - участок 2

																							PAM										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23											
5'	G	G	T	C	G	A	G	C	T	G	G	A	C	G	G	C	G	A	C	G	T	A	A	A	C	G	G	C	C	A	3'	эталон	
	G	G	T	C	G	A	G	C	T	G	G	A	C	G	G	C	G	A	C	G	T	A	A	A	A	C	G	G	C	C	A		97.95%
	G	G	T	C	G	A	G	C	T	G	G	A	C	G	G	C	G	A	C	G	T	A	A	A	A	C	G	G	C	C	A		0.74%

## ФИГ.3С

EGFP - участок 3

																							PAM								
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23									
5'	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G	3'	эталон
	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G		97.78%
	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G		0.40%
	A	G	T	T	C	A	T	C	T	G	C	A	C	C	A	C	G	G	C	A	A	G	C	T	G	C	C	C	G		0.33%

## ФИГ.3D

EGFP - участок 4

																							PAM								
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23									
5'	C	G	T	G	C	C	C	T	G	G	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	C	T	3'	эталон
	C	G	T	G	C	C	C	T	G	G	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	C	T		97.93%
	C	G	T	G	C	C	C	T	G	G	C	C	A	C	C	C	T	C	G	T	G	A	C	C	A	C	C	C	T		0.35%

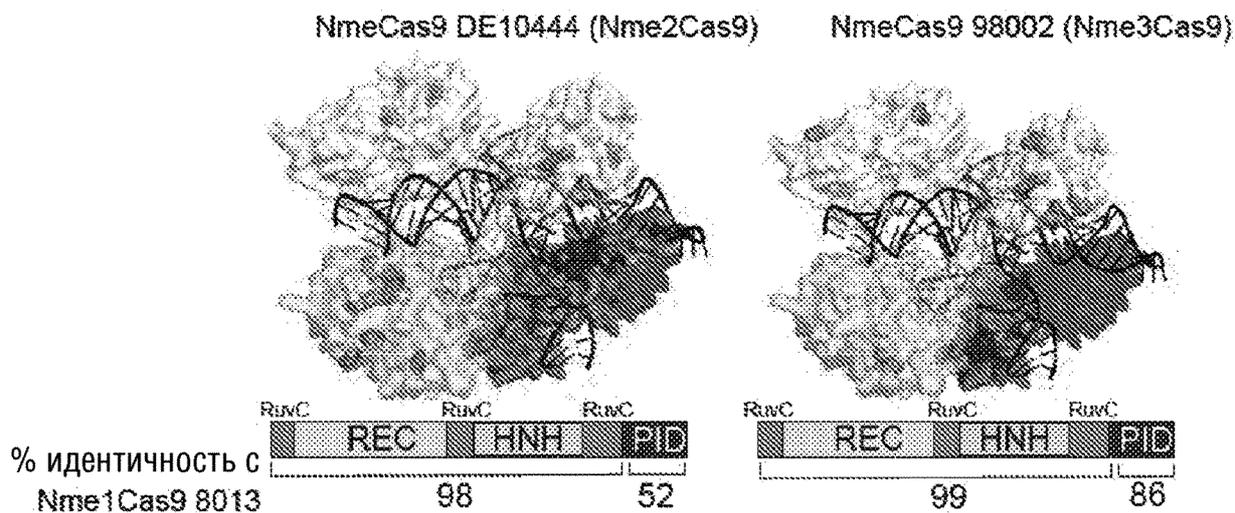


## ФИГ.4

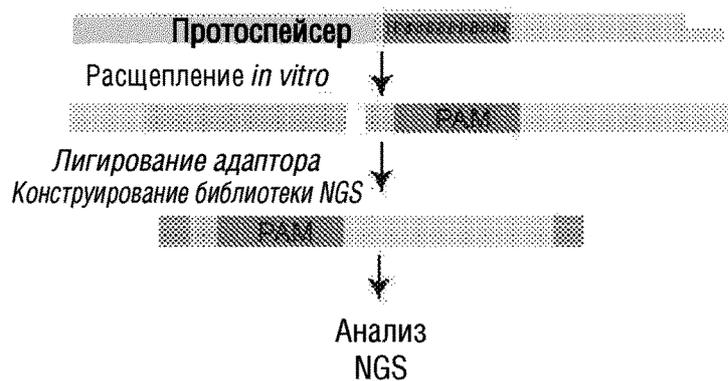


ФИГ.5

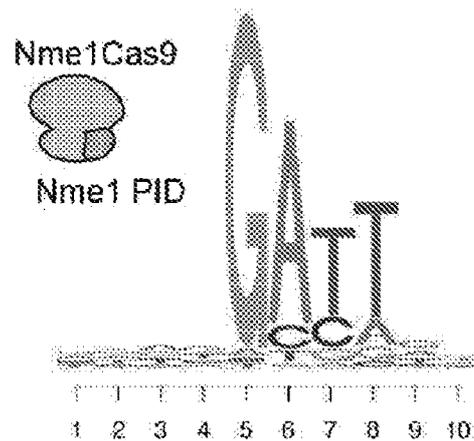
ФИГ.5А



ФИГ.5В

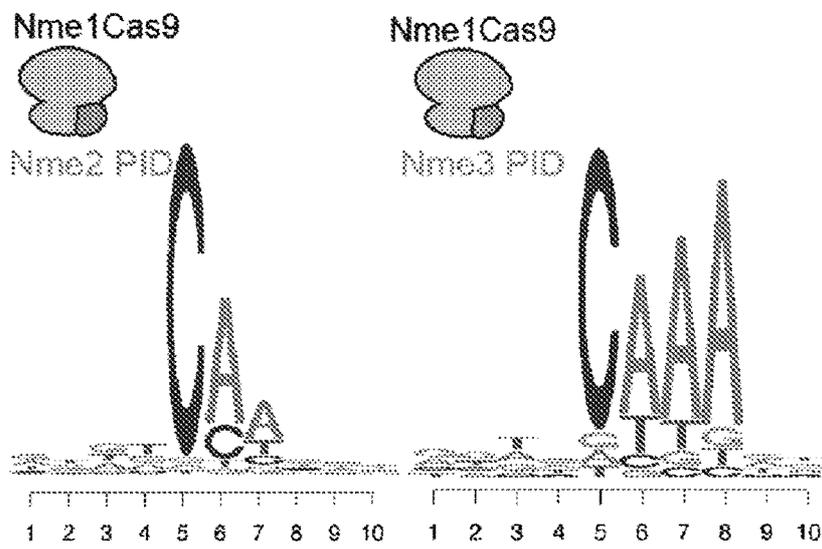


ФИГ.5С

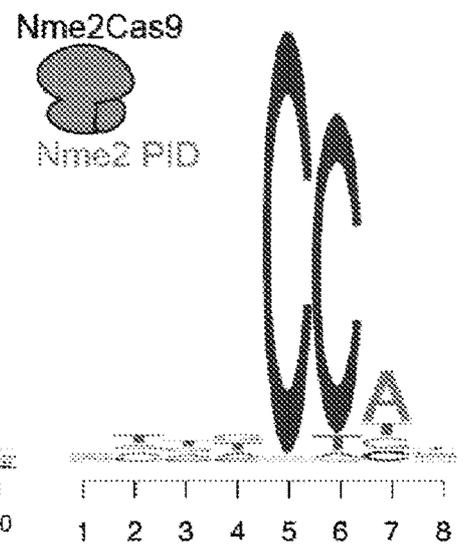


## ФИГ.5 (продолжение)

ФИГ.5D



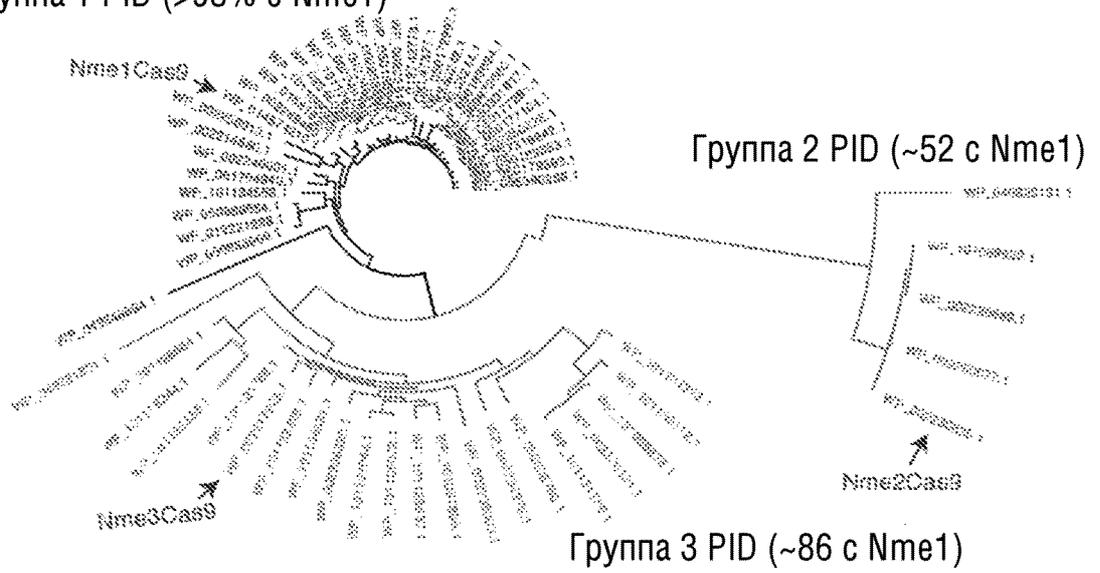
ФИГ.5E



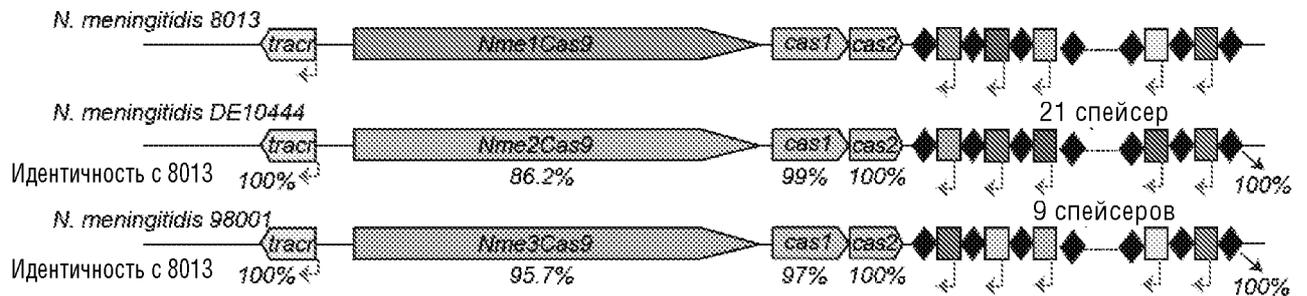
ФИГ.6

ФИГ.6А

Группа 1 PID (>98% с Nme1)

