



(51) МПК
C12N 15/09 (2006.01)
C12N 15/90 (2006.01)
C12N 5/10 (2006.01)
A61K 48/00 (2006.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
C12N 15/09 (2021.08); *C12N 15/90* (2021.08); *C12N 5/10* (2021.08)

(21)(22) Заявка: 2019114217, 25.07.2014

(24) Дата начала отсчета срока действия патента:
25.07.2014

Дата регистрации:
21.01.2022

Приоритет(ы):

(30) Конвенционный приоритет:
26.07.2013 US 61/858,866

Номер и дата приоритета первоначальной заявки,
из которой данная заявка выделена:
2016106649 26.07.2013

(43) Дата публикации заявки: 23.08.2019 Бюл. № 24

(45) Опубликовано: 21.01.2022 Бюл. № 3

Адрес для переписки:
**101000, Москва, ул. Мясницкая, д. 13, стр. 5,
ООО "Союзпатент", С.Б. Фелициной**

(72) Автор(ы):

**ЧЁРЧ, Джордж М. (US),
ЯН, Лухан (US),
КАРГОЛ, Марк Гуэль (US),
ЯН, Джойс Личи (US)**

(73) Патентообладатель(и):

**ПРЕЗИДЕНТ ЭНД ФЭЛЛОУЗ ОФ
ХАРВАРД КОЛЛИДЖ (US)**

(56) Список документов, цитированных в отчете
о поиске: PABLO PEREZ-PINERA et al., RNA-
guided gene activation by CRISPR-Cas9-based
transcription factors, Nature Methods, 25.07.2013,
vol. 10, No. 10, pp. 973-976, DOI:10.1038/
NMETH.2600. US 2010076057 A1, 25.03.2010.
DICARLO J.E. et al., Genome engineering in
Saccharomyces cerevisiae using CRISP-Cas
systems, Nucleic Acids Research, 04.03.2013, 41(7),
(см. прод.)

C2

RU 2764757

(54) ГЕНОМНАЯ ИНЖЕНЕРИЯ

(57) Реферат:

Настоящее изобретение относится к биотехнологии и клеточной инженерии, в частности к способу изменения ДНК-мишени в стволовой клетке, экспрессирующей фермент Cas из системы CRISPR типа II, который формирует комплекс колокализации с направляющей РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом и к стволовой клетке для РНК-направленного основанного на Cas редактирования генома, которая не является клеткой зародышевой линии человека или эмбриональной клеткой человека. Для осуществления указанного способа сначала в указанную стволовую клетку вводят направляющую РНК, комплементарную ДНК-

мишени, причём РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени. Далее в ту же стволовую клетку вводят донорную последовательность нуклеиновой кислоты. Вследствие направляющая РНК и фермент Cas колокализуются с ДНК-мишенью, фермент Cas расщепляет ДНК-мишень и донорная последовательность нуклеиновой кислоты встраивается в ДНК-мишень. Таким образом получают измененную ДНК в указанной стволовой клетке. Настоящее изобретение позволяет применять специально сконструированные нуклеазы при конструировании человеческих индуцированных плюрипотентных стволовых клеток. 2 н. и 33 з.п. ф-лы, 50 ил., 9 табл., 20 пр.

R U 2 7 6 4 7 5 7 C 2

(56) (продолжение):
4336-4343, doi:10.1093/nar/gkt135. ПУГАЧ К.С. и др., CRISPR-системы адаптивного иммунитета прокариот, Молекулярная биология, 2012, 46(2), 195-203.

R U 2 7 6 4 7 5 7 C 2

R U 2 7 6 4 7 5 7 C 2

R U 2 7 6 4 7 5 7 C 2

RUSSIAN FEDERATION



(19) RU (11)

2 764 757⁽¹³⁾ C2

(51) Int. Cl.
C12N 15/09 (2006.01)
C12N 15/90 (2006.01)
C12N 5/10 (2006.01)
A61K 48/00 (2006.01)

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) ABSTRACT OF INVENTION

(52) CPC
C12N 15/09 (2021.08); *C12N 15/90* (2021.08); *C12N 5/10* (2021.08)

(21)(22) Application: 2019114217, 25.07.2014

(24) Effective date for property rights:
25.07.2014

Registration date:
21.01.2022

Priority:

(30) Convention priority:
26.07.2013 US 61/858,866

Number and date of priority of the initial application,
from which the given application is allocated:
2016106649 26.07.2013

(43) Application published: 23.08.2019 Bull. № 24

(45) Date of publication: 21.01.2022 Bull. № 3

Mail address:

101000, Moskva, ul. Myasnitskaya, d. 13, str. 5,
OOO "Soyuzpatent", S.B. Felitsinoj

(72) Inventor(s):

CHURCH, George M. (US),
YANG, Luhan (US),
CARGOL, Marc Guell (US),
YANG, Joyce Lichi (US)

(73) Proprietor(s):

PRESIDENT AND FELLOWS OF HARVARD
COLLEGE (US)

(54) GENOMIC ENGINEERING

(57) Abstract:

FIELD: biotechnology; cellular engineering.

SUBSTANCE: present invention relates to biotechnology and cellular engineering, in particular to a method for modifying a target DNA in a stem cell expressing Cas enzyme from a CRISPR type II system, which forms a colocalization complex with a guide RNA, complementary to the target DNA, and which cleaves the target DNA in a site-specific manner, and to a stem cell for RNA-directed Cas-based genome editing, which is not a human germ line cell or a human embryonic cell. To implement the specified method, the guide RNA, complementary to the target DNA, is first introduced into the specified stem cell, wherein

RNA and Cas enzyme are elements of the colocalization complex on the target DNA. Next, a donor sequence of nucleic acid is introduced into the same stem cell. As a result, the guide RNA and Cas enzyme colocalize with the target DNA, Cas enzyme cleaves the target DNA, and the donor sequence of nucleic acid is embedded in the target DNA. Thus, altered DNA is obtained in the specified stem cell.

EFFECT: present invention allows for the use of specially structured nucleases in the construction of human induced pluripotent stem cells.

35 cl, 50 dwg, 9 tbl, 20 ex

R U 2 7 6 4 7 5 7 C 2

Данные связанных заявок

Данная заявка испрашивает приоритет предварительной заявки на патент США №. 61/858866, поданной 26 июля 2013 г. и включенной в данный документ ссылкой в полном объеме и для всех целей.

5 Заявление государственных интересов

Это изобретение было осуществлено при государственной поддержке согласно P50 HG003170 от Национального научно-исследовательского центра генома человека за выдающиеся достижения в геномике. Правительство имеет определенные права на это изобретение.

10 Предшествующий уровень техники

Известно редактирование генома с помощью специфических к последовательности нуклеаз. Смотрите источники 1, 2 и 3, которые включены в данный документ ссылкой в полном объеме. Опосредованный нуклеазой разрыв двухцепочечной ДНК (дцДНК) в геноме может быть устранен с помощью двух основных механизмов: негомологичного

15 соединения концов (NHEJ), который часто приводит к введению неспецифических вставок и делеций (indels), или направляемой гомологией репарации (HDR), которая использует гомологичную цепь в качестве шаблона для репарации. См. источник 4 включенный ссылкой в полном объеме. Когда специфичная к последовательности

20 нуклеаза доставляется вместе с гомологичной донорной ДНК-конструкцией, содержащей искомые мутации, эффективности направленного воздействия на ген повышаются в 1000 раз по сравнению только с доставкой донорной конструкции отдельно. См. источник 5, включенный в данный документ ссылкой в полном объеме. Сообщалось об использовании одноцепочечных олигодезоксирибонуклеотидов («ssODNs») в качестве

25 ДНК-доноров. См. источники 21 и 22, включенные в данный документ ссылкой в полном объеме.

Несмотря на большие достижения в области технологий редактирования генов, остаются нерешенными многие проблемы и многие вопросы относительно применения специально сконструированных нуклеаз при конструировании человеческих индуцированных плюрипотентных стволовых клеток («hiPSC»). Во-первых, несмотря на простоту своей конструкции TALEN (эффекторные нуклеазы, подобные активаторам транскрипции) нацеливаются на определенные последовательности ДНК тандемными копиями доменов RVD (домены вариабельных двух остатков). См. источник 6, включенный в данный документ ссылкой в полном объеме. В то время как модульный характер RVDs упрощает дизайн TALEN, их повторяющиеся последовательности усложняют способы синтеза ДНК-конструкций на их основе (см. источники 2, 9 и 15-19, включенные в данный документ ссылкой в полном объеме), а также делает практически невозможным их совместное применение вместе со средствами доставки генов на основе лентивирусов. См. источник 13, включенный в данный документ ссылкой в полном объеме.

40 В современной практике NHEJ и HDR часто оцениваются с помощью отдельных анализов. Тесты с эндонуклеазой, чувствительной к некомлементарности (см. источник 14, включенный в данный документ ссылкой в полном объеме) часто используются для оценки NHEJ, но количественная точность этого способа переменчива, и чувствительность ограничивается частотами NHEJ выше ~ 3%. См. источник 15,

45 включенный в данный документ ссылкой в полном объеме. HDR часто оценивается с помощью клонирования и секвенирования, совершенно иной и часто громоздкой процедуры. Чувствительность по-прежнему является проблемой, поскольку, несмотря на высокие частоты редактирования порядка 50% о которых часто сообщалось в случае

некоторых типов клеток, таких как U20S и K562 (см. источники 12 и 14, включенные в данный документ ссылкой в полном объеме), в случае hiPSC частоты, как правило, ниже. См. Источник 10, включенный в данный документ ссылкой в полном объеме. В последнее время высокие частоты редактирования были зарегистрированы в hiPSC и hESC, достигнутые с помощью TALENs (см. источник 9, включенный в данный документ ссылкой в полном объеме), и еще более высокие частоты были достигнуты с помощью системы CRISPR Cas9-гидРНК (см. источники 16-19, включенные в данный документ ссылкой в полном объеме. Тем не менее, показатели редактирования на различных участках по всей видимости, сильно отличаются (см. источник 17, включенный в данный документ ссылкой в полном объеме), а в некоторых участках редактирование иногда не обнаруживается вообще (см. источник 20, включенный в данный документ ссылкой в полном объеме).

Бактериальные и архейные системы CRISPR-Cas основываются на коротких направляющих РНК (гидРНК) в комплексе с Cas-белками, которые управляют деградацией комплементарных последовательностей, присутствующих во вторгающихся чужеродных нуклеиновых кислотах. См. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011); Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2579-2586 (2012); Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012); Sapranauskas, R. et al. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic acids research* 39, 9275-9282 (2011); and Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual review of genetics* 45, 273-297 (2011).

Недавнее воссоздание *in vitro* системы CRISPR II типа из *S. pyogenes* показало, что crRNA («CRISPR РНК»), слитая, как правило, с транс-кодируемой tracrRNA («транс-активированная CRISPR РНК») достаточны для направления белка Cas9 к специфически расщепляемой последовательности ДНК-мишени, соответствующей crRNA. Экспрессия гидРНК гомологичных сайту-мишени приводит к рекрутированию Cas9 и деградации ДНК-мишени. См. H. Deveau et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190, 1390 (Feb, 2008).

Сущность изобретения

Аспекты настоящего описания относятся к применению модифицированных нуклеаз

TALEN (эффекторные нуклеазы, подобные активатору транскрипции) для генетической модификации клетки, такой как соматическая клетка или стволовая клетка. Как известно, TALEN включают повторяющиеся последовательности. Аспекты настоящего раскрытия относятся к способу изменения ДНК-мишени в клетке, в том числе введением в клетку TALEN, утратившую последовательности повторов в 100 п.о. или длиннее, где TALEN расщепляет ДНК-мишени и клетка подвергается негомологичному соединению концов с получением измененной ДНК в клетке. В соответствии с некоторыми аспектами, повторяющиеся последовательности искомой длины были удалены из TALEN. В соответствии с некоторыми аспектами, TALEN была лишена повторяющихся последовательностей определенной искомой длины. В соответствии с некоторыми аспектами, TALEN предоставляется с удаленными повторяющимися последовательностями искомой длины. В соответствии с некоторыми аспектами, TALEN модифицируется, для удаления повторяющихся последовательностей искомой длины. В соответствии с некоторыми аспектами, TALEN сконструирована с удалением

повторяющихся последовательностей искомой длины.

Аспекты настоящего изобретения включают способы изменения ДНК-мишени в клетке, в том числе объединением в клетке TALEN, утратившей повторяющиеся последовательности в 100 п.о. или длиннее, и последовательности донорной нуклеиновой

5 кислоты, где TALEN расщепляет ДНК-мишень и донорная последовательность нуклеиновой кислоты встраивается в ДНК клетки. Аспекты настоящего раскрытия направлены на вирус, включающий последовательность нуклеиновой кислоты, кодирующую TALEN, утратившую повторяющиеся последовательности в 100 п.о. или длиннее. Аспекты настоящего раскрытия относятся к клетке, включающей

10 последовательность нуклеиновой кислоты, кодирующую TALEN, утратившую повторяющиеся последовательности в 100 п. о. или длиннее. В соответствии с некоторыми аспектами, описанными в данном документе, TALEN утратила повторяющиеся последовательности в 100 п.о. или длиннее, 90 п.о. или длиннее, 80 п.о. или длиннее, 70 п.о. или длиннее, 60 п.о. или длиннее, 50 п.о. или длиннее, 40 п.о. или 15 длиннее, 30 п.о. или длиннее, 20 п.о. или длиннее, 19 п.о. или длиннее, 18 п.о. или длиннее, 17 п.о. или длиннее, 16 п.о. или длиннее, 15 п.о. или длиннее, 14 п.о. или длиннее, 13 п.о. или длиннее, 12 п.о. или длиннее, на 11 п.о. или длиннее, или 10 пар или длиннее.

Аспекты настоящего раскрытия направлены на изготовление TALE, включающего объединение эндонуклеазы, ДНК-полимеразы, ДНК-лигазы, экзонуклеазы, множество 20 блоков нуклеотидных димеров, кодирующих RVD-домены (домены вариабельных двух остатков) и каркасного вектора TALE-N/TF, включающего сайт расщепления эндонуклеазой, активацию эндонуклеазы для разрезания каркасного вектора TALE-N/TF по сайту разрезания эндонуклеазой, с получением первого конца и второго конца, активацию экзонуклеазы, для создания 3' и 5' выступов на каркасном векторе TALE-N/TF и множество блоков нуклеотидных димеров и соединения каркасного вектора TALE-N/TF и множества блоков нуклеотидных димеров в искомом порядке, активации ДНК-полимеразы и ДНК-лигазы для соединения каркасного вектора TALE-N/TF и множества блоков нуклеотидных димеров. Специалист в данной области техники легко поймет на основании настоящего описания, как идентифицировать подходящие эндонуклеазы, 25 ДНК-полимеразы, ДНК-лигазы, экзонуклеазы, блоки нуклеотидных димеров, кодирующих RVD-домены и каркасные векторы TALE-N/TF.