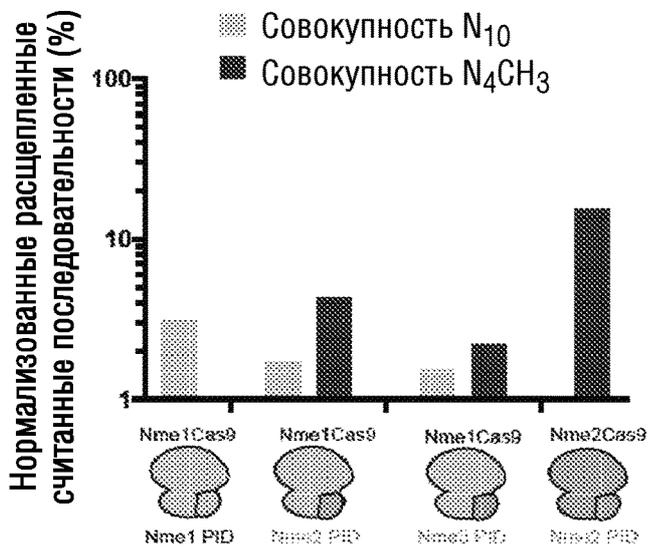


ФИГ.6В

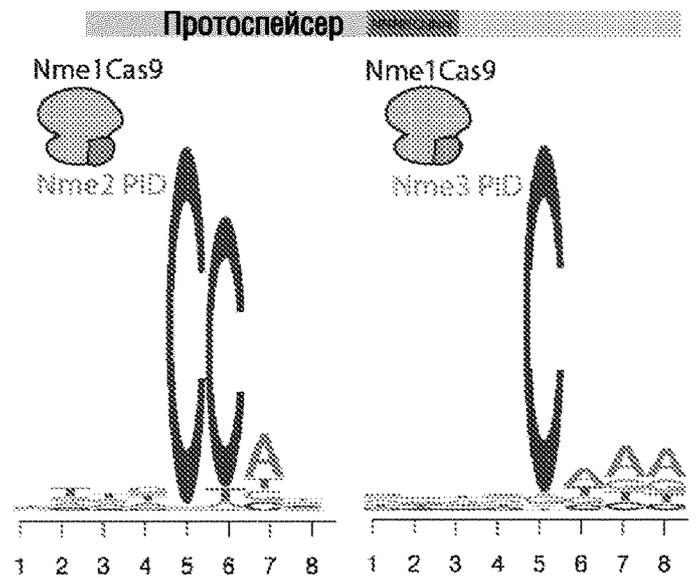


## ФИГ.6 (продолжение)

ФИГ.6С



ФИГ.6D

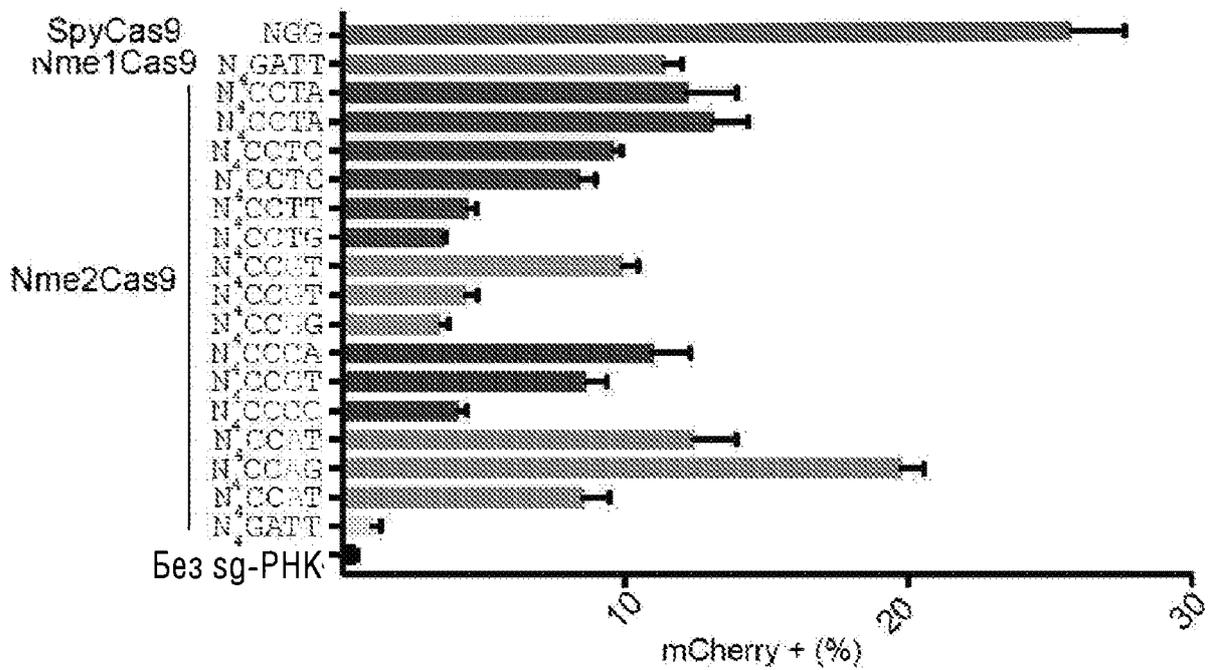


## ФИГ.7

## ФИГ.7А

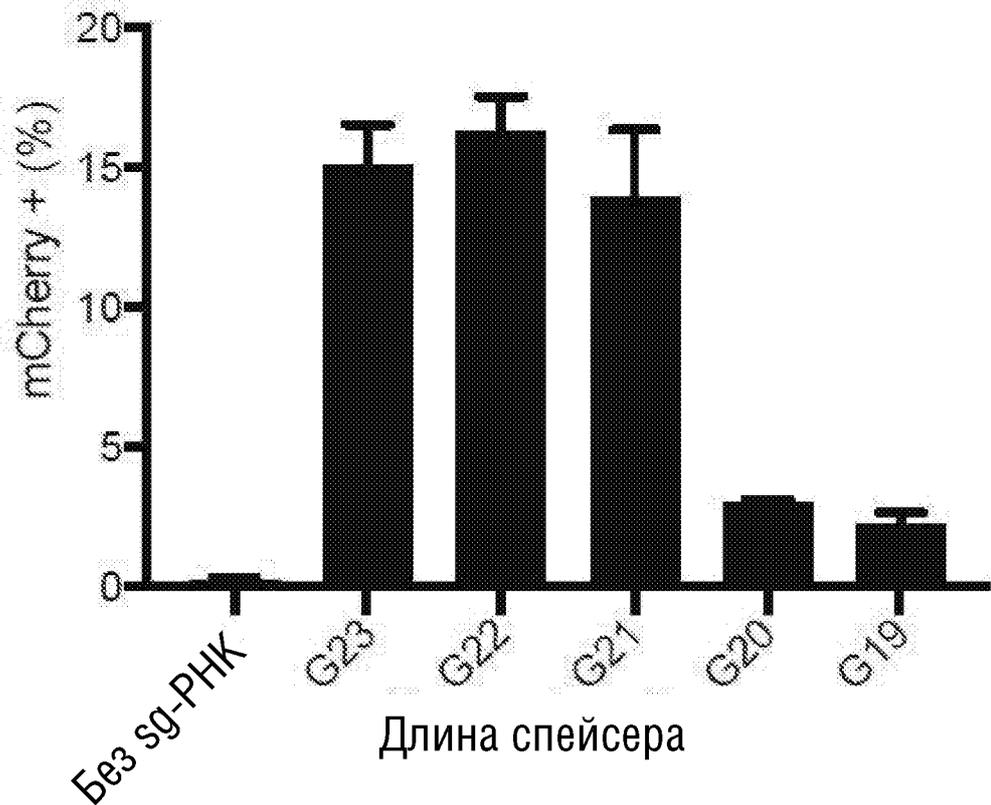


## ФИГ.7В



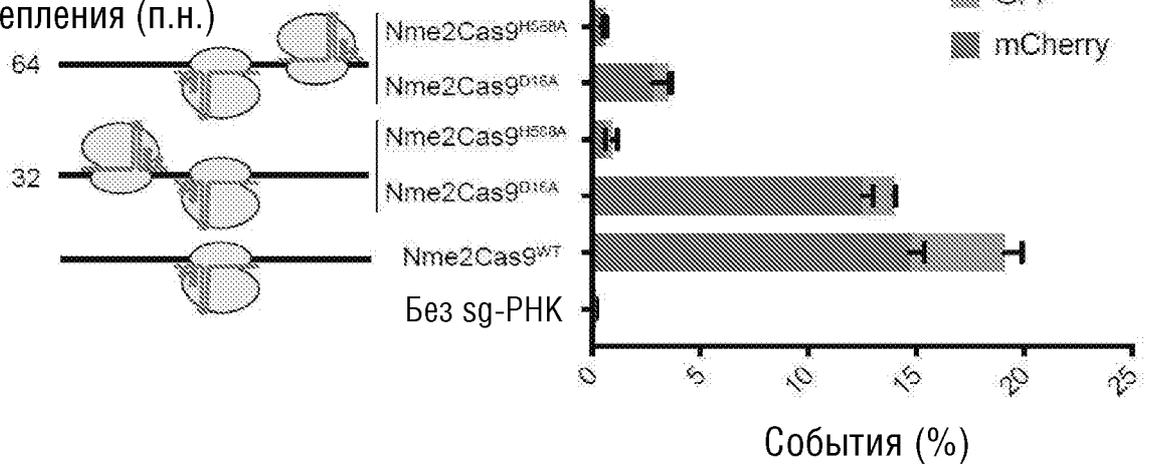
## ФИГ.7 (продолжение)

ФИГ.7С



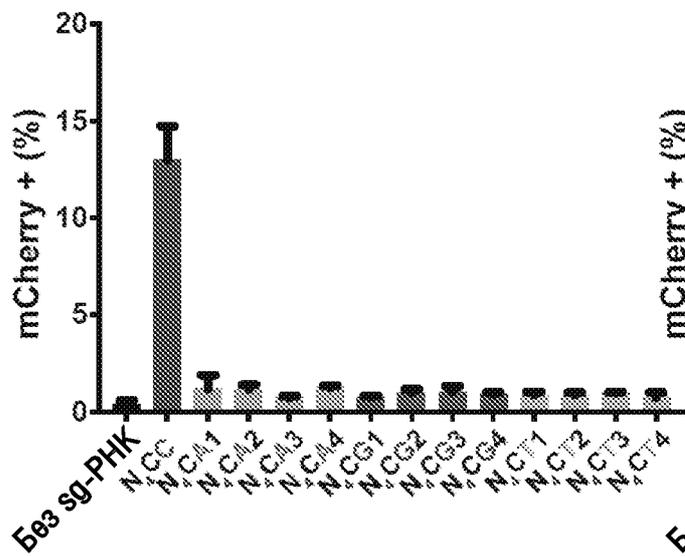
ФИГ.7D

Расстояние от участка  
расщепления (п.н.)