Аспекты настоящего описания относятся к способу изменения ДНК-мишени в стволовых клетках, экспрессирующих фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сай́т-специфическим образом, включающий (а) введение в стволовую клетку первой чужеродной нуклеиновой кислоты, кодирующей РНК, комплементарную ДНК-мишени, который направляет фермент на ДНК-мишень, где РНК и фермент являются членами комплекса колокализации на ДНК-мишени, введение в стволовую клетку второй чужеродной нуклеиновой кислоты, кодирующей последовательность донорной нуклеиновой кислоты, где РНК и донорная нуклеотидная последовательность экспрессируются, где РНК и фермент колокализуются с ДНК-мишенью, фермент расщепляет ДНК-мишени и донорная нуклеиновая кислота встраивается в ДНК-мишени, с получением измененной ДНК в стволовой клетке.

40 Аспекты настоящего раскрытия направлены на стволовую клетку, включающую первую чужеродную нуклеиновую кислоту, кодирующую фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сай́т-специфическим образом.

Аспекты настоящего раскрытия относятся к клетке, включающей первую

чужеродную нуклеиновую кислоту, кодирующую фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом и включающую индуцируемый промотор для содействия экспрессии фермента. Таким образом, экспрессия может регулироваться, 5 например, она может быть запущена и может быть остановлена.

Аспекты настоящего раскрытия нацелены на клетку, включающую первую чужеродную нуклеиновую кислоту, кодирующей фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом, где первая чужеродная нуклеиновая кислота 10 является удаляемой из геномной ДНК клетки с помощью удаляющего фермента, такого как транспозаза.

Аспекты настоящего раскрытия относятся к способу изменения ДНК-мишени в клетке, экспрессирующей фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт- 15 специфическим образом, который включает (а) введение в клетку первой чужеродной нуклеиновой кислоты, кодирующей последовательность донорной нуклеиновой кислоты, введение в клетку из среды, окружающей клетку, РНК, комплементарной ДНК-мишени, которая направляет фермент на ДНК-мишень, где и РНК и фермент являются членами комплекса колокализации на ДНК-мишени, где последовательность донорной 20 нуклеиновой кислоты экспрессируется, где РНК и фермент колокализованы на ДНК-мишени, фермент расщепляет ДНК-мишень и донорная нуклеиновая кислота встраивается в ДНК-мишень с получением измененной ДНК в клетке.

Аспекты настоящего раскрытия направлены на использование направляемого РНК ДНК-связывающего белка для генетической модификации стволовых клеток. В одном 25 аспекте, стволовая клетка генетически модифицирована для включения нуклеиновой кислоты, кодирующей направляемый РНК ДНК-связывающий белок и стволовая клетка экспрессирует направляемый РНК ДНК-связывающий белок. Согласно конкретному аспекту, донорные нуклеиновые кислоты для введения специфических 30 мутаций оптимизированы для редактирования генома с использованием либо модифицированных TALEN, либо направляемого РНК ДНК-связывающего белка.

Аспекты настоящего раскрытия направлены на модификацию ДНК, например, мультиплексную модификацию ДНК, в стволовой клетке с помощью одной или нескольких направляющих РНК (рибонуклеиновых кислот), для того, чтобы направить фермент, обладающий нуклеазной активностью, экспрессируемый стволовой клеткой, 35 например, ДНК-связывающий белок, имеющий активность нуклеазы, в указанное положение на ДНК (дезоксирибонуклеиновой кислоте), где фермент разрезает ДНК и экзогенная донорная нуклеиновая кислота встраивается в ДНК, например, путем гомологичной рекомбинации. Аспекты настоящего изобретения включают циклические или повторяющиеся стадии модификации ДНК стволовой клетки для создания стволовой 40 клетки, имеющей множество модификаций ДНК внутри клетки. Модификации могут включать введение экзогенных донорных нуклеиновых кислот.

Множественные экзогенные вставки нуклеиновых кислот могут быть осуществлены одиночной стадией введения в стволовую клетку, которая экспрессирует фермент, нуклеиновых кислот, кодирующих множество РНК и множество экзогенных донорных 45 нуклеиновых кислот, например, путем котрансформации, где РНК экспрессируются и где каждая РНК в совокупности направляет фермент на конкретный участок ДНК, фермент разрезает ДНК и одна из множества экзогенных нуклеиновых кислот встраивается в ДНК в месте разреза. Согласно этому аспекту, многие изменения или

модификации ДНК в клетке создаются в одиночном цикле.

Множественные экзогенные вставки нуклеиновых кислот в стволовую клетку могут быть осуществлены путем повторных стадий или циклов введения в стволовую клетку, которая экспрессирует фермент, одной или нескольких нуклеиновых кислот,

- 5 кодирующих одну или несколько РНК или множество РНК и одну или несколько экзогенных нуклеиновых кислот или множество экзогенных нуклеиновых кислот, где РНК экспрессируется и направляет фермент на конкретный участок ДНК, фермент разрезает ДНК и экзогенная нуклеиновая кислота встраивается в ДНК в месте разреза, при этом получается клетка, обладающая несколькими изменениями или вставками
- 10 экзогенной ДНК в ДНК внутри стволовой клетки. Согласно одному аспекту, стволовая клетка, экспрессирующая фермент, была генетически изменена для экспрессии фермента, например, путем введения в клетку нуклеиновой кислоты, кодирующей фермент, которая может экспрессироваться стволовой клеткой. Таким образом, аспекты настоящего изобретения включают циклические стадии введения РНК в стволовую клетку, которая
- 15 экспрессирует фермент, введения экзогенной донорной нуклеиновой кислоты в стволовую клетку, которая экспрессирует РНК, образования комплекса колокализации РНК, фермента и ДНК, ферментативное разрезание ДНК ферментом, и вставку донорной нуклеиновой кислоты в ДНК. Цикл или повторение вышеуказанных стадий приводит к мультиплексной генетической модификации стволовых клеток во множестве
- 20 локусов, т. е. приводит к образованию стволовой клетки, имеющей множество генетических модификаций.

В соответствии с некоторыми аспектами, ДНК-связывающие белки или ферменты в пределах объема настоящего изобретения включают белок, который образует комплекс с направляющей РНК, и с помощью направляющей РНК комплекс

- 25 направляется к двухцепочечной последовательности ДНК, где комплекс связывается с последовательностью ДНК. Согласно одному аспекту, фермент может быть РНК направляемым ДНК-связывающим белком, таким как РНК направляемый ДНК-связывающий белок системы CRISPR II типа, который связывается с ДНК и направляется РНК. Согласно одному аспекту, направляемый РНК ДНК-связывающий белок
- 30 представляет собой белок Cas9.

Этот аспект настоящего изобретения может быть отнесен к колокализации РНК и ДНК-связывающего белка с двухцепочечной ДНК. Таким образом, комплекс ДНК-связывающего белка-направляющей РНК может быть использован для расщепления множества участков на двухцепочечной ДНК, с образованием таким образом стволовой клетки с несколькими генетическими модификациями, например, множественными вставками экзогенной донорной ДНК.

- 35 В соответствии с некоторыми аспектами, предлагается способ внесения нескольких изменений в ДНК-мишень в стволовой клетке, экспрессирующей фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который
- 40 расщепляет ДНК-мишень сайт-специфическим образом, который включает (а) введение в стволовую клетку первой чужеродной нуклеиновой кислоты, кодирующей одну или несколько РНК, комплементарных ДНК-мишени, которые направляют фермент к ДНК-мишени, где одна или несколько РНК и фермент являются членами комплекса колокализации с ДНК-мишенью, введение в стволовую клетку второй чужеродной нуклеиновой кислоты, кодирующей одну или несколько последовательностей донорной нуклеиновой кислоты, где одна или несколько РНК и фермент колокализуются с ДНК-мишенью, фермент расщепляет
- 45

ДНК и донорная нуклеиновая кислота встраивается в ДНК-мишень, с получением таким образом измененной ДНК в стволовой клетке, и повторение стадии (а) несколько раз, для внесения множества изменений в ДНК стволовой клетки.

Согласно одному аспекту, длина РНК составляет от около 10 до около 500

⁵ нуклеотидов. Согласно одному аспекту, длина РНК составляет от около 20 до около 100 нуклеотидов.

Согласно одному аспекту, одна или несколько РНК представляют собой направляющую РНК. Согласно одному аспекту, одна или несколько РНК представляют собой химеру tracrRNA-crRNA.

¹⁰ Согласно одному аспекту, ДНК представляет собой геномную ДНК, митохондриальную ДНК, вирусную ДНК или экзогенную ДНК.

Согласно одному аспекту, клетка может быть генетически модифицирована для обратимого включения нукleinовой кислоты, кодирующей ДНК-связывающий фермент, с использованием вектора, который может быть легко удален с помощью фермента.

¹⁵ Полезные векторы и способы известны специалистам в данной области техники и включают лентивирусы, аденоассоциированный вирус, опосредованные нуклеазами и интегразами способы направленной вставки и способы вставки, опосредованные транспозонами. Согласно одному аспекту, нукleinовая кислота, кодирующая ДНК-связывающую фермент, который был добавлен, например, с помощью кассеты или ²⁰ вектора, к примеру, может быть удалена полностью вместе с кассетой и вектором, не оставляя часть такой нукleinовой кислоты, кассеты или вектора в геномной ДНК. Такое удаление упоминается в данной области техники как «не оставляющее шрамов» удаление, поскольку геном становится таким же, каким он был до добавления нукleinовой кислоты, кассеты или вектора.

²⁵ Одним примерным воплощением вставки и удаления без шрамов является вектор PiggyBac, коммерчески доступный у System Biosciences.

Дополнительные признаки и преимущества определенных воплощений настоящего изобретения станут более очевидными в следующем описании воплощений и чертежей, и из формулы изобретения.

³⁰ Краткое описание чертежей

Вышеуказанные и другие признаки и преимущества настоящих воплощений будут более понятны из нижеследующего подробного описания иллюстративных воплощений, рассматриваемых вместе с сопроводительными чертежами, где:

³⁵ Фиг. 1 относится к функциональным тестам ге-TALENs в соматических и стволовых клетках человека.

(а) Схематическое изображение экспериментального проекта для тестирования эффективности направленного воздействия на геном. Интегрированная в геном последовательность, кодирующая GFP, нарушается вставкой стоп-кодона и 68 п.о. геномного фрагмента, полученного из локуса AAVS1 (внизу). Восстановление

⁴⁰ последовательности GFP с помощью нуклеазо-опосредованной гомологичной рекомбинации с помощью донора tGFP (вверху) приводит к образованию клеток GFP⁺, которые могут быть количественно оценены с помощью FACS. Re-TALENs и TALENs направлено воздействуют на одинаковые последовательности в пределах фрагментов AAVS1.

⁴⁵ (б) Столбиковая диаграмма, демонстрирующая процент GFP⁺ клеток, полученных с помощью только tGFP-донора отдельно, с помощью TALENs с tGFP-донором и ге-TALENs с tGFP-донором в локусе-мишени, согласно измерениям с помощью FACS. (N = 3, планки погрешности = SD). Репрезентативные графики FACS показаны ниже.

(c) Схематический обзор, изображающий стратегию направленного воздействия, для нативного локуса AAVS1. Плазмида-донор содержит акцептор сплайсинга (SA) - 2A (саморасщепляющиеся пептиды), ген устойчивости к пуромицину (PURO) и GFP (см. источник 10, включенный в данный документ ссылкой в полном объеме). Места расположения ПЦР-праймеров, используемых для обнаружения успешных событий редактирования, изображены в виде голубых стрелок.

(d) Подвергнутые успешному направленному воздействию клоны PGP1 hiPSCs отбирали на пуромицине (0,5 мкг/мл) в течение 2 недель. Показаны микроскопические изображения трех представительных GFP⁺ клонов. Клетки также окрашивали по 10 маркерам плюрипотентности TRA-1-60. Измерительная линейка: 200 мкм.

(e) Анализы ПЦР, выполненные на этих моноклональных GFP⁺ hiPSC клонах, продемонстрировали успешное встраивание донорных кассет в участок AAVS1 (дорожки 1,2,3), в то время как обычные hiPSCs не демонстрировали никаких признаков успешного встраивания (дорожка C).

(f) Отображение последовательности SEQ ID NO: 1.

Фиг. 2 относится к сравнению эффективности направленного воздействия на геном с помощью reTALENs и Cas9-гидРНК на CCR5 в iPSCs.

(a) Схематическое изображение эксперимента по геномной инженерии. В целевой участок пары ге-TALEN или Cas9-гидРНК в PGP1 hiPSCs доставляли 90-мер ssODN, несущий 2 п.о. несоответствие относительно геномной ДНК, вместе с конструкциями reTALEN или Cas9-гидРНК. Участки расщепления нуклеазами изображены на фигуре в виде красных стрелок.

(b) Анализ глубоким секвенированием эффективностей HDR и NHEJ для пар ге-TALEN (CCR5 #3) и ssODN, или Cas9-гидРНК и ssODN. Изменения в геноме hiPSCs анализировали на основе данных о последовательности, полученной высокопроизводительными способами, с помощью GEAS. Вверху: HDR количественно оценивали из фракции ридов, содержащих 2 п.о. точечные мутации, встроенных в центр ssODN (синий), а активность NHEJ количественно оценивали из фракции делеций (серые)/вставок (красные) в каждом конкретном положении в геноме. Для графиков reTALEN и ssODN, зеленый пунктир нанесен для того, чтобы отметить внешнюю границу участков связывания пар ге-TALENs, которые находятся в положениях -26 п.о. и + 26 п.о. относительно центра двух участков связывания ге-TALEN. Для графиков Cas9-гидРНК и ssODN, зеленый пунктир отмечает внешнюю границу участка направленного воздействия гидРНК, который находится в положениях -20 и -1 п.о. по отношению к последовательности РАМ. Внизу: распределение по размерам Делеция/Вставка в hiPSCs анализировали по всей популяции NHEJ с обработкой, указанной выше.

(c) Эффективность редактирования генома re-TALENs и Cas9-гидРНК, направленно воздействовавших на CCR5 в PGP1 hiPSCs.

Вверху: схематическое представление участков, подвергшихся направленному редактированию геномов в CCR5. 15 участков для направленного воздействия показаны синими стрелками снизу. Для каждого участка клетки котрансфектировали парой ге-TALENs и их соответствующим донорным ssODN, несущим 2 п.о. несовпадение относительно геномной ДНК. Эффективность редактирования генома анализировали через 6 дней после трансфекции. Аналогичным образом, 15 Cas9-гидРНК трансфектировали с соответствующими ssODNs индивидуально в PGP1-hiPSCs для направленного воздействия не те же 15 участков и анализировали эффективность через 6 дней после трансфекции. Внизу: эффективность редактирования генома ге-TALENs и Cas9-гидРНК при направленном целевом воздействии на CCR5 в PGP1 hiPSCs. Панели

1 и 2 демонстрируют эффективности NHEJ и HDR, опосредованные reTALENs. Панель 3 и 4 демонстрируют эффективности NHEJ и HDR, опосредованные Cas9-гидРНК. Показатели NHEJ рассчитывали по частоте геномных аллелей, несущих делеции или вставки в подвергнутом направленному воздействию участке; показатели HDR

5 рассчитывали по частоте геномных аллелей, несущих 2 п.о. несоответствия. Панель 5, профиль DNaseI HS для клеточной линии hiPSC из базы данных ENCODE (Duke DNase HS, iPS NIH i7 DS). Следует отметить, что масштабы панелей отличаются.

Фиг. 3 направлена на изучение функциональных параметров, регулирующих ssODN-опосредованную HDR с re-TALENs или Cas9-гидРНК в PGP1 hiPSCs.

10 (a) PGP1 hiPSCs котрансфецировали с помощью пары re-TALENs (#3) и ssODNs различной длины (50, 70, 90, 110, 130 150 170 нуклеотидов). Все ssODNs обладали идентичным 2 п.о. несовпадением относительно геномной ДНК в середине своей последовательности. 90-мерный ssODN достиг оптимального уровня HDR в подвергнутом направленному воздействию геноме. Оценка HDR-, NHEJ-опосредованной 15 эффективности делеции и вставки описана в данном документе.

20 (b) 90-мерные ssODNs, соответствующие паре re-TALEN # 3, каждая из которых содержит 2 п.о. несовпадение (A) в центре и дополнительное 2 п.о. несовпадение (B) в разных положениях, смешенных от A (где смещения варьировали от -30 п.о. -> 30 п.о.), были использованы для тестирования последствий отклонения от гомологии вдоль ssODN. Эффективность редактирования генома каждого ssODN оценивали в PGP1 hiPSCs. Нижняя столбиковая диаграмма демонстрирует частоту включения только A, только B и A + B в подвергнутом направленному воздействию геноме. Показатели HDR уменьшались по мере увеличения расстояния от центра участка с отклонением по гомологии.

25 (c) ssODNs, нацеленные на участки с варьирующими расстояниями (-620 п.о. ~ 480 п.о.) от участка-мишени пары #3 re-TALEN, были испытаны для оценки максимального расстояния, в пределах которого ssODNs могут быть помещены для введения мутаций. Все ssODNs несли 2 п.о. несоответствие в середине своих последовательностей. Наблюдали минимальную эффективность HDR (<= 0,06%), когда несовпадение в ssODN 30 располагалось на расстоянии 40 п.о. от середины участка связывания пары re-TALEN.

(d) PGP1 hiPSCs котрансфецировали Cas9-гидРНК (AAVS1) и ssODNs различной ориентации (Ос: комплементарная гидРНК; On: некомплémentарная гидРНК) и разной длины (30, 50, 70, 90, 110 нуклеотидов). Все ssODNs обладали идентичными 2 п.о. несовпадениями относительно геномной ДНК в середине своих последовательностей. 35 70-мерный Os достигал оптимального значения HDR в подвергнутом направленному воздействию геноме.

Фиг. 4 направлен на использование re-TALENs и ssODNs для получения моноклональных hiPSCc отредактированным геномом без селекции.

(a) Временная шкала эксперимента.

40 (b) Эффективность редактирования генома пары re-TALENs и ssODN (# 3), оценивали с помощью платформы NGS, описанной на фигуре 2b.

(c) Результаты секвенирования по Сенгеру колоний моноклональных hiPSC после редактирования генома. 2 п.о. гетерогенный генотип (CT/CT->TA/CT) был успешно внедрен в геном колоний PGP 1-IPS-3-11, PGPI-IPS-3-13.

45 (d) Иммунофлуоресцентное окрашивание подвергнутых направленному воздействию PGPI-IPS-3-11. Клетки окрашивали на маркеры плюрипотентности TRA-1-60 и SSEA4.

(e) Гематоксилин-эозиновое окрашивание срезов тератом, полученных из моноклональных клеток PGPI-IPS-3-11.

Фиг. 5. Дизайн reTALE. (а) Выравнивание последовательности исходного мономера TALE RVD относительно мономеров в re-TALE-16.5 (re-TALE-M1>re-TALE-M17).

Нуклеотидные изменения относительно исходной последовательности выделены серым,

(б) Тест повторяемости re-TALE с помощью ПЦР. Верхняя панель показывает структуру 5 re-TALE/TALE и позиции праймеров для ПЦР-реакции. Нижняя панель показывает полосы ПЦР в условиях, указанных ниже. Электрофоретическая картина фрагментов ПЦР представлена с помощью исходного шаблона TALE (правая дорожка).

Фиг. 6. Дизайн и практическое осуществление сборки TASA (TALE Single-incubation Assembly), (а) Схематическое представление библиотеки димерных блоков re-TALE для

10 сборки TASA. Представлена библиотека 10 димерных блоков re-TALE, кодирующих два RVDs. Внутри каждого блока, все 16 димеров имеют одну и ту же последовательность, за исключением последовательностей, кодирующих RVD; Димеры в различных блоках имеют отличающиеся последовательности, но спроектированы таким образом, чтобы они имели общие 32 п.о. перекрывания с соседними блоками.

15 Нуклеотидная и аминокислотная последовательности одного димера (Блок 6_AC) представлены справа.

(б) Схематическое представление сборки TASA. На левой панели показан способ сборки TASA: однореакторная реакция инкубации проводится с ферментной смесью/ re-TALE блок s/re-TALE-N/TF каркасными векторами. Продукт реакции может быть

20 использован непосредственно для трансформации бактерий. Правая панель демонстрирует механизм TASA. Вектор назначения линеаризуется эндонуклеазой при 37°C для того, чтобы отрезать контрселективную кассету ccdB; экзонуклеаза, которая обрабатывает конец блоков и линеаризованных векторов, экспонирует оцДНК выступы на конце фрагментов, что позволяет блокам и векторным каркасам соединяться в 25 назначенном порядке. При повышении температуры до 50°C, полимеразы и лигазы работают вместе для того, чтобы заделать разрыв, что позволяет получить конечные конструкции, готовые к трансформации.

30 (с) Эффективность сборки TASA для re-TALEs, обладающих разными длинами мономеров. Используемые для сборки блоки показаны слева, а эффективность сборки представлена справа.

Фиг. 7. Функциональность и целостность последовательности Lenti-reTALEs. Клетки 293T, трансдуцированные lenti-re-TALE-TF показали 36x активацию экспрессии репортера по сравнению с только отрицательным репортером.

Фиг. 8. Чувствительность и воспроизводимость GEAS.

35 (А) Информационный анализ предела обнаружения HDR. С учетом набора данных re-TALENs (# 10)/ssODN, были определены «риды», содержащие ожидаемое редактирование (HDR) и эти HDR-«риды» систематически удалялись для создания различных искусственных наборов данных с «разведенным» редактирующим сигналом. Наборы данных с 100, 99,8, 99,9, 98,9, 97,8, 89,2, 78,4, 64,9, 21,6, 10,8, 2,2, 1,1, 0,2, 0,1, 0,02

40 и 0% удалением HDR-«ридов» были сгенерированы для получения искусственных наборов с эффективностью HR в диапазоне 0-0,67%. Для каждого отдельного набора данных, была оценена взаимная информация (mutual information, MI) фонового сигнала (обозначено фиолетовым) и сигнала, полученного в участке направленного воздействия (обозначено зеленым). MI в участке направленного воздействия заметно выше, чем 45 фон, при этом эффективность HDR выше 0,0014%. По оценке предел обнаружения HDR находится в диапазоне от 0,0014% до 0,0071%. Расчет MI описан в данном документе.

(б) Тест воспроизводимости системы оценки редактирования генома. Пары графиков (верхний и нижний) демонстрируют результаты оценки HDR и NHEJ в двух повторах

с парой re-TALENs и типом клеток, указанным выше. Для каждого эксперимента независимо были проведены нуклеофекция, направленная амплификация генома, глубокое секвенирование и анализ данных. Разброс повторов при оценке редактирования генома рассчитывали как V2 (|HDR1-HDR2|)/((HDR+HDR2)/2) = AHDR/HDR и V2 (|NHEJ1-NHEJ2|)/((NHEJ1+NHEJ2)/2) = ANHEJ/NHEJ, а результаты по разбросу приведены ниже графиков. Средний разброс системы составил $(19\% + 11\% + 4\% + 9\% + 10\% + 35\%)/6 = 15\%$. Факторы, которые могут способствовать разбросу, включают статус клеток при нуклеофекции, эффективность нуклеофекции, охват и качество секвенирования.

Фиг. 9. Статистический анализ эффективностей NHEJ и HDR при использовании

reTALENs и Cas9-гидРНК на CCR5.

(а) Корреляция эффективностей HR и NHEJ при использовании reTALENs на одинаковых участках в iPSCs ($i=0,91$, $P < 10^{-5}$).

(б) Корреляция эффективностей HR и NHEJ, опосредованных Cas9-гидРНК на одинаковых участках в iPSCs ($i=0,74$, $p=0,002$).

(с) Корреляция эффективности NHEJ, опосредованной Cas9-гидРНК и температуры Т_m целевого участка для гидРНК в iPSCs ($r=0,52$, $p=0,04$)

Фиг. 10. Корреляционный анализ эффективности редактирования генома и эпигенетического состояния. Корреляции Пирсона использовали для изучения возможных связей между чувствительностью ДНКазы I и эффективностью геномной инженерии (HR, NHEJ). Наблюдаемую корреляцию сравнивали с рандомизированным набором ($N = 100000$). Наблюдаемые корреляции выше 95-го процентиля или ниже 5-го процентиля моделируемого распределения считались потенциальными ассоциациями. Не наблюдали заметной корреляции между чувствительностью ДНКазы I и эффективностью NHEJ/HR.

Фиг. 11. Влияние гомологичного спаривания при ssODN-опосредованном редактировании генома.

(а) В эксперименте, описанном на рисунке 3b, общая HDR снижалась, что измерялось по темпам, с которыми происходило включение среднего 2 п.о. несовпадения (A), по мере того, как росла дистанция на которую вторичные несовпадения B удалялись от A (относительное положение B по отношению к A, к варьирует от -30 до 30 п.о.). Более высокие темпы включения, когда B находится всего лишь в 10 п.о. от A (-10 п.о. и + 10 п.о.) могут отражать меньшую потребность в спаривании ssODN против геномной ДНК, проксимальной к двухцепочечному разрыву ДНК.

(б) Распределение длин генной конверсии вдоль ssODN. На каждом расстоянии B от A, доля событий HDR включает только A, тогда как друга доля включает как A, так и B. Эти два события могут интерпретироваться в терминах трактов конверсии генов (Elliott et al., 1998), в результате чего A+B-события представляют длинные конверсионные тракты, которые выходят за рамки только B и только A событий, представляющие более короткие пути, которые не достигают B. Согласно данной интерпретации, может быть оценено распределение трактов генной конверсии в обоих направлениях вдоль олигонуклеотида (середина ssODN определяется как 0, конверсионные тракты по направлению к 5'-концу ssODN - как «-» направление, а к 3'-концу - как «+» направление). Тракты генной конверсии постепенно уменьшаются по мере увеличения их длин, результат очень похожий на распределение трактов генной конверсии, наблюдаемое с дцДНК-донорами, но на сильно сжатой шкале дистанций в десятки п.о. для одноцепочечных олигодезоксинуклеотидов по сравнению с сотнями пар оснований для дцДНК-доноров.

(с) Анализы трактов генной конверсии с использованием одиночного ssODN, который

содержит ряд мутаций, и измерение непрерывных последовательностей включений. Использовали ssODN-донор с тремя парами 2 п.о. несовпадений (оранжевый), разнесенных с интервалом в 10 нуклеотидов в обе стороны от центрального 2 п.о. несовпадения (вверху). Было обнаружено незначительное количество «ридов» геномного

5 секвенирования (см. источник 62, включенный в данный документ ссылкой в полном объеме), несущих ≥ 1 несовпадения, определенных с помощью ssODN среди $> 300\,000$ «ридов» при секвенировании этой области. Все эти «риды» были нанесены на график (внизу), а последовательность «ридов» была закодирована условной окраской.

Оранжевый: определенные несоответствия; зеленый: последовательность дикого типа.

10 Геномное редактирование с этими ssODN дало паттерн, в котором средняя мутация отдельно была включена в 85% (53/62) моментов времени, с множеством В-несовпадений, включенных в другие моменты. Хотя количество событий В-включений было слишком низким для оценки распределения длин путей > 10 п.о., ясно, что преобладает область коротких путей от -10 до 10 п.о.

15 Фиг. 12. Эффективности редактирования генома Cas9-гидРНК нуклеазой и никазами. PGP1 iPSCs котрансфектировали комбинацией нуклеазы (C2) (Cas9-гидРНК) или никазы (Cc) (Cas9D10A-гидРНК) и ssODNs различной ориентации (Oc и On). Все ssODNs обладали идентичным 2 п.о. несовпадением относительно геномной ДНК в середине своей последовательности. Оценка HDR описана в данном документе.

20 Фиг. 13. Дизайн и оптимизация последовательности ге-TALE.

Последовательность ге-TALE подвергали эволюции в течение нескольких циклов дизайна для устранения повторов. В каждом цикле оценивали синонимичные последовательности из каждого повтора. Отбирали те из них, которые имели наибольшее расстояние Хэмминга к эволюционирующему ДНК. Окончательная последовательность 25 имела $ca i = 0,59 \Delta G = -9,8$ ккал/моль. Для выполнения общей основы для дизайна синтетического белка использовали пакет R.

Фиг. 14 представляет собой изображение геля, демонстрирующее проверку с помощью ПЦР геномной вставки Cas9 в клетки PGP1. Дорожки 3, 6, 9, 12 являются продуктами ПЦР обычных клеточных линий PGP1.