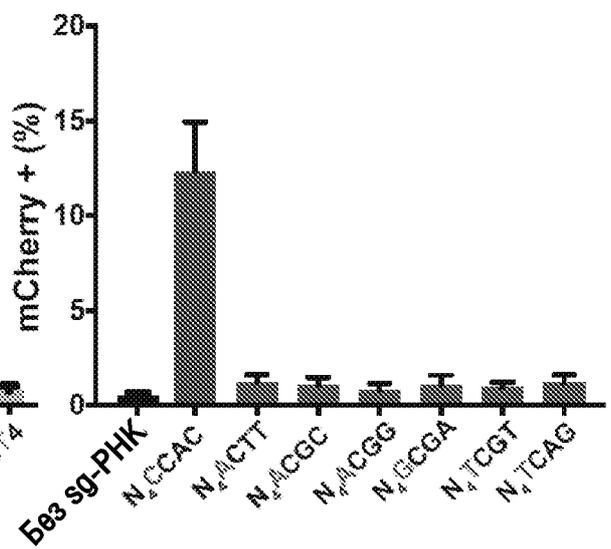


## ФИГ.8

ФИГ.8А



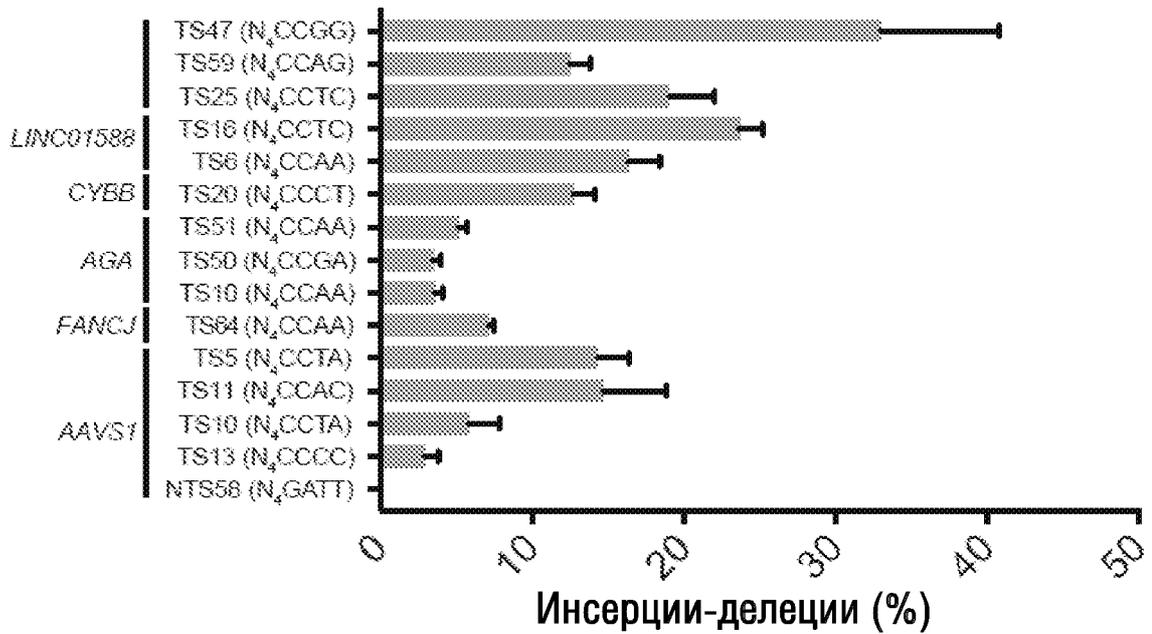
ФИГ.8В



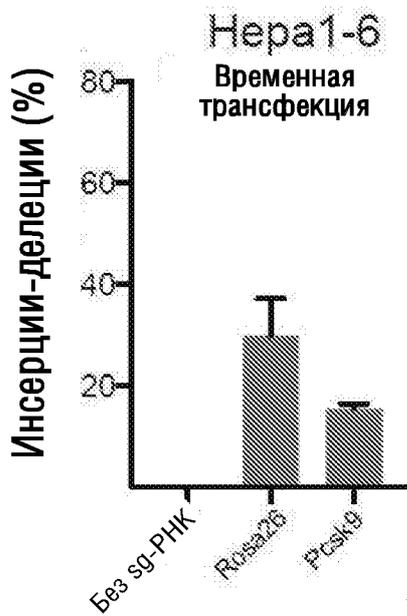


ФИГ.9

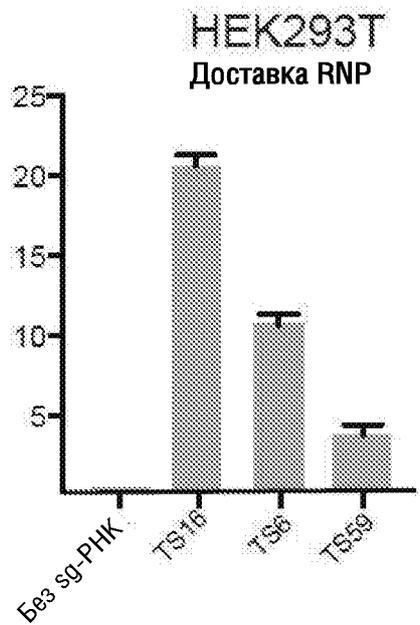
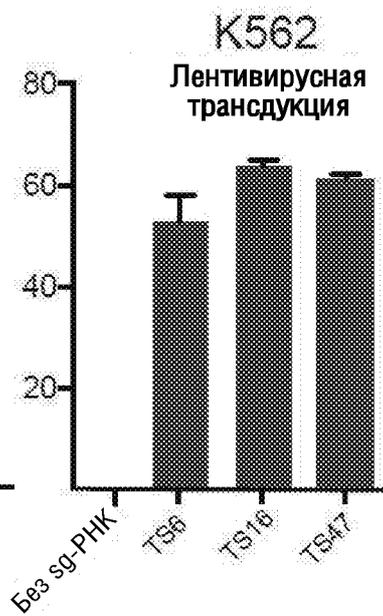
ФИГ.9А