30 Фиг. 15 представляет собой график уровня экспрессии мРНК Cas9 при индукции.

Фиг. 16 представляет собой график, показывающий эффективность направленного воздействия на геном РНК с различным дизайном.

Фиг. 17 представляет собой график, показывающий эффективность направленного воздействия на геном 44% гомологичной рекомбинации, достигнутой с помощью химеры 35 направляющей РНК-донорной ДНК.

Фиг. 18 является схемой, демонстрирующей генотип изогенных клеточных линий PGP 1, полученных с помощью системы, описанной в данном документе. PGP1-IPS-BTHN имеет фенотип с одноклеточными делециями, как у пациентов с BTHN. PGP 1-NHEJ имеет 4 п.о. делеции, которые образуют мутации со сдвигом рамки считывания 40 другим образом.

Фиг. 19 представляет собой график, показывающий, что кардиомиоциты, полученные из изогенных PGP1 iPS, воспроизводят дефектную выработку АТФ и специфическую активность F1F0 АТФазы, как продемонстрировано в пациент-специфических клетках.

Фиг. 20 показывает последовательности для последовательности ге-TALEN-основной 45 и основной последовательности ге-TALE-TF.

Подробное описание

Аспекты настоящего изобретения относятся к применению TALEN, которая утратила определенные повторяющиеся последовательности, для инженерии нукleinовых кислот,

например, путем расщепления двухцепочечной нуклеиновой кислоты. Применение TALEN для расщепления двухцепочечной нуклеиновой кислоты может привести к негомологичному соединению концов (NHEJ) или к гомологичной рекомбинации (HR). Аспекты настоящего раскрытия также предусматривают применение TALEN, которая 5 утратила повторяющиеся последовательности, для инженерии нуклеиновых кислот, например, путем расщепления двухцепочечной нуклеиновой кислоты в присутствии донорной нуклеиновой кислоты и введения донорной нуклеиновой кислоты в двухцепочечную нуклеиновую кислоту, например, путем негомологичного соединения концов (NHEJ) или гомологичной рекомбинации (HR).

10 Эффекторные нуклеазы, подобные активаторам транскрипции (TALEN) известны в данной области техники и включают искусственные ферменты рестрикции, полученные слиянием ДНК-связывающего домена TAL-эффектора с ДНК-расщепляющим доменом. Ферменты рестрикции являются ферментами, которые разрезают цепи ДНК в определенной последовательности. Эффекторы, подобные активаторам транскрипции 15 (TALE) могут быть сконструированы для связывания с искомой последовательностью ДНК. См. Boch, Jens (February 2011). "TALEs of genome targeting". Nature Biotechnology 29 (2): 135-6, включенный в данный документ ссылкой в полном объеме. Путем объединения такого сконструированного TALE с ДНК-расщепляющим доменом (который разрезает нити ДНК), получается TALEN, которая является ферментом рестрикции, специфичным 20 к любой искомой последовательности ДНК. В соответствии с некоторыми аспектами, TALEN вводят в клетку для целевого редактирования нуклеиновой кислоты *in situ*, например, для редактирования генома *in situ*.

Согласно одному аспекту, неспецифический ДНК-расщепляющий домен с конца эндонуклеазы FokI может быть использован для создания гибридных нуклеаз, которые 25 активны в клетках дрожжей, растений и животных. Домен FokI функционирует в виде димера, требующего две конструкции с уникальными ДНК-связывающими доменами для участков в геноме-мишени с правильной ориентацией и расстоянием. Как число аминокислотных остатков между TALE ДНК-связывающим доменом и расщепляющим 30 доменом FokI, так и количества оснований между двумя отдельными участками связывания TALEN оказывают влияние на активность.

Взаимоотношение между аминокислотной последовательностью и распознаванием ДНК связывающим доменом TALE позволяет создавать конструируемые белки. Компьютерные программы, такие как «DNA Works» могут быть использованы для разработки TALE-конструкций. Другие способы разработки TALE-конструкций известны 35 специалистам в данной области техники. См. Cermak, T.; Doyle, E. L.; Christian, M.; Wang, L.; Zhang, Y.; Schmidt, C; Bailer, J. A.; Somia, N. V. et al. (2011). "Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting". Nucleic Acids Research. doi:10. 1093/nar/gkr218; Zhang, Feng; et. al. (February 2011). "Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription", Nature Biotechnology 29 (2): 149-53; Morbitzer, R.; Elsaesser, J.; Hausner, J.; Lahaye, T. (2011). "Assembly of custom 40 TALE-type DNA binding domains by modular cloning". Nucleic Acids Research. doi:10'1093/nar/gkrl51; Li, T.; Huang, S.; Zhao, X.; Wright, D.A.; Carpenter, S.; Spalding, M.H.; Weeks, D.P.; Yang, B. (2011). "Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes". Nucleic Acids Research. doi"10. 1093/nar/gkrl88; 45 GeipTer, R.; Scholze, H.; Hahn, S.; Streubel, J.; Bonas, U.; Behrens, S.E.; Boch, J. (2011). "Transcriptional Activators of Human Genes with Programmable DNA-Specificity". In Shiu, Shin-Han. PLoS ONE 6 (5): e19509; Weber, E.; Graetzner, R.; Werner, S.; Engler, C; Marillonnet, S. (2011). "Assembly of Designer TAL Effectors by Golden Gate Cloning". In Bendahmane,

Mohammed. PLoS One 6 (5): e19722, которые включены в данный документ ссылкой в полном объеме.

В соответствии с типичным аспектом, сразу после сборки гены TALEN можно вставлять в плазмиды, в соответствии с определенными воплощениями; плазмиды затем используются для трансфекции клетки-мишени, при этом генные продукты экспрессируются и входят в ядро для доступа к геному. Согласно типичным аспектам, TALEN, как описано в данном документе, могут быть использованы для редактирования целевой нуклеиновой кислоты, например, генома, с помощью индукции двухцепочечных разрывов (DSB), на которые клетки реагируют посредством механизмов репарации.

Примерные механизмы репарации включают негомологичное соединение концов (NHEJ), который восстанавливает ДНК по обе стороны от двухцепочечного разрыва, где имеется весьма незначительное перекрывание для отжига, или такое перекрывание отсутствует вовсе. Этот репаративный механизм индуцирует ошибки в геноме, через вставку или делецию (indel), или хромосомную перестройку; такие ошибки могут сделать генные продукты, кодируемые в данном месте, нефункциональными. См. Miller, Jeffrey; et. al. (February 2011). "A TALE nuclease architecture for efficient genome editing". Nature Biotechnology 29 (2): 143-8, включенный в данный документ ссылкой в полном объеме. Поскольку эта активность может изменяться в зависимости от используемых вида, типа клеток, гена-мишени, и нуклеазы, активность можно контролировать с помощью анализа гетеродуплексного расщепления, который не обнаруживает никакой разницы между двумя аллелями, амплифицированными с помощью ПЦР. Продукты расщепления могут быть визуализированы на простом агарозном геле или блочной гелевой системе.

В ином случае, ДНК может быть введена в геном посредством NHEJ в присутствии экзогенных фрагментов двухцепочечной ДНК. Управляемая гомологией репарация также может ввести чужеродную ДНК в DSB, поскольку трансфекционные двухцепочечные последовательности используются в качестве шаблонов для ферментов репарации. В соответствии с некоторыми аспектами, TALEN, описанные в данном документе, могут быть использованы для получения стабильно модифицированных человеческих эмбриональных стволовых клеток и клонов индуцированных плюрипотентных стволовых клеток (iPSCs). В соответствии с некоторыми аспектами, TALEN, описанные в данном документе, могут быть использованы для получения нокаутных видов, таких как *C. elegans*, нокаутные крысы, нокаутные мыши или нокаутные данио-перио.

Согласно одному аспекту настоящего раскрытия, воплощения нацелены на применение экзогенной ДНК, нуклеазных ферментов, таких как ДНК-связывающие белки и направляющие РНК, для колокализации с ДНК в стволовой клетке и гидролиза или разрезания ДНК с вставкой экзогенной ДНК. Такие ДНК-связывающие белки хорошо известны специалистам в данной области техники, для связывания с ДНК для различных целей. Такие ДНК-связывающие белки могут быть природными. ДНК-связывающие белки, включенные в объем настоящего изобретения, включают те, которые могут направляться РНК, которые называются в данном документе как направляющие РНК. Согласно этому аспекту, направляющая РНК и РНК-направляемый ДНК-связывающий белок образуют комплекс колокализации на ДНК. Такие ДНК-связывающие белки, обладающие нуклеазной активностью, известны специалистам в данной области техники, и включают природные ДНК-связывающие белки, имеющие нуклеазную активность, такие как белки Cas9, присутствующие, например, в системах CRISPR II типа. Такие белки Cas9 и системы CRISPR типа II хорошо описаны в данной

области техники. См. Makarova et al., *Nature Reviews, Microbiology*, Vol. 9, June 2011, pp. 467-477 включая всю дополнительную информацию, включенный в данный документ ссылкой в полном объеме.

Типичные ДНК-связывающие белки имеют функцию нуклеазной активности для

- 5 внесения одноцепочечных разрывов (ников) или разрезания двухцепочечной ДНК. Такая нуклеазная активность может быть результатом действия ДНК-связывающего белка, имеющего одну или несколько полипептидных последовательностей, демонстрирующих нуклеазную активность. Такие типичные ДНК-связывающие белки могут иметь два отдельных нуклеазных домена, при этом каждый домен отвечает за
- 10 разрезание или внесение одноцепочечных разрывов конкретной цепи двухцепочечной ДНК. Типичные полипептидные последовательности, имеющие нуклеазную активность, известные специалистам в данной области, включают домен, родственный нуклеазе McrA-HNH и домен RuvC-подобной нуклеазы. Соответственно, типичными ДНК-связывающими белками являются те, которые по своей природе содержат один или
- 15 несколько из числа домена, родственного нуклеазе McrA-HNH и домен RuvC-подобной нуклеазы.

Типичный ДНК-связывающий белок представляет собой РНК управляемый ДНК-связывающий белок системы CRISPR II типа. Типичный ДНК-связывающий белок представляет собой белок Cas9.

- 20 В *S. pyogenes*, Cas9Cas9 образует двухцепочечной разрыв с тупыми концами на 3 п.о. выше мотива PAM (прилегающий к протоспайсеру мотив) с помощью процесса, опосредованного двумя каталитическими доменами в белке: доменом HNH, который расщепляет комплементарную нить ДНК и доменом подобного RuvC, который расщепляет некомплектарную цепь. См. Jinke et al., *Science*, 337, 816-821 (2012),
- 25 включенный в данный документ ссылкой в полном объеме. Белки Cas9, как известно, существуют во многих системах CRISPR типа II, включая те, которые определены в дополнительных документах к Makarova et al., *Nature Reviews, Microbiology*, Vol. 9, June 2011, pp. 467-477: *Methanococcus maripaludis* C7; *Corynebacterium diphtheriae*; *Corynebacterium efficiens* YS-314; *Corynebacterium glutamicum* ATCC 13032 Kitasato; *Corynebacterium*
- 30 *glutamicum* ATCC 13032 Bielefeld; *Corynebacterium glutamicum* R; *Corynebacterium kroppenstedtii* DSM 44385; *Mycobacterium abscessus* ATCC 19977; *Nocardia farcinica* IFM10152; *Rhodococcus erythropolis* PR4; *Rhodococcus jostii* RHA1; *Rhodococcus opacus* B4 uid36573; *Acidothermus cellulolyticus* 11B; *Arthrobacter chlorophenolicus* A6; *Kribbella flava* DSM 17836 uid43465; *Thermomonospora curvata* DSM 43183; *Bifidobacterium dentium* Bdl;
- 35 *Bifidobacterium longum* DJO10A; *Slackia heliotrinireducens* DSM 20476; *Persephonella marina* EX HI; *Bacteroides fragilis* NCTC 9434; *Capnocytophaga ochracea* DSM 7271; *Flavobacterium psychrophilum* JIP02 86; *Akkermansia muciniphila* ATCC BAA 835; *Roseiflexus castenholzii* DSM 13941; *Roseiflexus RSI*; *Synechocystis* PCC6803; *Elusimicrobium minutum* Peil91; бактериальный филотип Rs D17 некультивированной группы 1 Termite; *Fibrobacter*
- 40 *succinogenes* S85; *Bacillus cereus* ATCC 10987; *Listeria innocua*; *Lactobacillus casei*; *Lactobacillus rhamnosus* GG; *Lactobacillus salivarius* UCC118; *Streptococcus agalactiae* A909; *Streptococcus agalactiae* NEM316; *Streptococcus agalactiae* 2603; *Streptococcus dysgalactiae* equisimilis GGS 124; *Streptococcus equi* zooepidemicus MGCS10565; *Streptococcus galloyticus* UCN34 uid46061; *Streptococcus gordonii* Challis subst CHI; *Streptococcus mutans* NN2025 uid46353; *Streptococcus*
- 45 *mutans*; *Streptococcus pyogenes* M1 GAS; *Streptococcus pyogenes* MGAS5005; *Streptococcus pyogenes* MGAS2096; *Streptococcus pyogenes* MGAS9429; *Streptococcus pyogenes* MGAS10270; *Streptococcus pyogenes* MGAS6180; *Streptococcus pyogenes* MGAS315; *Streptococcus pyogenes* SSI-1; *Streptococcus pyogenes* MGAS 10750; *Streptococcus pyogenes* NZ131; *Streptococcus*

thermophiles CNRZ1066; Streptococcus thermophiles LMD-9; Streptococcus thermophiles LMG 18311; Clostridium botulinum A3 Loch Maree; Clostridium botulinum B Eklund 17B; Clostridium botulinum Ba4 657; Clostridium botulinum F Langeland; Clostridium cellulolyticum H10; Finegoldia magna ATCC 29328; Eubacterium rectale ATCC 33656; Mycoplasma gallisepticum;

5 Mycoplasma mobile 163K; Mycoplasma penetrans; Mycoplasma synoviae 53; Streptobacillus moniliformis DSM 12112; Bradyrhizobium BTail; Nitrobacter hamburgensis X14; Rhodopseudomonas palustris BisB18; Rhodopseudomonas palustris BisB5; Parvibaculum lavamentivorans DS-1; Dinoroseobacter shibae DFL 12; Gluconacetobacter diazotrophicus Pal 5 FAPERJ; Gluconacetobacter diazotrophicus Pal 5 JGI; Azospirillum B510 uid46085;

10 Rhodospirillum rubrum ATCC 11170; Diaphorobacter TPSY uid29975; Verminephrobacter eiseniae EF01-2; Neisseria meningitidis 053442; Neisseria meningitidis alphal4; Neisseria meningitidis Z2491; Desulfovibrio salexigens DSM 2638; Campylobacter jejuni doylei 269 97; Campylobacter jejuni 81116; Campylobacter jejuni; Campylobacter lari RM2100; Helicobacter hepaticus; Wolinella succinogenes; Tolumonas auensis DSM 9187; Pseudoalteromonas atlantica

15 T6c; Shewanella pealeana ATCC 700345; Legionella pneumophila Paris; Actinobacillus succinogenes 130Z; Pasteurella multocida; Francisella tularensis novicida U112; Francisella tularensis holarktika; Francisella tularensis FSC 198; Francisella tularensis tularensis; Francisella tularensis WY96-3418; и Treponema denticola ATCC 35405. Соответственно, аспекты настоящего раскрытия направлены на белок Cas9, присутствующий в системе CRISPR

20 II типа.

Белок Cas9 может быть назван специалистом в данной области литературы как Csn1. Белок Cas9 из *S. pyogenes* показан ниже. См. Deltcheva et al., Nature 471, 602-607 (2011), включенный в данный документ ссылкой в полном объеме.

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDGET
 25 AEATRLKRTARRRYTRRKNRICYLQEIFSNEAKVDDSFHRLEESFLVEEDKKHER
 HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAIDLRIYLALAHMIKFRGHFIEGD
 LNPDNSDVKLFQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG
 EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDLDNLLAQIGDQYAD
 LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
 30 KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
 DNGSIPHQIHLGELHAILRRQEDFYPFLKDNRKIEKILTFRIPYYVGPLARGNSRAFW
 MTRKSEETITPWNFEEVVDKGASAQSIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN
 ELTKVKYVTGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVE
 ISGVEDRFNASLGTYHDLLKIJKDKDFLDNEENEDILEDIVLTLFEDREMIERLKTY
 35 AHLFDDKVMQLKRRRTGWGRSLRKLINGIRDQSGKTILDPLKSDGFANRNFQ
 QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDELVKVM
 GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQLKEHPVENTQLQNE
 KLYLYLQNNGRDMYVDQELDINRLSDYDWDHVQPSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLNAKLITQRKFNDNLTKAERGGLSELDKAGFIKRQLV
 40 ETRQITKHVAQILDMSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINN
 YHHAHDAYLNAVVTALIKKYPKLESEFVYGYDVKVYDVRKMIAKSEQEIGKATAK
 YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVN
 IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWPDKYGGFDSPTVAYSVLVVAKVE
 KGKSKKLKSVKELLGITMERSSEKPNPIDFLEAKGYKEVKKDLIILPKYSLFELEN
 45 GRKRMLASAGELQKGNELALPSKYVNFLYASHYEKLKGSPEDNEQKQLFVEQHK
 HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHDKPIREQAENIIHLFTLTLNGAP
 AAFKYFDTTIDRKRYTSTKEVLDATLIHQHSITGLYETRIDLSQLGGD-

Согласно одному аспекту, управляемый РНК ДНК-связывающий белок включает

гомологи и ортологи Cas9, которые сохраняют способность белка связываться с ДНК, быть управляемыми РНК и разрезать ДНК. Согласно одному аспекту, белок Cas9 включает последовательность, которая представлена в естественном белке Cas9 из *S. ryogenes* и белковых последовательностях, гомологичных, по меньшей мере, на 30%, 5 40%, 50%, 60%, 70%, 80%, 90%, 95 %, 98% или 99% к этому белку, и, являющихся ДНК-связывающим белком, таким как направляемый РНК ДНК-связывающий белок.

Согласно одному аспекту предлагается сконструированная система Cas9-гидРНК, которая позволяет осуществить направляемое РНК разрезание генома сайт-специфическим образом в стволовой клетке, при необходимости, и модификацию генома 10 стволовой клетки путем вставки экзогенных донорных нуклеиновых кислот.

Направляющие РНК являются комплементарными к участкам-мишеням или локусам-мишеням на ДНК. Направляющие РНК могут быть химерами crRNA-tracrRNA.

Направляющие РНК могут быть введены из среды, окружающей клетки. Таким образом, предлагается способ непрерывной модификации клетки, которая осуществляется при 15 внесении направляющих РНК в окружающую среду и при поглощении данных направляющих РНК и при обогащении среды дополнительными направляющими РНК. Обогащение может идти непрерывно. Cas9 связывается с или около ДНК-мишени в геноме. Одна или несколько направляющих РНК связываются с или около ДНК-мишени 20 в геноме. Cas9 разрезает ДНК-мишень в геноме и экзогенная донорская ДНК встраивается в ДНК в месте разреза.

Соответственно, способы направлены на применение направляющей РНК с белком Cas9 и экзогенной донорной нуклеиновой кислотой для мультиплексных вставок экзогенных донорных нуклеиновых кислот в ДНК стволовой клетки, экспрессирующей Cas9, посредством циклической вставки нуклеиновой кислоты, кодирующей РНК (или 25 путем предоставления РНК из окружающей среды) и экзогенной донорной нуклеиновой кислоты, экспрессирующей РНК (или путем поглощения РНК), колокализации РНК, Cas9 и ДНК таким образом, чтобы происходило разрезание ДНК, и вставки экзогенной донорной нуклеиновой кислоты. Стадии способа могут быть циклически повторены любое необходимое количество раз для получения необходимого количества 30 модификаций ДНК. Способы настоящего описания соответственно направлены на редактирование генов-мишений с помощью белков Cas9 и направляющих РНК, описанных в данном документе, с целью осуществления мультиплексной генетической и эпигенетической инженерии стволовых клеток.

Дальнейшие аспекты настоящего описания направлены на применение ДНК-35 связывающих белков или систем (например, модифицированных TALENs или Cas9, описанных в данном документе) в целом для мультиплексной вставки экзогенных донорных нуклеиновых кислот в ДНК, такую как геномная ДНК стволовой клетки, например, человеческой стволовой клетки. Специалист в данной области техники легко определит типичные ДНК-связывающие системы на основании настоящего описания.

40 Клетки в соответствии с настоящим описанием, если не указано иное, включают любую клетку, в которую могут быть введены и в которой могут быть экспрессированы чужеродные нуклеиновые кислоты, как описано в данном документе. Следует понимать, что основные принципы настоящего раскрытия, описанные в данном документе, не ограничиваются типом клеток. Клетки, согласно настоящему изобретению, включают 45 соматические клетки, стволовые клетки, эукариотические клетки, прокариотические клетки, клетки животных, растительные клетки, клетки грибов, клетки, клетки архей, клетки эубактерий и т.п. Клетки включают эукариотические клетки, такие как дрожжевые клетки, растительные клетки и клетки животных. Конкретные клетки

включают клетки млекопитающих, такие как клетки человека. Кроме того, клетки включают любую клетку, в которой было бы выгодно или желательно модифицировать ДНК.

Нуклеиновой кислоты включают любую последовательность нуклеиновой кислоты,

- 5 в случае которой TALEN или направляемый РНК ДНК-связывающий белок, обладающий нуклеазной активностью, как описано в данном документе, могут быть полезны для внесения одноцепочных разрывов или разрезания. Нуклеиновые кислоты-мишени включают любую последовательность нуклеиновой кислоты, в случае которой комплекс колокализации, как описано в данном документе, может быть полезен для
- 10 внесения одноцепочных разрывов или разрезания. Нуклеиновые кислоты-мишени включают гены. Для целей настоящего описания, ДНК, такая как двухцепочечная ДНК, может включать нуклеиновую кислоту-мишень, а комплекс колокализации может связываться или иным образом колокализоваться с ДНК или TALEN может в ином случае связываться с ДНК в или в непосредственной близости или вблизи нуклеиновой
- 15 кислоты-мишени таким образом, чтобы комплекс колокализации или TALEN оказывали искомый эффект на нуклеиновую кислоту-мишень. Такие нуклеиновые кислоты-мишени могут включать эндогенную (или естественную) нуклеиновую кислоту и экзогенную (или чужеродную) нуклеиновую кислоту. Специалист на основе настоящего описания легко идентифицирует или разработает направляющие РНК и белки Cas9, которые
- 20 колокализуются с ДНК или TALEN, которые связываются с ДНК, включающей нуклеиновую кислоту-мишень. Специалист также будет способен идентифицировать транскрипционные регуляторные белки или домены, такие как транскрипционные активаторы или транскрипционные репрессоры, которые аналогичным образом колокализуются с ДНК, включающей нуклеиновую кислоту-мишень. ДНК включает
- 25 геномную ДНК, митохондриальную ДНК, вирусную ДНК или экзогенную ДНК. Согласно одному аспекту, материалы и способы, используемые при практическом осуществлении настоящего изобретения, включают те, которые описаны в Di Carlo, et al., Nucleic Acids Research, 2013, vol. 41, No. 7 4336-4343, включенном в данный документ ссылкой в полном объеме и для всех целей, включая типичные штаммы и среды,
- 30 плазмидные конструкции, трансформацию плазмид, электропорацию транзиторной гидРНК-кассеты и донорных нуклеиновых кислот, трансформацию гидРНК-плазмиды с донорной ДНК в Cas9-экспрессирующие клетки, индукцию Cas9 галактозой, определение мишений CRISPR-Cas в геноме дрожжей, и т.д. Дополнительные источники, включая информацию, материалы и методы, полезные для специалиста в данной области
- 35 при осуществлении изобретения, приведены в Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J. E. and Church, G.M. (2013) RNA-Guided human genome engineering via Cas9. Science, 10. 1126/science. 1232033; Storici, F., Durham, C.L., Gordenin, D.A. and Resnick, M.A. (2003) Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. PNAS, 100, 14994-14999 and 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, каждый из которых включен в данное описание ссылкой в полном объеме и для всех целей.

Чужеродные нуклеиновые кислоты (т.е. те, которые не являются частью природной композиции нуклеиновых кислот клетки) могут быть введены в клетку с помощью

- 45 любого способа, известного специалистам в данной области техники для такого введения. Такие способы включают трансфекцию, трансдукцию, вирусную трансдукцию, микроинъекции, липофекцию, нуклеофекцию, бомбардировку наночастицами, трансформацию, коньюгацию и т. п. Специалисту в данной области техники легко

понять и адаптировать такие способы, с помощью легко идентифицируемых литературных источников.

Донорные нуклеиновые кислоты включают любую нуклеиновую кислоту, которая будет вставлена в последовательность нукleinовой кислоты, как описано в данном

5 документе.

Следующие примеры приведены в качестве типичных образцов настоящего описания. Эти примеры не должны быть истолкованы как ограничивающие объем настоящего описания, поскольку они и другие эквивалентные воплощения будут очевидны в свете настоящего описания, чертежей и сопровождающей формулы изобретения.

10 Пример I

Сборка направляющей РНК

19 п.о. выбранной последовательности-мишени (т. е. 5'-N19 5'-N19-NGG-3') были включены в два комплементарных олигонуклеотида 100-мера (TTCTTGGCTTTATAT

ATCTTGTGGAAAGGACGAAACACCGN19GTTTAGAGCTAGAAATAGCAAG

15 ТТААААТАААГГГСТАГТСС). Каждый олигонуклеотид 100-мер суспендировали в концентрации 100 мМ в воде, смешивали с равным объемом и гибридизовали в термоциклере (95°C, 5 мин; линейное уменьшение до 4°C, 0,1°C/сек). Для приготовления вектор назначения, клонирующий вектор с гидРНК (Addgene plasmid ID 41824)

линеаризовали с помощью AfIII и очищали вектор. Реакцию сборки (10 мкл) гидРНК 20 проводили с 10 нг соединившегося 100 п.о. фрагмента, 100 нг каркаса назначения, IX реакционной смеси для сборки Gibson (New England Biolabs) при 50°C в течение 30 мин. Продукт реакция может быть сразу же трансформирован в бактерии для получения колоний отдельных сборок.

Пример II

25 Дизайн и Сборка перекодированных TALE

re-TALEs были оптимизированы на разных уровнях для облечения сборки, а также для улучшения экспрессии. ДНК-последовательности вначале вместе оптимизировали для приближения к частоте использования кодонов у человека, и низкой энергии сворачивания мРНК на 5'-конце (GeneGA, Bioconductor). Полученную

30 последовательность подвергали эволюции в течение нескольких циклов для устраниния повторов (прямых или инвертированных) длиннее, чем 11 п.о. (см. Фиг. 12). В каждом цикле оценивали синонимичные последовательности для каждого повтора. Отбирали те, у которых было наибольшее расстояние Хэмминга к эволюционирующему ДНК. Последовательность одного из re-TALE, обладающего 16,5 мономерами представлена

35 ниже

СТААССССТГААСАГГТАГТСГСТААГСТГАААААТСГГГГГСААГСААГСА
СТТГАГАССГТСААСГАСТССТГССАГТГСТСТГССААГСССАТГГАТГАСТСС
ГАГСААГТСГТСГСГАТСГСГАСГААССГСГГГГГСААГСАГГСГСТГГААААСТГ
ТТСАГАГАСТГСТГСТГАСТТГСТГСАГГСГСАТГГСТСТГССССГАААСАГГТТГ

40 ТСГСААТАГСААГТААТААГГСГГТААГСААГСССТАГАГАСТГСААСГССТГ
СТССССГТСТГТГСТГСАГГСТСАСГГСТСАСГСАААГТГСТГСГСАТАГСС
АГТСАСГАСГГГГААААСАААГСТСАГААААСГГТСАААГГТГСТГСССГТС
ГТГССААГСАСАТГГСТТААСАСССГААААГСТСАААГГТГСТГСССГТС

45 САСГГАСТТАСГССАГАААСАГГТГГСТГСААТГСССТССААСАТСГГСГГГАААСА
АГСГТГГААААСТГТГСАГАГАСТСССТГТГСТГСААГСССАСГГСТГС
ГССТГСАГСАГГТГГССАТСГСТАГССААСГАСГГААГСАГГСТС
ССГТААСАГСАСТСССАГТТГСТГСААГСТСАСГГСТС