ФИГ.9В

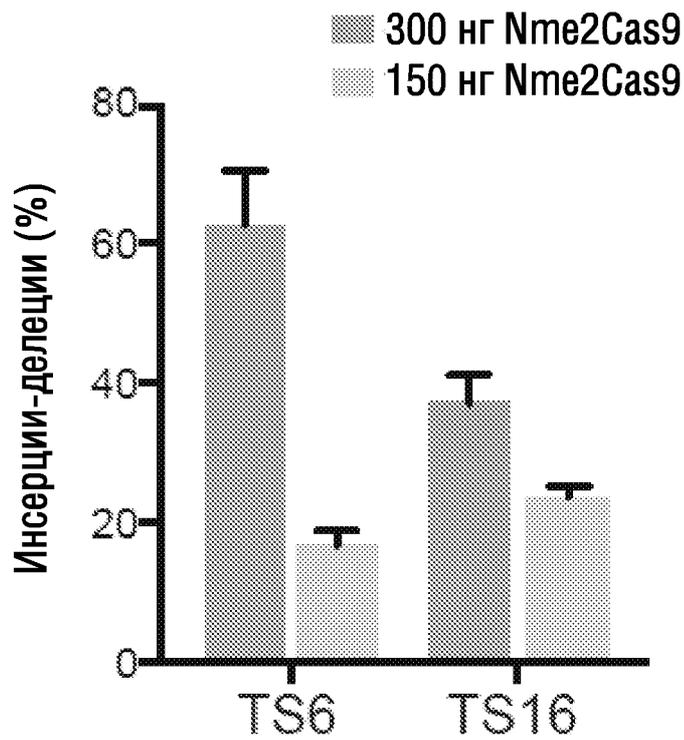


ФИГ.9С

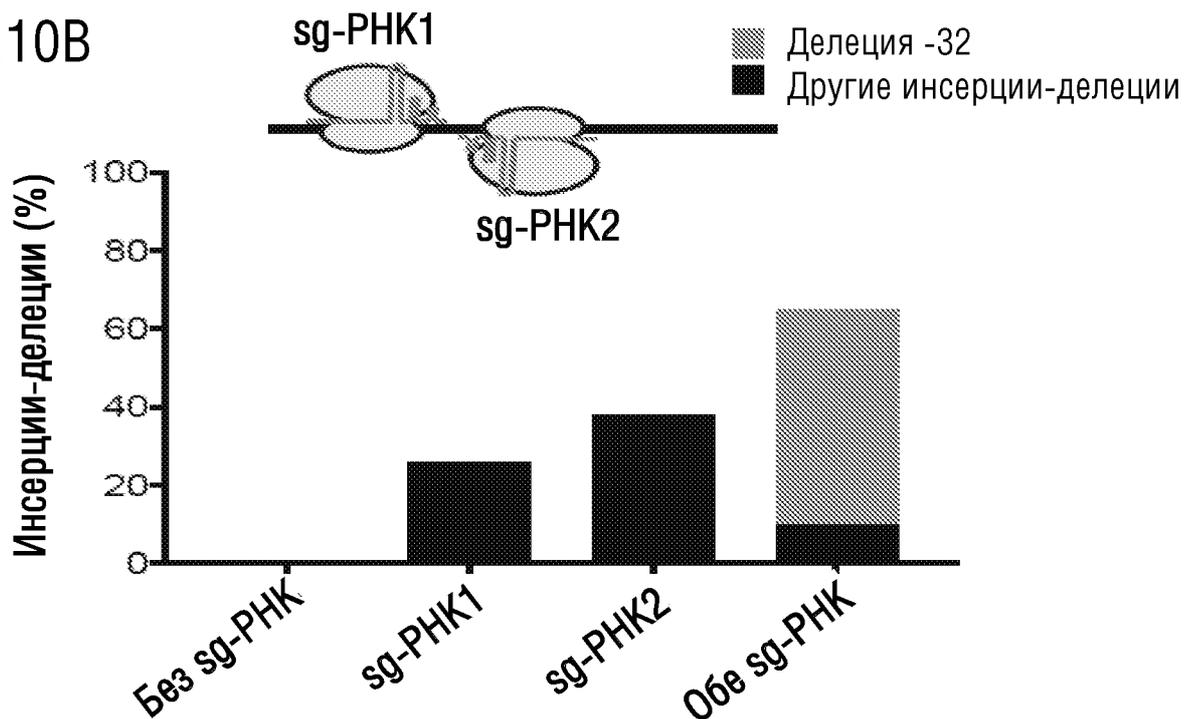


ФИГ.10

ФИГ.10А

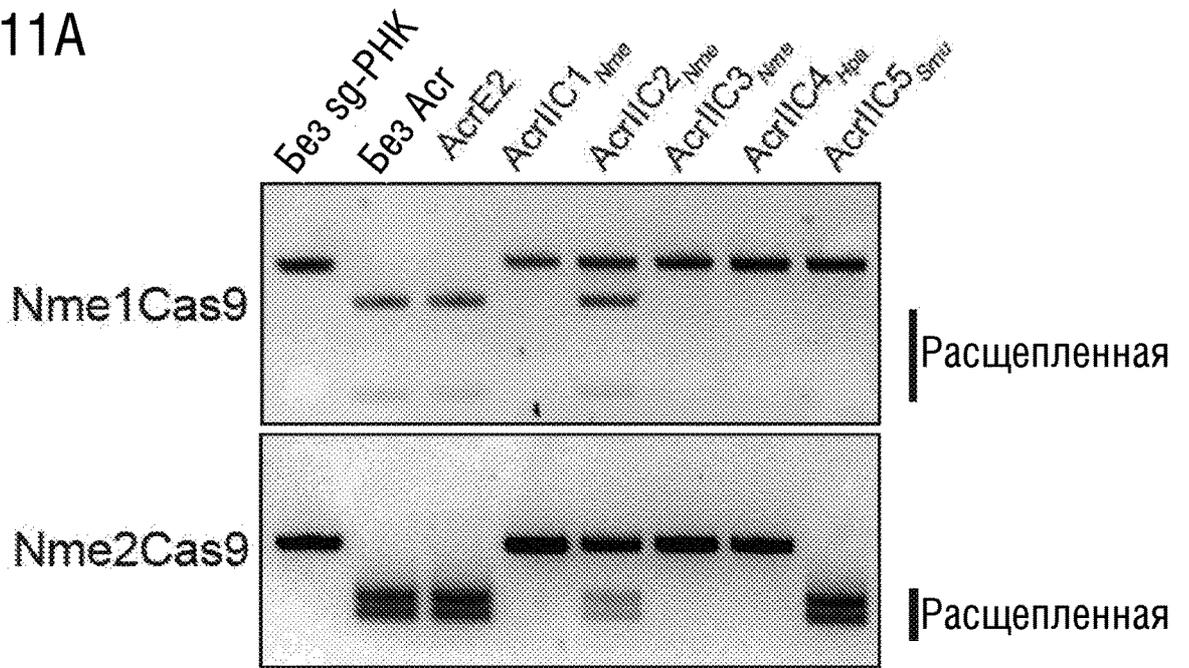


ФИГ.10В

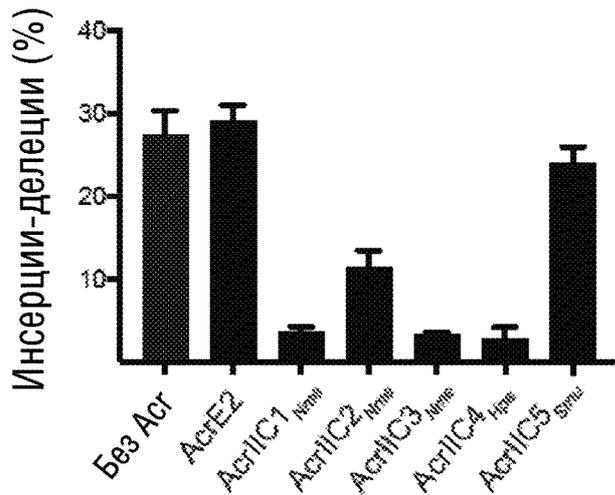


ФИГ.11

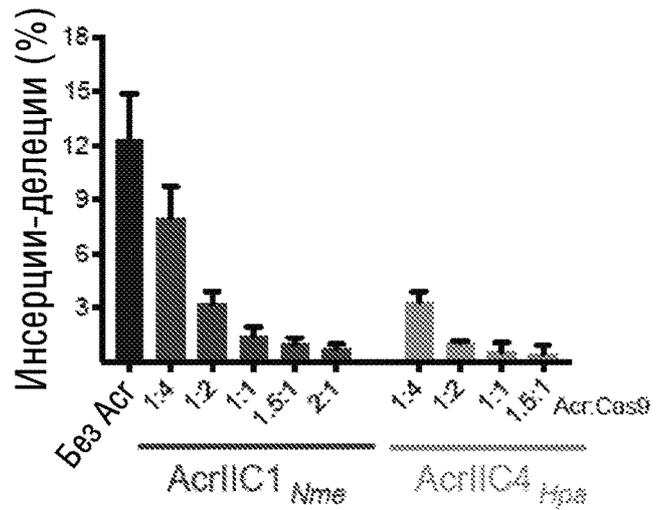
ФИГ.11А



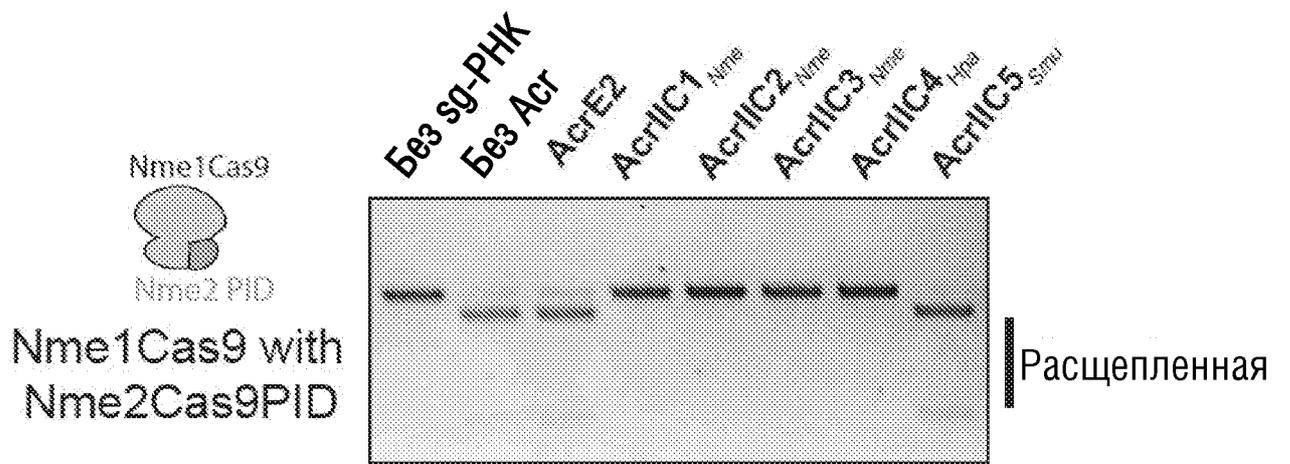
ФИГ.11В



ФИГ.11С