TAGTTGCCATAGCAAGCAACGGAGGAGGAAAACAGGCATTAGAAACAGTTCAGCG
 CTTGCTCCCGGTACTCTGTCAGGCACACGGTCTAACTCCGGAACAGGTCGTAGCCAT
 TGCTTCCCCTGATGGCGCAAACAGGCGTAGAGACAGTCCAGAGGCTTGCCTG
 TGTTATGCCAGGCACATGGCCTCACCCCGGAGCAGGTCGTTGCCATGCCAGTAATA
 5 TCGGCAGGAAAGCAAGCTCTCGAAACAGTACAACGGCTGTTGCCAGTCCTATGTCAA
 GCTCATGGACTGACGCCCGAGCAGGTAGTGGCAATCGCATCTCACGATGGAGGTA
 ACAAGCACTCGAGACTGTCCAAGAATTGTTACCCGTACTATGCCAAGCGCATGGTT
 AACCCCCAGAGCAAGTTGTTGCTATTGCATCTAACGGCGTGGCAAACAAAGCCTGG
 AGACAGTGCAACGATTACTGCCTGTCTTATGTCAGGCCATGGCCTACTCCTGAGC
 10 AAGTCGTAGCTATGCCAGCAACATAGGTGGAAACAGGCCCTGGAAACCGTACAA
 CGTCTCCTCCCAGTACTTGTCAAGCACACGGGTTGACACCGGAACAAGTGGTGGC
 GATTGCGTCCAACGGCGGAGGCAAGCAGGCAGTGGAGACCGTCCAACGGCTTCTTC
 CGGTTCTTGCCAGGCTCATGGCCTACGCCAGAGCAGGTGGTAGCAATAGCGTCG
 AACATCGGTGGTAAGCAAGCGCTGAAACGGTCCAGCGTCTGCCCCGGTGGTGC
 15 CAGGCGCACGGACTCACACCAGAACAAAGTGGTGTCTATTGCTAGTAACAACGGTGG
 AAAGCAGGCCCTCGAGACGGTGCAGAGGTTACTTCCCCTGTCAAGCGCACG
 GCCTCACTCCAGAGCAAGTGGTGCATCGCTCAAACAAATGGTGGAAAGACCTGCC
 CTGGAA

- В соответствии с некоторыми аспектами, могут быть использованы TALEs, имеющие,
- 20 по меньшей мере, 80% идентичность последовательности, по меньшей мере, 85% идентичность последовательности, по меньшей мере, 90% идентичность последовательности, по меньшей мере, 95% идентичность последовательности, по меньшей мере, 98% идентичность последовательности, или, по меньшей мере, 99% идентичность последовательности с приведенной выше последовательностью.
- 25 Специалист легко поймет, где указанная последовательность может варьировать при сохранении ДНК-связывающей активности TALE.

Димерные блоки ге-TALE кодирующие два RVDs (см. Фиг. 6А) получали с помощью двух раундов ПЦР в стандартных условиях Кара HIFI (KPAP), при которых первый раунд ПЦР вводил RVD-кодирующую последовательность, а второй раунд ПЦР

30 приводил к образованию целых димерных блоков с 36 п.о. перекрытием с соседними блоками. ПЦР-продукты очищали с помощью QIAquick 96 PCR Purification Kit (QIAGEN), а концентрации измеряли с помощью Nano-drop. Последовательности праймеров и матриц перечислены в таблице 1 и таблице 2 ниже.

Таблица 1. последовательности блоков ге-TALE

Блок 0	CGCAATGCGCTACGGGAGCACCCCTCAACCTAACCCCTGAACAGGTAGTCGCTATAGCTTCANNNNNNGGGGC AAGCAAGCACTTGAGACCGTTCAACGACTCTGCCAGTGCTCTGCCAACGCCATGGATTGACTCCGGAGCAAGTC GTCGCGATCCGAGCNNNNNNGGGGAAACAGGGCTGGAAACTGTTAGCAAGGACTGCTGCTGTACTTTGTCA GCGCATGGTCTC
Блок 1	AGACTGCTGCCGTACTTGTCAAGCGCATGGTCTCACCCCGAACAGGTTGTCGCAATAGCAAGmNNNNNGGCC GTAAGCAAGCCCTAGAGACTGTCAACGCCCTGCTCCCGTGTGTCAGGCTCACGGTCTGACACCTGAACAAG TTGTCGCGATAGCCAGTNNNNNNNGGGGAAACAAAGCTCTAGAAACGGTTCAAAGGTTGTTGCCGTTCTGTG AACGACATGGGTTA
Блок 1'	TGCGCTCACGGGAGCACCCCTCAACCTCACCCCGAACAGGTTGTCGCAATAGCAAGTNNNNNNGGCGGTAAAGC AAGCCCTAGAGACTGTCAACGCCCTGCTCCCGTGTGTCAGGCTCACGGTCTGACACCTGAACAAGTTG CGATAGCCAGTNNNNNNNGGGGAAACAAAGCTCTAGAAACGGTTCAAAGGTTGTTGCCGTTCTGTG CAAGCATGGGTTA
Блок 2	AGGTTGTTGCCGTTCTGTGCAAGCACATGGGTTAACACCCGAACAAGTAGTAGCGATAGCGTCANNNNNNGG GGTAAACAGGCTTGGAGACGGTACAGCGGTTATTGCGCGTCTGCGCAGGCCACGGACTTACGCCAGAACAG GTGGTTGCAATTGCCCTCCNNNNNNGGGGAAACAAAGCGTTGGAAACTGTGCAAGAGACTCCTCCTGTTTG CAAGCCCACGGCTGACGCT
Блок 3	AGACTCCTCCTGTTTGTGCAAGCCCACGGCTTGACGCCAGCAGGTTGTCAGGCCATCGCTAGC ANNNNNNGGGGAAAGCAGGCTTGTGAAACCGTACAGCGACTTCTCCAGTTTGTGCAAGCTCACGGCTAACCCCC GAGCAA GTAGTTGCCATAGCAAGC NNNNNNGGAGGAAACAGG CATTAGAAACAGTT CAGCGCTGCTCCGGTACTCTGT CAGGCACACGGCTA

	Блок 4	CGCTTGCTCCCGTACTCTGTCAGGCACACGGTCTAACCTCCGAACAGGTCGTAGCCATTGCTCCNNNNNGGC GGCAAACAGGCCTAGAGACCGTCCAGAGGCCTTGCCCTGTGTTATGCCAGGCACATGCCCATCCCCGGAGCAG GTCGTTGCCATGCCAGTNNNNNNGCGGAAAGCAAGCTCTCGAAACAGTACAACGGCTTGCAGTCATGT CAAGCTCATGGACTG
5	Блок 5°	CGGCTGTTGCCAGTCCTATGTCAGCTCATGGACTGACGCCCGAGCAGGTAGTGGCAATCGCATTCTNNNNNGGA GGTAACACAAGCACTCGAGACTGTCAAAGATTGTTACCCGTAATGCCAAGCGCATGGTTAACCCCAGAGCAA GTTGTTGGCTATTGCATCTTSnSINNNNNGGCGGAAACAGCTCTGGAGACCGTGCAACGATTACTGCCTGTCTTATG TCAGGCCATGCCCT
	Блок 6	CGATTACTGCCTGTCTTATGTCAGGCCATGCCCTACTCCTGAGCAAGTCTAGCTATGCCAGCN>NNNNNGGT GGCAAGCAAGCACTGGAAACAGTCCAGCGTTGCTTCCAGTACTTGTCAAGCACAGGGTTGACACCGGAACAAG GTGGTGGCTATGCCTCANNNNNGGAGGAAAGCAGGGCTGGAAACCGTCCAACGTCTTACCGGTGTTG CAGGCGCACGGGCTC
10	Блок 6'	CGATTACTGCCTGTCTTATGTCAGGCCATGCCCTACTCCTGAGCAAGTCTAGCTATGCCAGCN>NNNNNGGT GGAAACAGGCCCTGAAACCGTACAACGTCCTCCAGTACTTGTCAAGCACAGGGTTGACACCGGAACAAG TGGTGGCATTGCGTCCNNNNNGGAGGAAAGCAGGGACTGGAGACCGTCCAACGGCTTCTCCGGTTCTTGCC AGGCTCATGGGCTC
	Блок 7	CGGCTTCTCCGGTTCTTGCCAGGCTCATGGCTCACCCAGAGCAGGTGGTAGCAATAGCGTCGN>NNNNNGGT GGTAAGCAAGCGCTGAAACGGTCCAGCGTCTCTGCGGTGTTGCGCAGGCGACGGACTCACACCAGAACAA GTGGTTGCTATTGCTAGTTSnnsnsnSINGGTGAAAGCAGGCCCTGAGACGGTGAGAGGTTACTCCCGTCTCT GTCAAGCGCACGGCCTC

Таблица 2. Последовательности праймеров блоков re-TALE

	Блок O-F	CGCAATGCGCTCACGGGAGCACCCCTCAACCTAACCCCTGAACAGGT*A*T*G
	Блок O-R	GAGACCATGCGCCTGACAAAGTACAGGCAGCAGTCTGAACAG*T*T
	Блок 1'-F	TGGCGCAATGCGCTCACGGGAGCACCCCTCA*A*C
	Блок 1-F	AGACTGCTGCCTGTACTTGTCAAGGCCATGGTCTCACCCCCAAC*A*G*G
20	Блок -R/Блок 1'-R	TAACCCATGTGCTTGGCACAGAACGGGAAACAACCTTGAACCG*T*T
	Блок 2-F	AGGTTGTTGCCCGTTCTGTGCAAGCACATGGTTAACACCCgaac*a*a
	Блок 2-R	AGGCGCTCAAGCCGTGGGCTTGACACAAAACAGGAAGGAGTCTCTGCACAG*T*T
	Блок 3-F	AGACTCCTCCTGTTGTCAAGCCCACGGCTTGACGCCG*A*G
	Блок 3-R	TAGACCGTGTGCCTGACAGAGTACCGGGAGCAAGCGCT*G*A
	Блок 4-F	CGCTTGCTCCGGTACTCTGTCAAGGCACAGGTCTAA*C*T
	Блок 4-R	CAGTCATGAGCTTGACATAGGACTGGCAACAGCCGTT*G*T
	Блок 5-F	CGGCTGTTGCCAGTCCTATGTCAGCTCATGGACTGA*C*G
	Блок 5-R	AAGGCCATGGCCTGACATAAGACAGGCAGTAATCGTT*G*C
	Блок 6-F	CGATTACTGCCTGTCTTATGTCAGGCCATGCCCTTA*C*T
	Блок 6-R	GAGCCCGTGCCTGGCAAAGCACCGGTAAGACGTTGA*C*G
	Блок 6'-F	CGATTACTGCCTGTCTTATGTCAGGCCATGCCCTTAACCTCTGAGCAA
	Блок 6'-R	GAGCCCATGAGCCTGGCAAAGAACCGGAAGAAGCCGTT*G*G
30	Блок 7-F	CGGCTTCTCCGGTTCTTGCCAGGCTCATGGCTCACGCCAGAGCAGG*T*T
	Блок 7-R	GAGGCCGTGCGCTTGACAGAGGACGGGAAGTAACCTCT*G*C

Re-TALENs и векторы назначения re-TALE-TF были сконструированы путем модификации TALE-TF и клонирующего каркаса TALEN (см. источник 24, включенный в данный документ ссылкой в полном объеме). 0,5 RVD-области на векторах были повторно закодированы, а в назначенный участок для клонирования re-TALE был включен сайт разрезания Sapi. Последовательности re-TALENs и каркасы re-TALE-TF представлены на Фиг. 20. Плазмиды могут быть предварительно обработаны с помощью Sapi (New England Biolabs) в рекомендованных производителем условиях, и очищены с помощью набора для очистки ПЦР QIAquick (QIAGEN).

Однореакторную (10 мкл) реакцию сборки TASA проводили с 200 нг каждого блока, 500 нг каркаса назначения, IX ферментной смеси TASA (2 Ед. Sapi, 100 Ед. амплигазы (Epicentre), 2,5 Ед. ДНК-полимеразы (New England Biolabs)) и IX буфер для изотермической реакции сборки, как описано ранее (см. источник 25, включенный в данный документ ссылкой в полном объеме) (5% PEG-8000, 100 мМ Трис-HCl, pH 7,5, 10 мМ MgCl₂, 10 мМ DTT, 0,2 мМ каждого из четырех dNTP и 1 мМ NAD).

Инкубирование проводили при 37°C в течение 5 мин и 50°C в течение 30 мин. Реакция может быть подвергнута непосредственно бактериальной трансформации для получения

колоний отдельных сборок. При таком подходе эффективность получения полноразмерного конструкции составила -20%. В ином случае эффективность > 90% может быть достигнута трехступенчатой сборкой. Во-первых, 10 мкл реакции сборки re-TALE осуществляли с 200 нг каждого блока, 1X ферментной смеси re-TALE, (100 Ед. амплигазы, 12,5 мЕд. T5-экзонуклеазы, 2,5 Ед. ДНК-полимеразы Phusion) и IX изотермического буфера для сборки при 50°C в течение 30 мин, а затем проводили стандартизированные реакции ПЦР Кара HIFI, электрофореза в агарозном геле и экстракции из геля QIAquick Gel extraction (Qiagen) для обогащения полноразмерных re-TALEs. 200 нг ампликонов re-TALE затем могут быть смешаны с 500 нг каркаса назначения, предварительно обработанного SapI, 1X изотермического реакционного буфера для сборки re-TALE и инкубировали при 50°C в течение 30 мин. Конечная реакция сборки re-TALE может быть подвергнута непосредственно бактериальной трансформации для получения колоний отдельных сборок. Специалист в данной области техники легко сможет выбрать эндонуклеазы, экзонуклеазы, полимеразы и лигазы из числа тех, которые известны для практического осуществления способов, описанных в данном документе. Например, можно использовать эндонуклеазы типа II_s, такие как: Fok 1, Bts I, Ear I, Sap I. Могут быть использованы титруемые экзонуклеазы, такие как лямбда-экзонуклеаза, T5-экзонуклеаза и экзонуклеаза III. Могут быть использованы полимеразы, применяемые без горячего старта, такие как, ДНК-полимераза Phusion, ДНК-полимераза Taq и ДНК-полимераза VentR. В данной реакции могут быть использованы термостойкие лигазы, такие как амплигаза, ДНК-лигаза rfu, ДНК-лигаза Taq. Кроме того, в зависимости от конкретного вида используемых ферментов для активации таких эндонуклеаз, экзонуклеаз, полимераз и лигаз могут быть использованы различные условия реакции.

25 Пример III

Клеточная линия и Клеточная культура

Клетки PGP1 iPS поддерживали в покрытых матригелем (BD Biosciences) планшетах в mTeSR1 (StemCell Technologies). Культуры пересевали каждые 5-7 дней с TrypLE Express (Invitrogen). Клетки 293T и 293FT выращивали и поддерживали в модифицированной 30 Дульбекко среде Игла (DMEM, Invitrogen) с высоким содержанием глюкозы с добавлением 10% фетальной бычьей сыворотки (FBS, Invitrogen), пенициллина/стрептомицина (pen/strep, Invitrogen), и незаменимых аминокислот (NEAA, Invitrogen). Клетки K562 выращивали и поддерживали в среде RPMI (Invitrogen) с добавлением 10% фетальной бычьей сыворотки (FBS, Invitrogen 15%) и пенициллина/стрептомицина (pen/strep, Invitrogen). Все клетки поддерживали при 37°C и 5% CO₂ в увлажненном 35 инкубаторе.

Была создана стабильная клеточная линия 293T для детекции эффективности HDR, как описано в источнике 26, включенном в данное описание ссылкой в полном объеме. В частности, репортерную клеточную линию, несущую интегрированные в геном 40 последовательности, кодирующие GFP, разрушали вставкой стоп-кодона и 68 п.о. геномного фрагмента, полученного из локуса AAVS1.

Пример IV

Тест активности re-TALENs

Репортерные клетки 293T высевали при плотности 2 x 10⁵ клеток на лунку в 24-45 луночный планшет и трансфецировали их 1 мкг каждой плазмида re-TALENs и 2 мкг плазмида с донорной ДНК с использованием Lipofectamine 2000 согласно протоколам производителя. Клетки собирали с помощью TrypLE Express (Invitrogen) через ~ 18 ч после трансфекции и ресуспендировали в 200 мкл среды для анализа проточной

цитометрией с использованием анализатора клеток LSRFortessa (BD Biosciences). Данные проточной цитометрии анализировали с использованием FlowJo (FlowJo). По меньшей мере, 25000 событий анализировали для каждого трансфекционного образца. Для экспериментов по направленному воздействию на эндогенный локус AAVS1 в клетках 5 293T, процедуры трансфекции были идентичны описанным выше, и отбор на пуромицине был проведен с концентрацией лекарственного средства 3 мкг/мл через 1 неделю после трансфекции.

Пример V

Оценка образования функциональных лентивирусов

- 10 Лентивирусные векторы были созданы с помощью стандартных методов ПЦР и клонирования. Лентивирусные плазмида трансфектировали Lipofectamine 2000 с помощью лентивирусной упаковочной смеси (Invitrogen) в культивируемые клетки 293FT (Invitrogen), с целью получения лентивируса. Надосадочную жидкость собирали через 48 и 72 ч после трансфекции, стерильно фильтровали, и 100 мкл фильтрованной 15 надосадочной жидкости добавляли к 5 x 10⁵ свежих клеток 293T с полибреном. Титр лентивируса рассчитывали на основании следующей формулы: титр вируса = (процент клеток GFP+ 293T * исходное количество клеток при трансдукции)/(объем исходной вирус-содержащей надосадочной жидкости, используемой в эксперименте трансдукции). Чтобы проверить работоспособность лентивирусов, через 3 дня после трансдукции, 20 лентивирус-трансдектированные клетки 293T трансфектировали 30 нг плазмида, несущей репортер mCherry и 500 нг плазмида pUC19 с помощью Lipofectamine 2000 (Invitrogen). Изображения клеток анализировали с помощью Axio Observer Z1 (Zeiss) через 18 часов после трансфекции, после чего клетки собирали с помощью TrypLE Express (Invitrogen) и ресуспендировали в 200 мкл среды для анализа проточной цитометрией с 25 использованием анализатора клеток LSRFortessa (BD Biosciences). Данные проточной цитометрии анализировали с использованием BD FACSDiva (BD Biosciences).

Пример VI

- Тест эффективности редактирования генома с помощью re-TALENs и Cas9-гидРНК
PGP1 iPSCs культивировали с ингибитором Rho-киназы (ROCK) Y-27632 (Calbiochem)
30 за 2 ч перед нуклеофекцией. Трансфекцию проводили с помощью P3 Primary Cell 4D-Nucleofector X Kit (Lonza). В частности, клетки собирали с помощью TrypLE Express (Invitrogen) и ресуспендировали 2x10⁶ клеток в 20 мкл смеси для нуклеофекции, содержащей 16,4 мкл раствора P3 Nucleofector, 3,6 мкл добавки, 1 мкг каждой плазмиды re-TALENs или 1 мкг Cas9 и 1 мкг гидРНК-конструкции, 2мкл 100 мкМ ssODN. Затем 35 смеси переносили в 20мкл полоски Nucleocuvette и проводили нуклеофекцию с помощью программы CB150. Клетки высевали в планшеты, покрытые матригелем в среде mTeSR1 среде с добавлением ингибитора ROCK в течение первых 24 часов. Для эксперимента направленного воздействия на эндогенный локус AAVS1 с помощью донорной двухцепочечной ДНК, проводили такую же процедуру за исключением того, что было 40 использовано 2 мкг донорной двухцепочечной ДНК, а среда mTeSR1 была дополнена пуромицином в концентрации 0,5 мкг/мл через 1 неделю после трансфекции.

Информация о reTALENs, гидРНК и ssODNs, используемых в данном примере, приведена в Таблице 3 и Таблице 4 ниже.

Таблица 3. Информация о парах re-TALEN/Cas9-гидРНК, нацеленных на CCR5

# Целевого участка	re-TALENs целевой участок пары (на- пары)	re-TALENs целевой участок пары (ко- пары)	re-TALEN-L целевая последователь- ность	re-TALEN-R целевая последователь- ность	гидРНК целевая последователь- ность	гидРНК на- чальная пози- ция целевой по- следовательно- сти

	чало)	нен)					
	/chr3:	/chr3:					
5	1	4640 9942	4640 9993	TCCCCACTTCTTGTGAA	TAACCACTCAGGACAGGG	CACTTCTTGTGAATCCTT	46409946
	2	4641 0227	4641 0278	TCACACAGCAAGTCAGCA	TAGCGGAGCAGGCTCGGA	TGGGCTAGCGGAGCAGGCT	4641 0264
	3	4641 1260	4641 1311	TACCCAGACGAGAAAGCT	TCAGACTGCCAAGCTGA	ACCCAGACGAGAAAGCTGA	4641 1261
	4	4641 1464	4641 1515	TCTTGTGGCTCGGGAGTA	TATTGTCAAGCAGAGCTGA	AGAGGGCATCTTGTGGCTC	4641 1456
	5	4641 1517	4641 1568	TTGAGATTTCAGATGTC	TATACAGTCATATCAAGC	ATCAAGCTCTTGGCGGT	4641 1538
10	6	4641 1634	4641 1685	TTCAGATAGATTATATCT	TGCCAGATACTAGGTGG	GCTTCAGATAGATTATATC	4641 1632
	7	4641	4641	TTATACTGTCT	TCAGCTCTTCT	ACGGATGTCTCA	4641
		2396	2447	ATATGAT	GGCCAGA	GCTCTTC	2437
	8	4641 2432	4641 2483	TGGCCAGAAGAGCTGAGA	TTACCGGGAGAGTTCT	CCGGGGAGAGTTCTTGT	4641 2461
	9	4641 2750	4641 2801	TTTGCAGAGAG ATGAGTC	TTAGCAGAA TAAGATT	GAAATCTTATCT TCTGCTA	4641 2782
15	10	4641 3152	4641 3203	TATAAGACTAA ACTACCC	TCGTCTGCCAC CACAGAT	AATGCATGACAT TCATCTG	4641 3172
	11	4641 4305	4641 4356	TAAAACAGTTT GCATTCA	TATAAAAGTCCT AGAATGT	AACAGTTGCAT TCATGGA	4641 4308
	12	4641 4608	4641 4659	TGCCCATCTCT GACCTGT	TAGTGAGCCC GAAGGGG	CCAGAAGGGGA CAGTAAGA	4641 4632
	13	4641 4768	4641 4820	TAGGTACCTGG CTGTCGT	TGACCGTCCTG GCTTTA	CTGACAATCGAT AGGTACC	4641 4757
	14	4641 5017	4641 5068	TGTCACTGGTCA TCTGCTA	TCGACACCGAA GCAGAGT	ACACCGAAGCA	4641 5046
20	15	4642 0034	4642 0084	TGCCCCCGCGA GGCCACA	TCTGGAAGTTG AACACCC	GGAAGTTGAAC ACCCTTGC	4642 0062
	25	Таблица 4. Дизайн ssODN для изучения ssODN-опосредованного редактирования генома					

30	Фигу- ра 3b	Дистанция между вто- рой мутаци- ей и DSB	90-*1	CTACTGTCAATTCAAGGGCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTCAAGCT TGGCAGTCTGACTACAGAGGCCACTGGCTT
			90-*2	CTACTGTCAATTCAAGGCCAATACCCCTAACGAGAAAAGCTGAGGGTATAACAGGTTCAAGCTT GGCAGTCTGACTACAGAGGCCACTGGCTT
			90-*3	CTACTGTCAATTCAAGGCCAATACCCAGACGAGAAAAGCTGAGGGTATAACAGGTTCAAGCT TGGCAGTCTGACTACAGAGGCCACTGGCTT
			90M-0	CTACTGTCAATTCAAGGCCAATACCCAGACGAGAAAAGCTGAGGGTATAACAGGTTCAAGCTT GGCAGTCTGACTACAGAGGCCACTGGCTT
			90-*4	CTACTGTCAATTCAAGGCCAATACCCAGACGAGAAAAGCTGAGGGTATAACAGGTTGTAGCTT GGCAGTCTGACTACAGAGGCCACTGGCTT
			90-*5	CTACTGTCAATTCAAGGCCAATACCCAGACGAGAAAAGCTGAGGGTATAACAGGTTCAAGCTT GGCTCTGACTACAGAGGCCACTGGCTT
			90-*6	CTACTGTCAATTCAAGGCCAATACCCAGACGAGAAAAGCTGAGGGTATAACAGGTTCAAGCTT GGCAGTCTGACTAGTGAGGCCACTGGCTT
40	Фигу- ра 3c	Дистанция между ssODN и DSB	L670 bp_9 OM	CACTTATATTCCCTGCTTAAACAGTCCCCCGAGGGTGGGTGCGGAAAGGCTCTACACT TGTTATCATTCCCTCTCCACCACAGGCAT
			L570 bp_9 OM	TAAGGCACAGAGCTCAATAATTGGTCAGAGCCAAGTAGCAG
			L480 bp_9 OM	GGAGGTTAAACCCAGCAGCATGACTGCAGTTCTTAATCAATGCCCCCTGAATTGCACATAT GGGATGAAGTAACTGACAT
			L394 bp_9 OM	CTCGATGATTGCTGCTCTGTTATGATTATGTTACTGAGCTACTGTAGCACAGACATAT GTCCTATATGGGGCGGGGGTGGGGGT
			L290 bp_9 OM	GGTGTCTTGATCGCTGGCTATTCTATACTGTTCTGGTTTCGGAAGCAGTCATTCTTT CTATTCTCAAGCACAGCAATTAGCTT
			L200 bp_9 OM	GCTTCTAGTTGCTGAAACTAATCTGCTATAGACAGAGACTCCGACGAACCAATTATTAA GGATTTGATCAAATAACTCTCTGACAA
45			L114 bp_9 OM	GAAAGAGTAACTAAGAGTTTGATGTTACTGAGTGCAATAGTATGCACTAGATGCTGCCGT GGATGCCTCATAGAATCCTCCCAACAAC