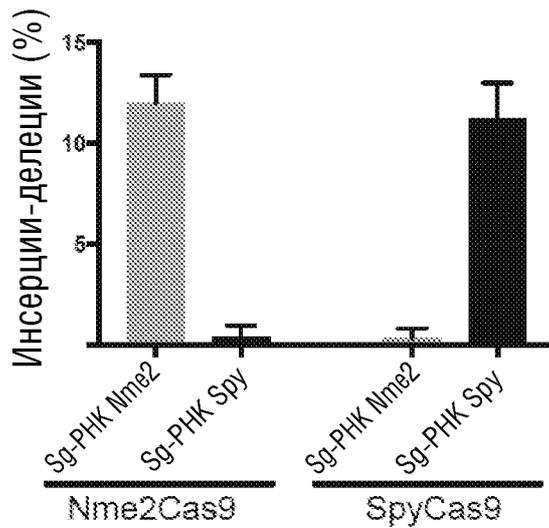


ФИГ.12

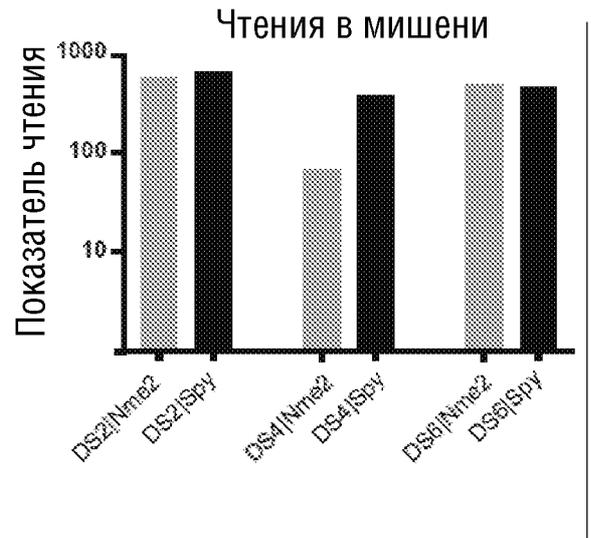


ФИГ.13

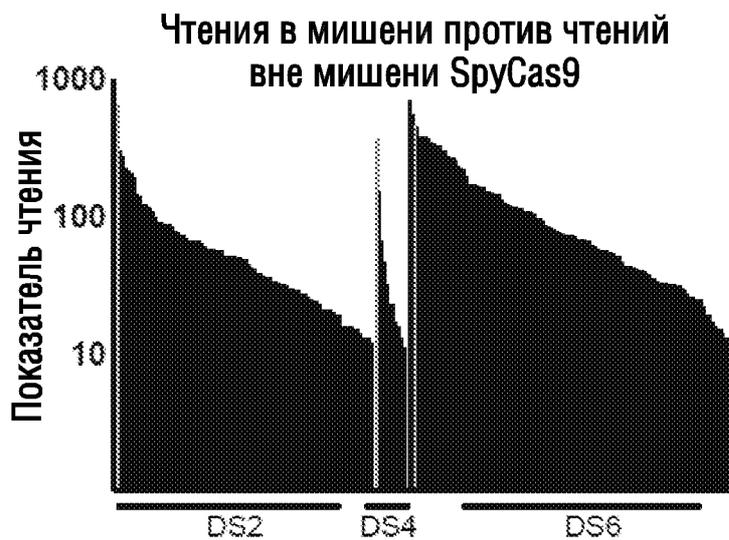
ФИГ.13А



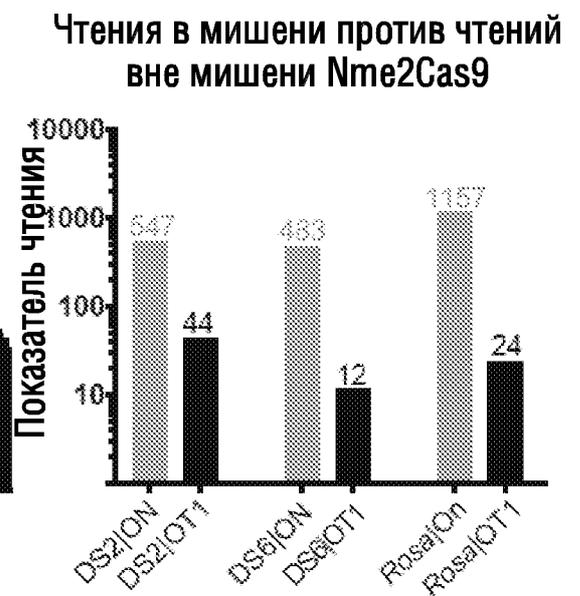
ФИГ.13В



ФИГ.13С

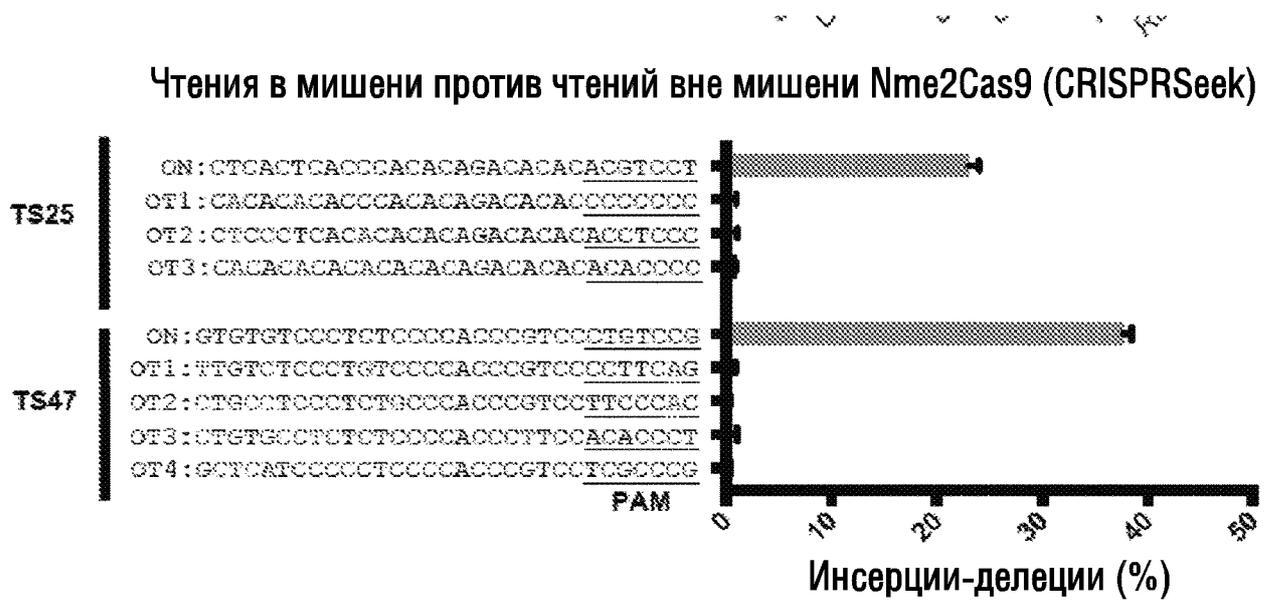


ФИГ.13D



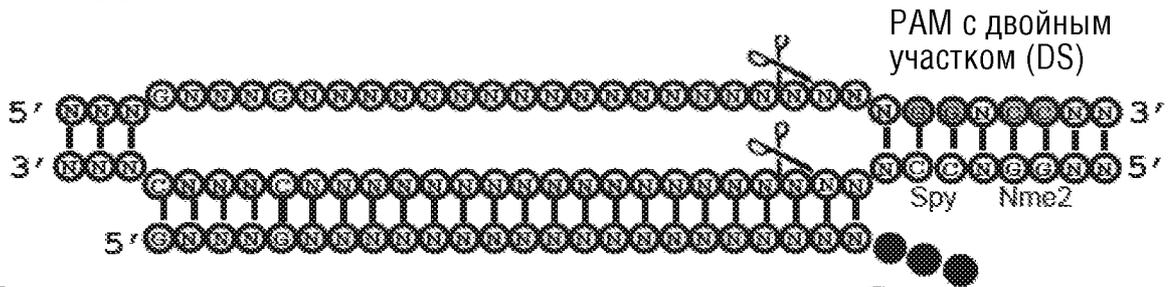
ФИГ.13 (продолжение)

ФИГ.13Е



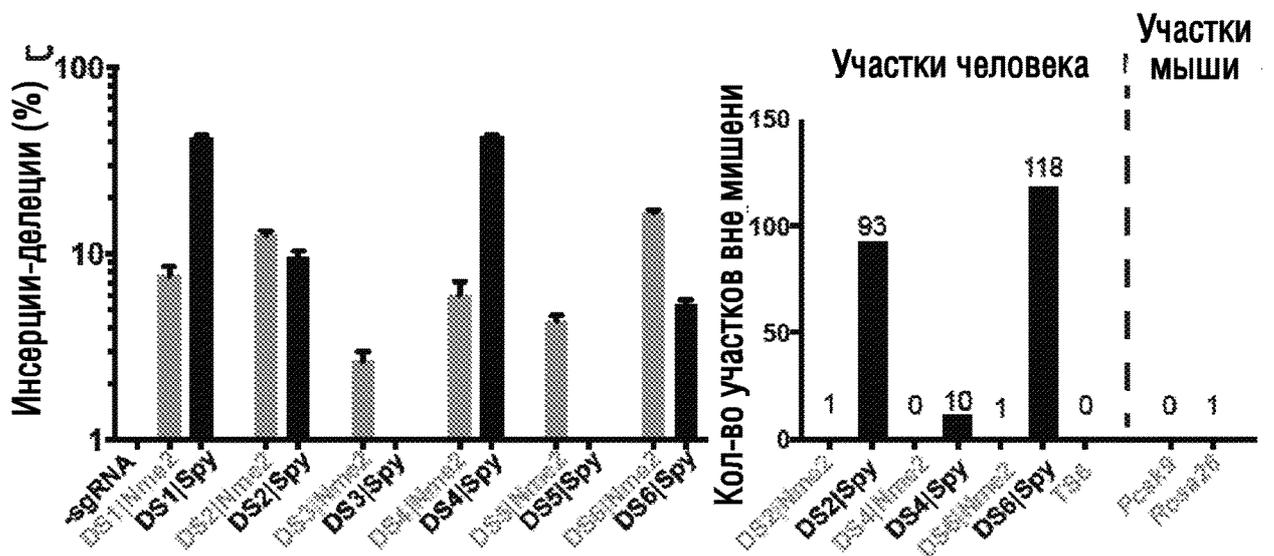
ФИГ.14

ФИГ.14А



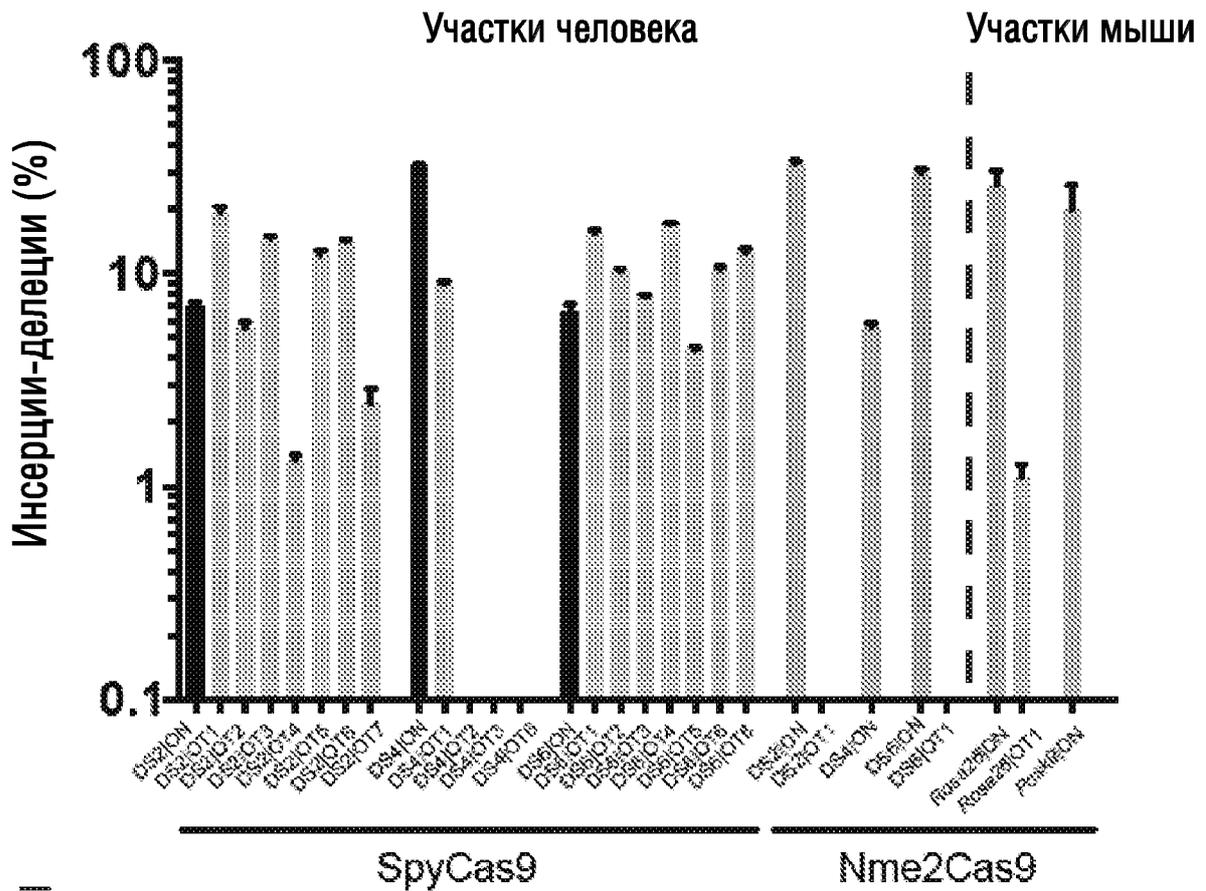
ФИГ.14В

ФИГ.14С



## ФИГ.14 (продолжение)

## ФИГ.14D



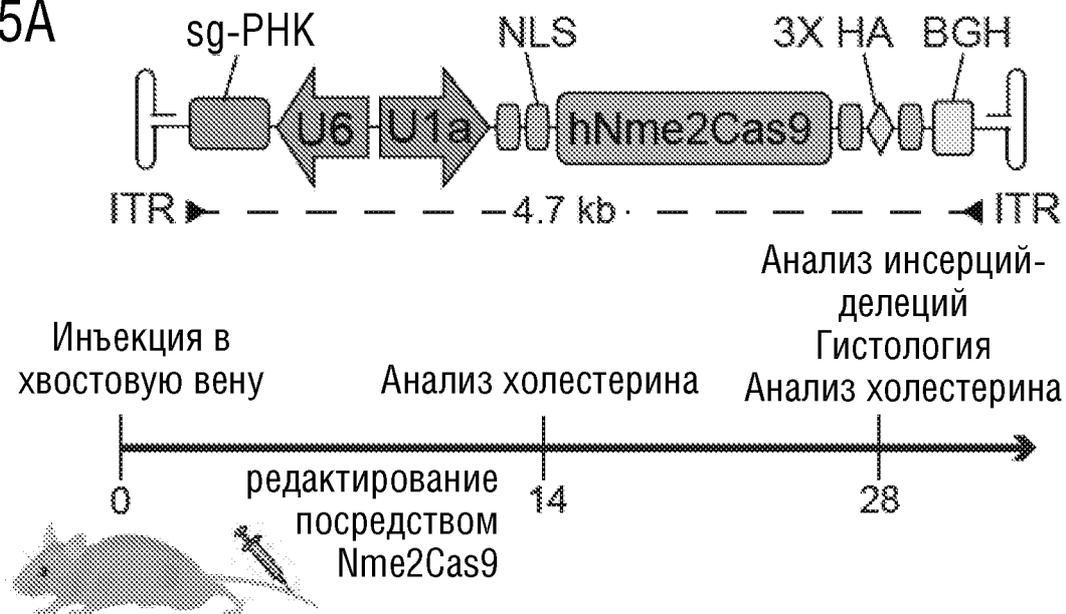
## ФИГ.14E

Rosa26|On: TGAGGACCGCCCTGGGCCTGGGAGAATCCCT

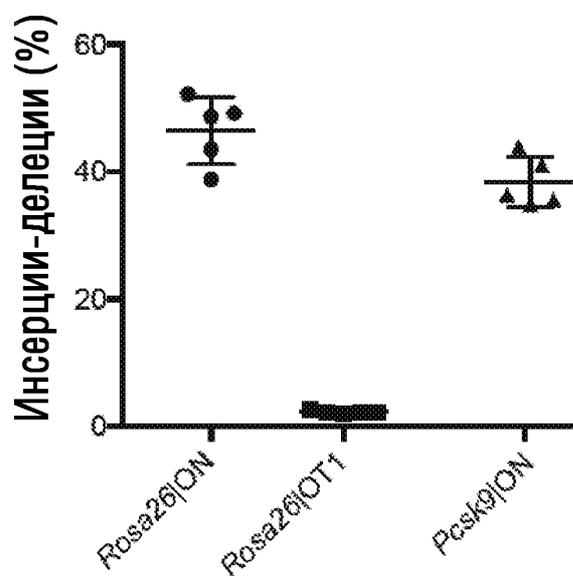
Rosa26|OT1: GAAGGACCAACCCTAGGCCTGGGAGACTCCCT

## ФИГ.15

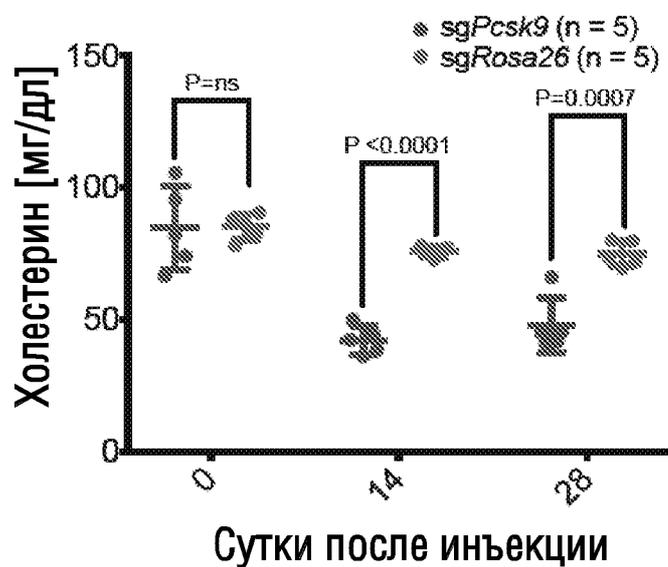
## ФИГ.15А



## ФИГ.15В

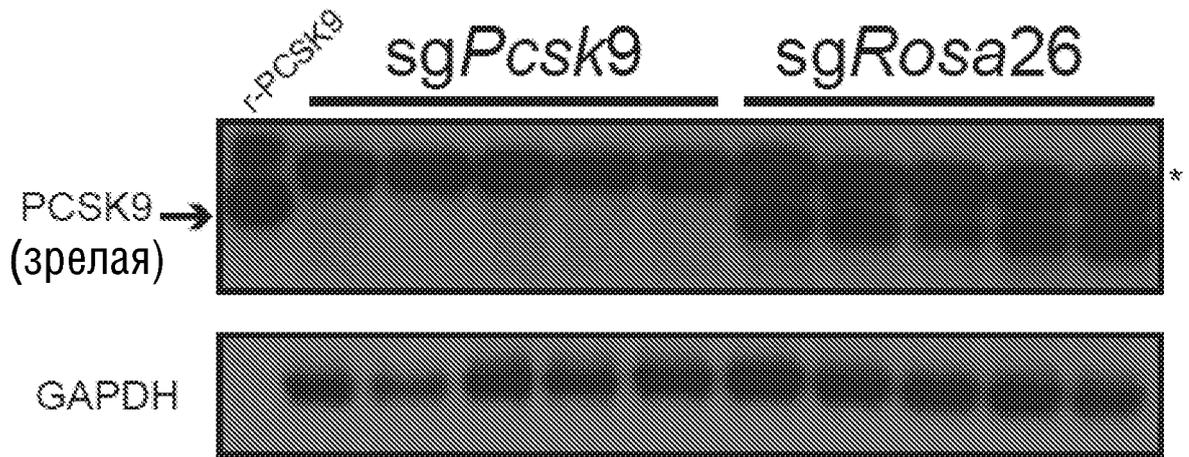


## ФИГ.15С

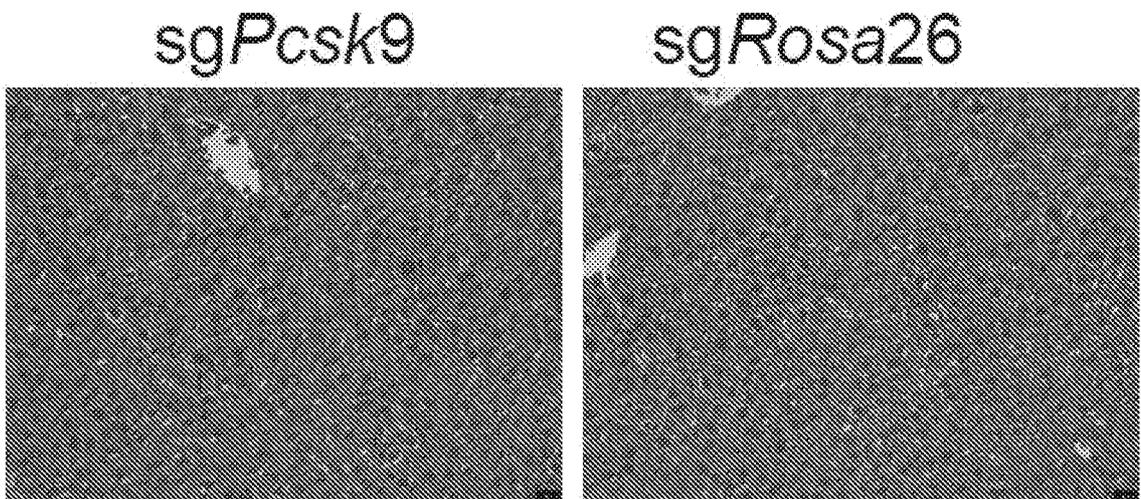


ФИГ.16

ФИГ.16А

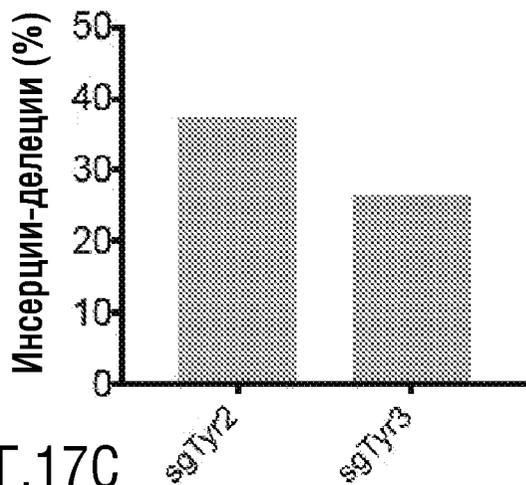


ФИГ.16В



ФИГ.17

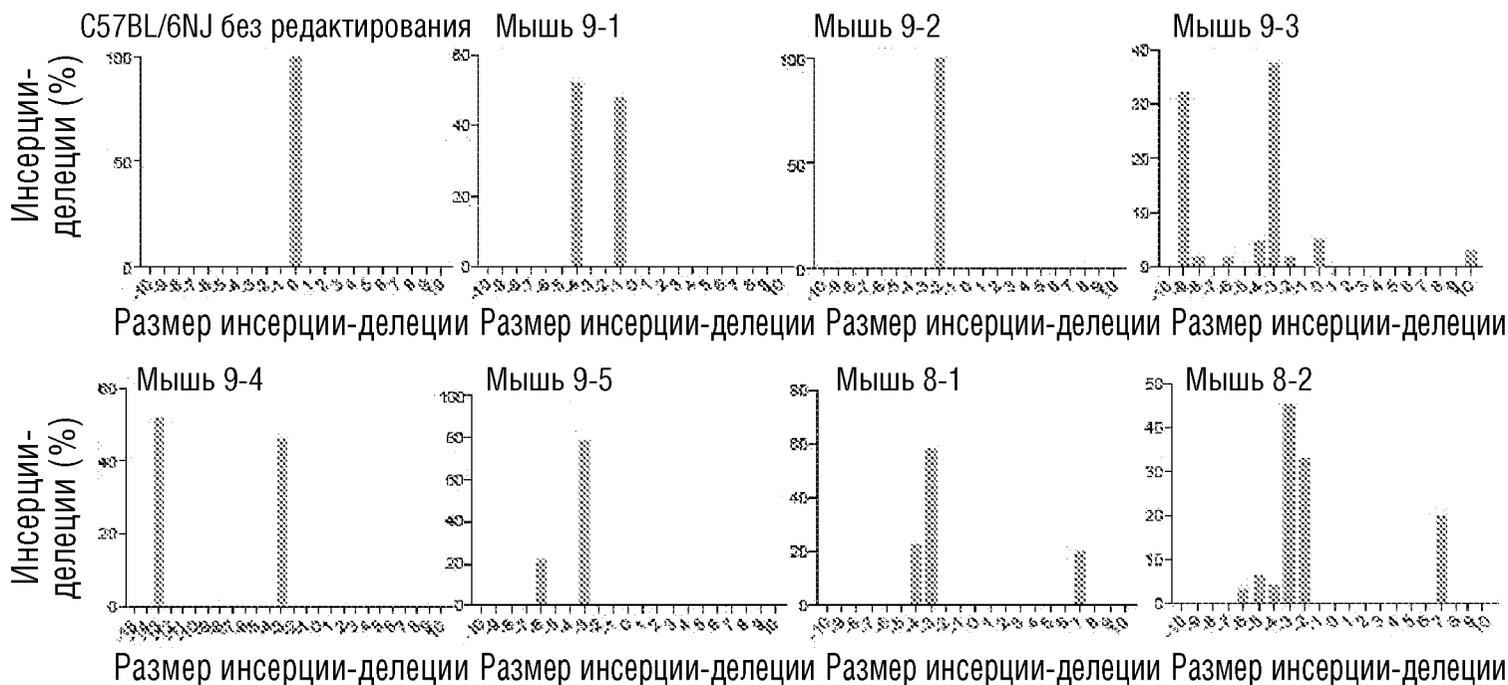
ФИГ.17А



ФИГ.17В

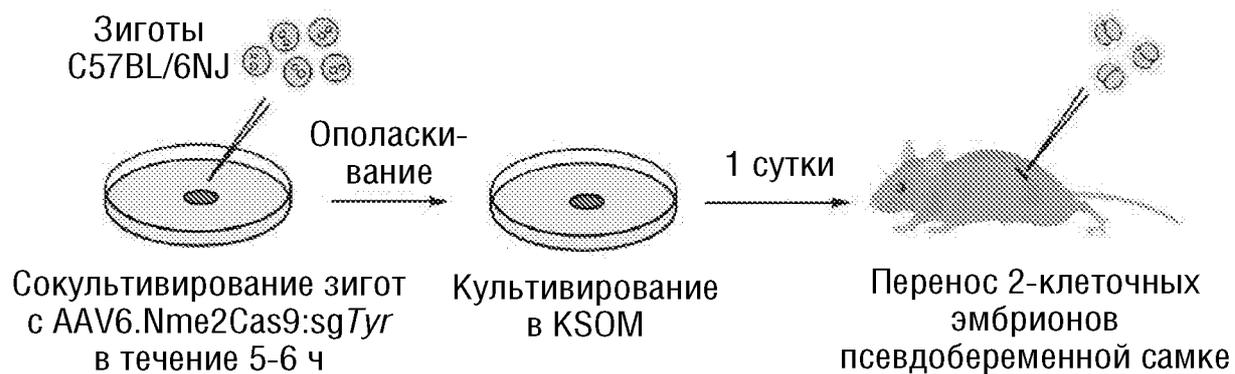
ID мыши	Дозировка AAV	Инсерции-делеции (%)	Цвет шерсти
9-1	$3 \times 10^9$ GC	100	Альбинос
9-2	$3 \times 10^9$ GC	100	Альбинос
9-3	$3 \times 10^9$ GC	84	мозаичные и гипоморфные
9-4	$3 \times 10^9$ GC	98	мозаичные и гипоморфные
9-5	$3 \times 10^9$ GC	100	мозаичные и гипоморфные
8-1	$3 \times 10^8$ GC	100	мозаичные и гипоморфные
8-2	$3 \times 10^8$ GC	96	мозаичные и гипоморфные

ФИГ.17С

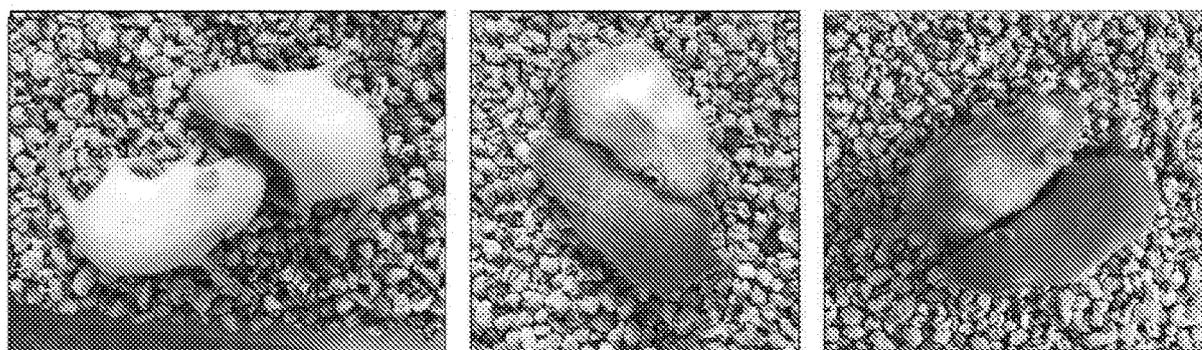


## ФИГ.18

## ФИГ.18А



## ФИГ.18В

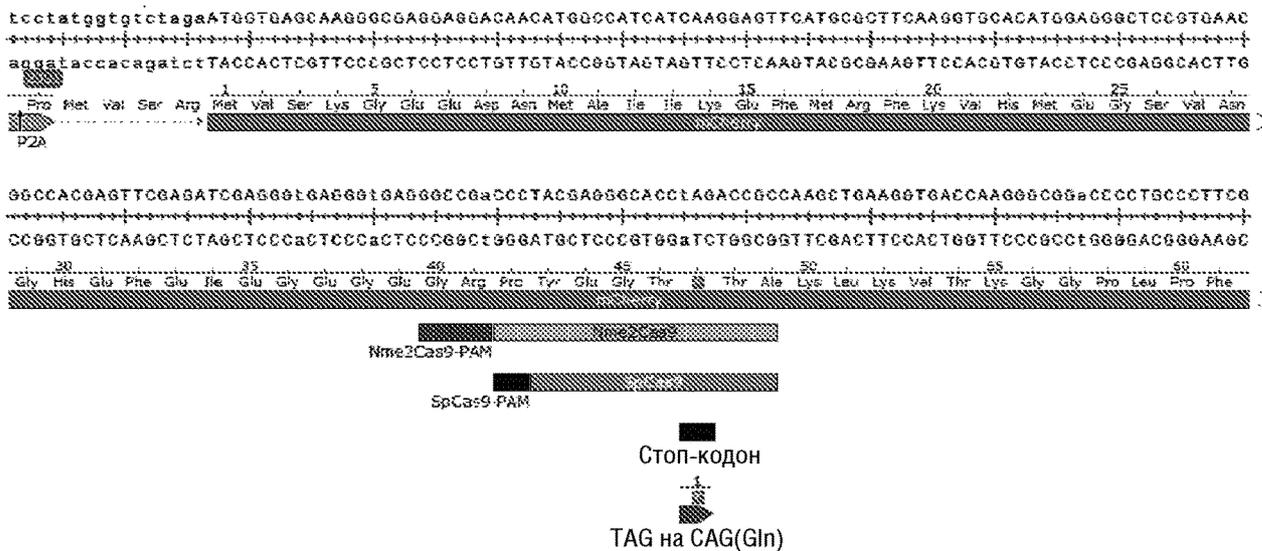
3 x 10<sup>9</sup> GC3 x 10<sup>8</sup> GC

## ФИГ.18С

Дозировка AAV	Количество перенесенных эмбрионов	Количество полученных детенышей	Количество детенышей с отредактированным Tug (альбиносы или шиншилловые)	Частота редактирования Tug (%)
3 x 10 <sup>9</sup> GC	50	5	5 (3)	100
3 x 10 <sup>8</sup> GC	28	2*	2 (1)	100

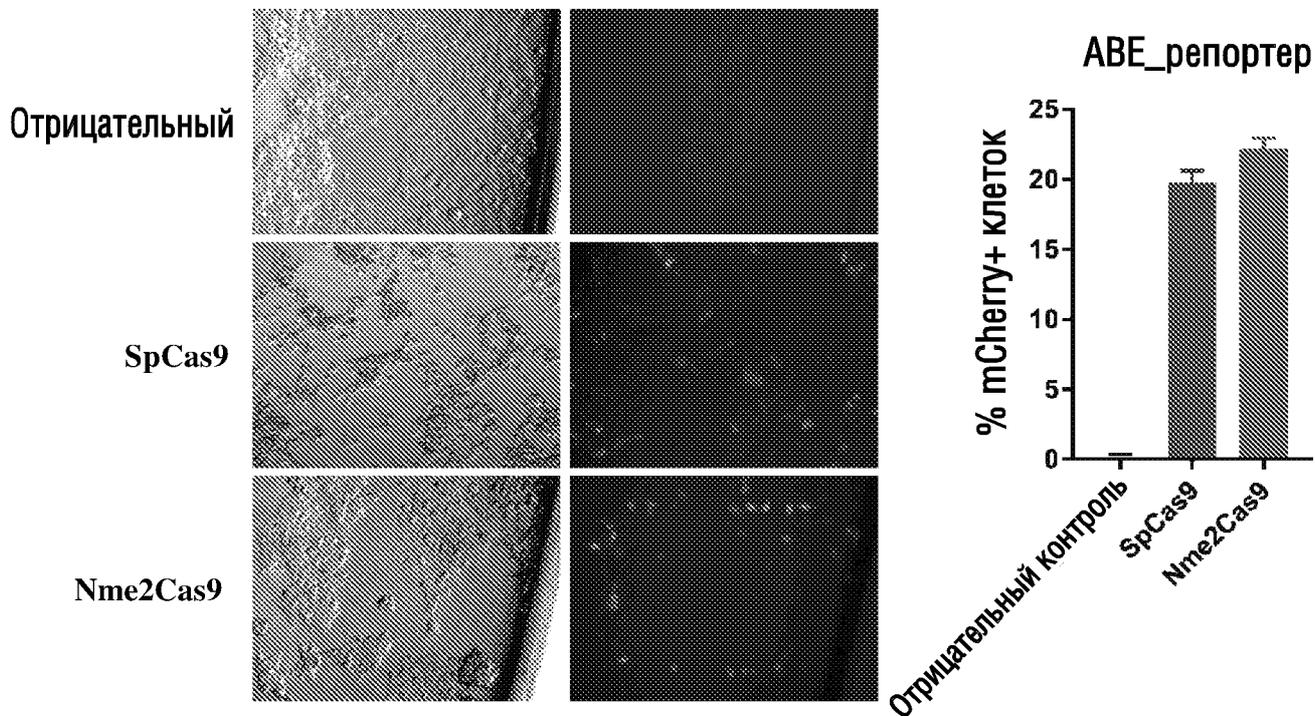
ФИГ.19

ФИГ.19А



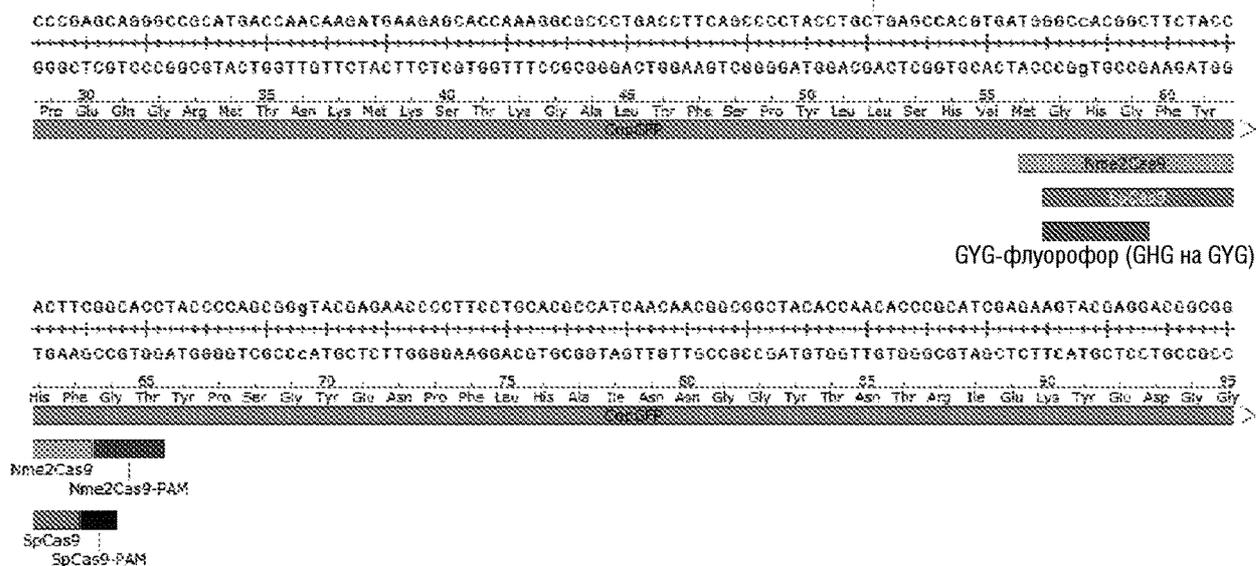
ФИГ.19В

ФИГ.19С



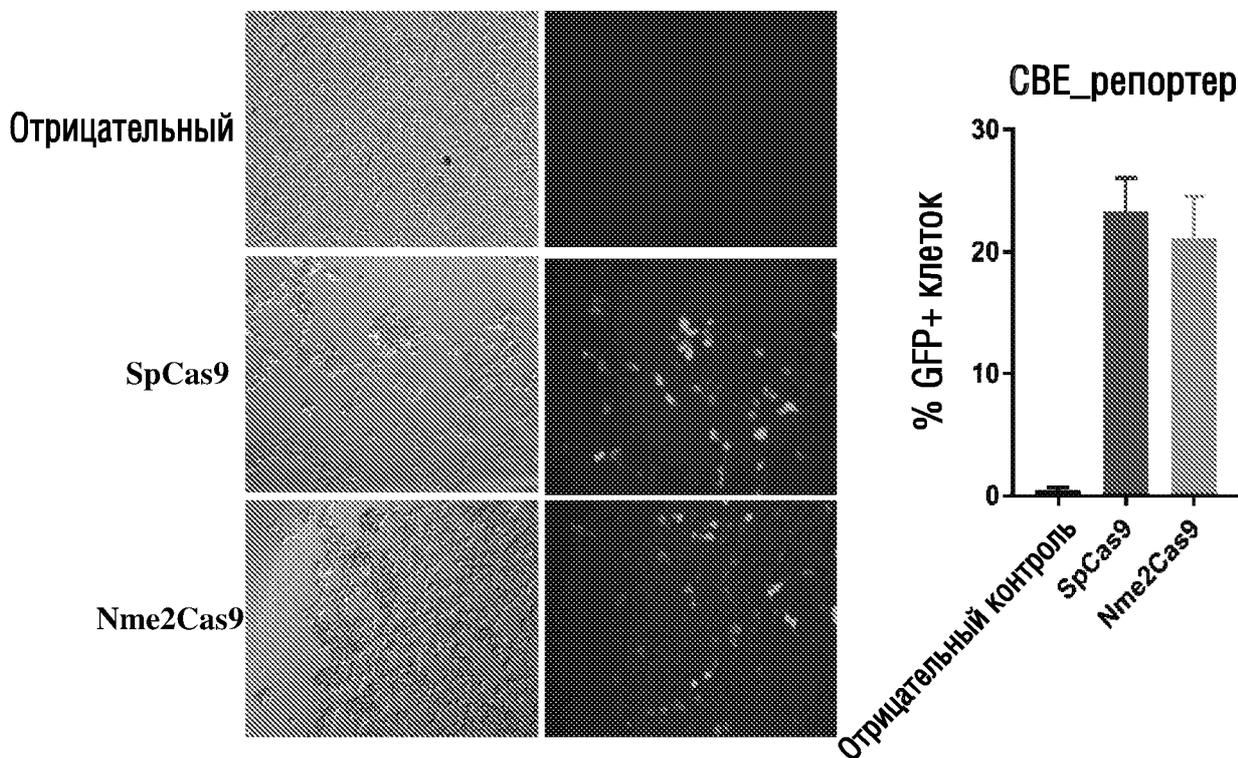
ФИГ.20

ФИГ.20А

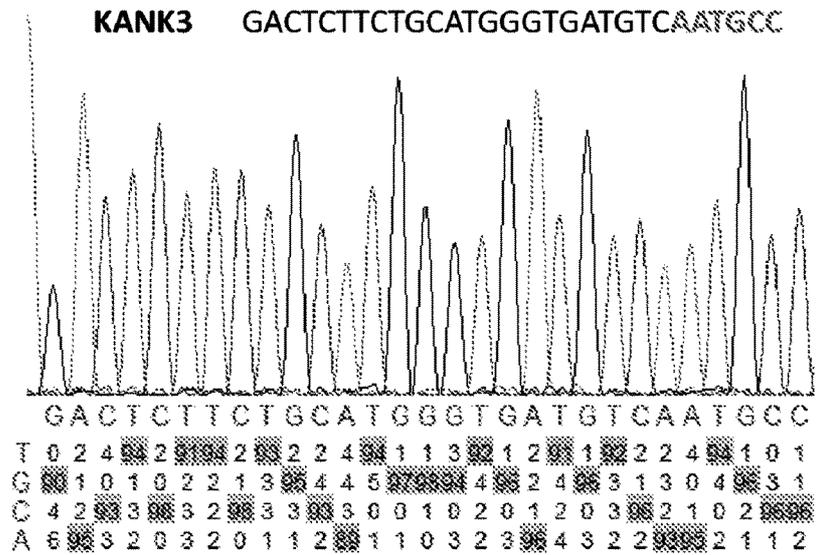


ФИГ.20В

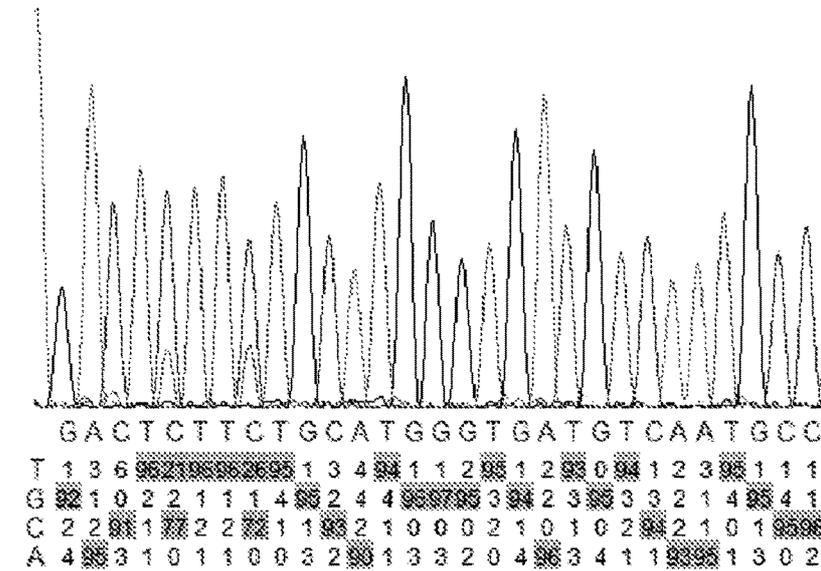
ФИГ.20С



ФИГ.21

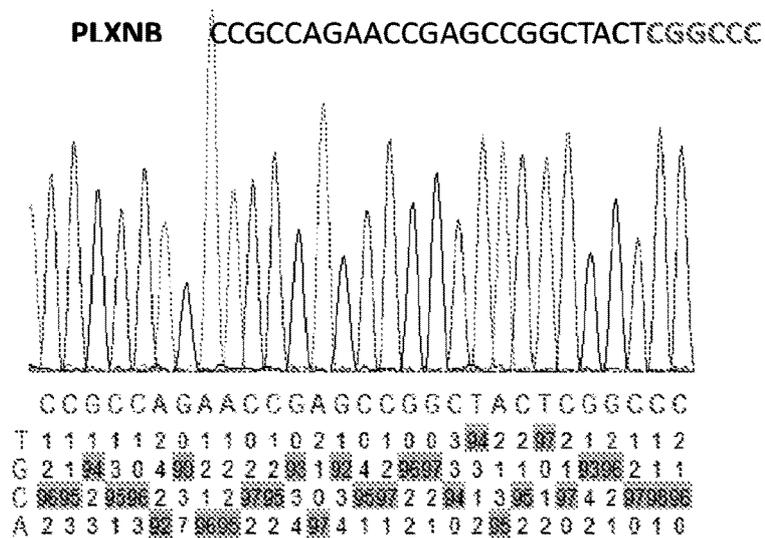


**Отрицательный**



**Nme2Cas9-CBE**

## ФИГ.22



## Отрицательный