		L45b p_90 М	GCTAGATGCTGGCCGTGGATGCCTCATAGAATCCTCCCAACAACCGATGAAATGACTACTG TCATTCAAGCCCCAATACCCAGACGAGAAAG
5		R40b p_90 М	ACAGGTTCAAGCTGGCAGTCTGACTACAGAGGCCACTGGCTTACCCCTGGGTTAGTCT GCCTCTGAGGATTGGGGCACGTAATTGCTGTTAAG
		R100 bp_90 М	TTAGTCTGCCCTGTAGGATTGGGGCACGTAATTGCTGTTAAG
		R200 bp_90 М	GGAAGCCCAGAGGCCATTTGTGGCTGGGAGTAGCTCTGCTACCTCTCAGCTGCT GACAATACCTGAGATTTCAGATGTCACC
		R261 bp_90 М	TCAGCTCTGCTGACAATACTTGAGATTTCAGATGTCACCAACCAGCAAGAGAGCTTGATA TGACTGTATATAGTATAGTCATAAAGAAC
10		R322 bp_90 М	CATAAAGAACCTGAACCTGACCATATACTTATGTCATGTGAAATCTCTCATAGCTTCAG ATAGATTATATCTGGAGTGAAGAACCTG
		R375 M_90 М	GTGGAAAATTCTCATAGCTTCAGATAGATTATATCTGGAGTGGAGCAATCCTGCCACCTAT GTATCTGGCATAGTGTAGTCCTCATAAA
		R448 bp_90 М	GGTTTGAAGGGCAACAAAATAGTGAACAGAGTGAAATCCCCACCTAGATCCTGGTCCA GAAAAAGATGGGAAACCTGTTAGCTCACC
		Комплементар- ный-30-мер	GGCCACTAGGGACAAAATTGGTGAacagaaa
15	Фигу- ра 3d	Комплементар- ный-50-мер	CCCACAGTGGGCCACTAGGGACAAAATTGGTGAacagaaaagccccatcc
		Комплементар- ный-70-мер	TCCCCCTCCACCCACAGTGGGCCACTAGGGACAAAATTGGTGAacag aaaagccccatcccttaggcetcc
		Комплементар- ный-90-мер	cttTTATCTGTCCCCCTCCACCCACAGTGGGCCACTAGGGACAAAAT TGGTGAacagaaaagcccc atcccttaggcetccctcccttag
		Длина и ориентация ssODN для направленно- го воздей- ствия Cas9- гидРНК	gttctgggtacttTTATCTGTCCCCCTCCACCCACAGTGGGCCACTAGGG A AAAATTGGTGAacag aaaagccccatcccttaggcetccctcccttagtctctgata
20		Комплементар- ный-110-мер	TTTCTGTCACCAATGGTGTCCCTAGTGGCC
		Некомплемен- тарный - 30-Мер	GGATGGGCTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGG
		Некомплемен- тарный -50-Мер	GGAGGCCTAAGGATGGGCTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGG TGGAGGGGA
25		Некомплемен- тарный -70-Мер	CTAGGAAGGAGGAGGCCATAAGGATGGGCTTTCTGTCACCAATGGTGTCCCTAGTGGCCC CACTGTGGGTGGAGGGGACAGATAAAAG
		Некомплемен- тарный-90-мер	TATCAGGAGACTAGGAAGGAGGAGGCCATAAGGATGGGCTTTCTGTCACCAATGGTGT CCTAGTGGCCCCACTGTGGGTGGAGGGGACAGATAAAAGTACCCAGAAC
30	Фигу- ра 2c	Cas9 гидРНК- CCR 5-1	TTCTAGTAACCACTCAGGACAGGGGGTTCAGCCAAAAATTACA AGAAAGTGGGACCCATGGAAAT
		Cas9 гидРНК- CCR 5-2	CAGCAAGTCAGCACAGCGTGTGACTCCGAGGGTGTCCGCT AGCCCCACATTGCCCTCTGGGGGT
		Cas9 гидРНК- CCR 5-3	GTCAGACTGCCAAGCTGAAACCTGTCTTACCCCTACTTCTCGTC TGGGTATTGGGTGAATGACAGT
		Cas9 гидРНК- CCR 5-4	CAGAGCTGAGAAGACAGCAGAGAGCTACTCCCGAAGCACAAGATG CCCTCTGGCTCCGTGACCTTGCG
		Cas9 гидРНК- CCR 5-5	CTGACAATACTTGAGATTTCAGATGTCACCAACGACCAAGAGAGC TTGATATGACTGTATAGTATAG
		Cas9 гидРНК- CCR 5-6	CAGATACATAGGTGGCAGGATTCTCACTCCAGACTTAATCTATCT GAAGCTATGAGAAATTTCACAT
35		Cas9 гидРНК- CCR 5-7	TATATGATTGATTGACAGCTCATCTGGCCAGATAAGCTGAGACA TCCGTTCCCTACAAGAAACTCTC
		Cas9 гидРНК- CCR 5-8	ATCTGGCCAGAAGAGCTGAGACATCCGTTCCCTGAAGAAA ACTCTCCCCGTAAGTAACC TCTCAGCTG
		Cas9 гидРНК- CCR 5-9	AGGCATCTCACTGGAGAGGGTTAGTTCTCTTAAGAGAAGATAAGATT CAAGAGGGAA GCTAAGACTC
40		Cas9 гидРНК- CCR 5-10	ATAATATAATAAAAAATGTTCTGTCTGCCACCAATA GAATGTCATGCATTCTGGTAGT TTAGTCTTA
		Cas9 гидРНК- CCR 5-11	TTTATAAAAGTCTAGAATGTTAGTTGCCCTCGTTGAATGCAA ACTGTTTATACATCAA TAGGTTT
		Cas9 гидРНК- CCR 5-12	GCTCAACCTGGCCATCTGACCTGTTCTCCACTGTCCCCTCTGG GCTACTATG CTGCCGCC
		Cas9 гидРНК- CCR 5-13	TTTTAAAGCAAACACAGCATGGACAGCCAGGCTCTATGATT GTCAGGAGGATGAT GAAGAAAGATT
		Cas9 гидРНК- CCR 5-14	GCTTGTCACTGGTCACTGCTACTCGGGAA CTCTTAATTACTCTGCTTCGGTGTGAAATGAG AAGAAAGAGG
45		Cas9 гидРНК- CCR 5-15	ATACTGCCCGCGAGGCCACATTGGCA AAACCCAGCTTGGGTGTTCAACTTCCAGACTTGGC CATGGAGAA
	Донорный ssODN для направлен-	reTALEN-CCR 5- 1	CTGAAGAATTCCCAGGGTCCCCACTTCTGTGAATCCTGGAGTGAACCCCCCTGT CCTGAGTGGTACTAGAACACACCTCTGGAC

		reTALEN-CCR 5-2	TGGAAGTATCTGCCGAGGTACACAGCAAGTCAGCAGCACAGCCAGTGTGACTCCGAGCCTGCTCCGCTAGCCACATTGCCCTCTGGG
НОГО ВОЗДЕЙСТВИЯ RETALENs на CC R5	reTALEN-CCR 5-3	CTACTGTCATT CAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTCAAGCTT GGCAGTCTGACTACAGAGGCCACTGGCTT	
	reTALEN-CCR 5-4	GGAAGCCCAGAGGGCATCTTGTGGCTGGAGTAGCTCTGCTACCTCTCAGCTCTGCTGACAATACTTGAGATTTCAGATGTCACC	
	reTALEN-CCR 5-5	TCAGCTCTGCTGACAATACTTGAGATTTCAGATGTCACCAACGCCAAGAGAGCTTGATA TGACTGTATAGTATAGTCATAAAGAAC	
	reTALEN-CCR 5-6	GTGGAAAATTCTCATAGCTCAGATAGATTATATCTGGAGTGAGCAATCCTGCCACCTAT GTATCTGGCATAGTGTGAGTCCTCATAAA	
	reTALEN-CCR 5-7	GAAACAGCATTCCACTTTTAACTGTCTATGATTGATTGGTCAGCTCATCTGCCAG AAGAGCTGAGACATCCGTCCCCCTACAA	
НОГО ВОЗДЕЙСТВИЯ RETALENs на CC R5	reTALEN-CCR 5-8	TTGATTGACAGCTCATCTGCCAGAAGAGCTGAGACATCCGTATCCCTACAAGAAACTCTCCCCGTTAAGTAACCTCTCAGCTGCTT	
	reTALEN-CCR 5-9	GGAGAGGGTTAGTTCTCCTTAGCAGAAGATAAGATTCTAAGATGAGAGCTAAGACTCAT CTCTCTGCAAATCTTCTTTGAGAGGTAA	
	reTALEN-CCR 5-10	TAATATAATAAAAAATGTTCCGTGCCACACAGATGAATGTCGAGCATTCTGGGTAGTT TAGTCTTATAACCAGCTGTCTTGCCTAGT	
	reTALEN-CCR 5-11	TTAAAAAACCTATTGATGTATAAAACAGTTGCATTCATGGAGGGTACTAAATACATTCTA GGACTTTATAAAAGATCACTTTTATTAA	
	reTALEN-CCR 5-12	GACATCTACCTGCTCACCTGGCCATCTGACCTGTTTCTTATTACTGCCCCCTCTGG GCTCACTATGCTGCCGCCAGTGGGAC	
	reTALEN-CCR 5-13	TCATCCTCCTGACAATCGATAGGTACCTGGCTGCGTCCATGCTACGTTGCTTAAAGCC AGGACGGTACCTTGGGGTGGTACAA	
	reTALEN-CCR 5-14	GGCTGGCCTGCCGCTGTTGTCATGGTCATCTGCTACTCGGGAGACCTAAAACCTCTGCTCGGTGTCGAAATGAGAAGAAGAGGCACA	
	reTALEN-CCR 5-15	GGCAAGCCTGGGTCAACTGCCCGCGAGGCCACATTGCGAAGTCAGCAAGGGTGTCAACTTCCAGACTTGCCCATGGAGAAGACAT	

Пример VII

Получение библиотеки ампликонов целевых участков

Собирали клетки через 6 дней после нуклеоинфекции и 0,1 мкл фермента-тканевой протеазы prepGEM (ZyGEM) и 1 мкл буфера prepGEM gold (ZyGEM) добавляли к 8,9 мкл 2~5 X 105 клеток в среде, затем по 1 мкл реакций добавляли в 9 мкл ПЦР-смеси, содержащей 5 мкл готовой смеси 2X KAPA HiFi Hotstart Readymix (KAPA Biosystems) и 100 нМ соответствующих пар праймеров для амплификации. Реакционные смеси инкубировали при 95°C в течение 5 мин с последующими 15 циклами 98°C, 20 с; 65°C, 20 с и 72°C, 20 с. Для добавления адаптерной последовательности Illumina, 5 мкл продуктов реакции затем добавляли в 20 мкл ПЦР-смеси, содержащей 12,5 мкл готовой смеси 2X KAPA HiFi Hotstart Readymix (KAPA Biosystems) и 200 нМ праймеров, несущих адаптерные последовательности Illumina. Реакционные смеси инкубировали при 95°C в течение 5 мин с последующими 25 циклами 98°C, 20 с; 65°C, 20 с и 72°C, 20 с. ПЦР-продукты очищали с помощью набора для очистки ПЦР QIAquick, смешивали приблизительно в одинаковых концентрациях, и секвенировали с помощью персонального секвенатора MiSeq. ПЦР-праймеры перечислены ниже в Таблице 5.

Таблица 5 Последовательности ПЦР-праймеров для участка направленного воздействия в CCR5

	# участка направ-ленного воздей-ствия в CCR5	название	последовательность праймера
	◦	Site1-F1	ACACTCTTCCCTACACGACGCTCTCCGATCTCGTGA TTT GCAGTGTGCGTTACTCC
	◦	Site1-F2	ACACTCTTCCCTACACGACGCTCTCCGATCA T CGT TT GCAGTGTGCGTTACTCC
	1	Site1-F3	ACACTCTTCCCTACACGACGCTCTCCGATCTGCC A TTT GCAGTGTGCGTTACTCC
	◦	Site1-F4	ACACTCTTCCCTACACGACGCTCTCCGATCTGGTC ATT GCAGTGTGCGTTACTCC CTGGCATTCC T GCTAACGCTCTCCGATCTCAAGCAACTAAGTCACAGCA
	◦	Site1-R	
	◦	◦	
	◦	Site2-F1	ACACTCTTCCCTACACGACGCTCTCCGATCTCGTGA TG AGGAAATGGAAGCTTG
	◦	Site2-F2	ACACTCTTCCCTACACGACGCTCTCCGATCA T CGAT GAGGAAATGGAAGCTTG

	2	Site2-F3	ACACTCTTCCCTACACGACGCTTCCGATCTGCCTAAAT GAGGAAATGGAAGCTTG
◦	Site2-F4	ACACTCTTCCCTACACGACGCTTCCGATCTGGTCAATG AGGAAATGGAAGCTTG	
◦	Site2-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTCATAGGG TATTGGAGGA	
◦	site3-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATAAT CCTCCAACAACCATCAT	
◦	site3-F2	ACACTCTTCCCTACACGACGCTTCCGATCATCGAA CCTCCAACAACCATCAT	
5	3	site3-F3	ACACTCTTCCCTACACGACGCTTCCGATCTGCCTAAA CCTCCAACAACCATCAT
◦	site3-F4	ACACTCTTCCCTACACGACGCTTCCGATCTGGTCAAT CCTCCAACAACCATCAT	
◦	site3_R	CTCGGCATTCTGCTGAACCGCTTCCGATCTCCAATCCT ACAGAGGCAG	
◦	site4-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATAA	
◦	site4-F2	ACACTCTTCCCTACACGACGCTTCCGATCATCGAA	
10	4	site4-F3	ACACTCTTCCCTACACGACGCTTCCGATCTGCCTAAA
◦	site4-F4	ACACTCTTCCCTACACGACGCTTCCGATCTGGTCAAA	
◦	site4_R	ACACTCTTCCCTACACGACGCTTCCGATCTAAGCAA GCTTTTATTCT	
◦	Site5-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATACT TTGTGGCTCGGGAGTAG	
◦	Site5-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGATC TTGTGGCTCGGGAGTAG	
◦	Site5-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTGGCAGGA TTCTTCACTCCA	
15	◦	site6-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATACTA TTTTGTGCCCTCAAA
6	site6-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGCTA TTTTGTGCCCTCAAA	
◦	site6-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTAACCTGAA CTTGACCATAACT	
◦	site7-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATACTA GCTGAGAGGTTACTTACC	
7	site7-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGCA GCTGAGAGGTTACTTACC	
◦	site7-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTAATGATTTA ACTCCACCCCTC	
◦	site8-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATACT CCACCCCTCCTCAAAAGA	
8	site8-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGACT CCACCCCTCCTCAAAAGA	
◦	site8-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTGGTTTG CCAAATGTCT	
◦	site9_F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATGG GCACATATTAGAAGGA	
9	site9_F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGGG GCACATATTAGAAGGA	
◦	site9_R	CTCGGCATTCTGCTGAACCGCTTCCGATCTAGTAAAG ACCTTAAAGGGAGCA	
◦	site10-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATCAC AATTAAGAGTTGTCTA	
10	site10-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGCA CAATTAAGAGTTGTCTA	
◦	site10-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTCTCAGCTA GAGCAGCTGAAC	
11	site11-F1	CTCGGCATTCTGCTGAACCGCTTCCGATCTGACACTTG ATAATCCATC	
	site11-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGTC	
		ATGTAGACATCTATGTAG	
30		site11-R	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATTCA ATGTAGACATCTATGTAG
	site12-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATACT GCAAAAGGCTGAAGAGC	
	site12-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGACT GCAAAAGGCTGAAGAGC	
12	site12-F3	ACACTCTTCCCTACACGACGCTTCCGATCTGCCAAACT GCAAAAGGCTGAAGAGC	
	site12-F4	ACACTCTTCCCTACACGACGCTTCCGATCTGGTCAACT GCAAAAGGCTGAAGAGC	
	site12-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTGCCATAA AATAGAGCCCTGTCAA	
35		site13-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATCTC TATTTTATAGGCTTCTC
13	site13-F2	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGCTC TATTTTATAGGCTTCTC	
	site13-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTAGCCACCA CCCAAGTGATC	
	site14-F1	ACACTCTTCCCTACACGACGCTTCCGATCTACATCGTC CAGACATTAAGATAGTC	
14	site14-F2	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATTTC CAGACATTAAGATAGTC	
	site14-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTACATGA TGGTGAAGATAAG	
40		site15-F1	ACACTCTTCCCTACACGACGCTTCCGATCTCGTATCCG GCAGAGACAAACATTA
15	site15-F2	ACACTCTTCCCTACACGACGCTTCCGATCTCCGGAGA GACAAACATTA	
	site15-R	CTCGGCATTCTGCTGAACCGCTTCCGATCTAGCTAGGA AGCCATGGCAAG	
45	адаптер Illumina	PE-PCR-F	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACA cgac*g*c
	PE-PCR-R	CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCT GCTGAACc*g*c	
		Мультиплексные силенсные праймеры для ПЦР	
3	site3-MF	ACACTCTTCCCTACACGACGCTTCCGATCTAGTGCATA GTATGTGCTAGATGCTG	
◦	site3-MR	GTGACTGGAGTTCAACGCGTGTGCTTCCGATCTGATCTC TAAGAAGGCAAATGAGAC	
Illumina адаптер	Index-PCR	CAAGCAGAAGACGGCATAACGAGATN1N2N3N4N5N6GTGAC TGGAGTTCAACGCGTGTGCTTCCGATCT	

universal-PCR	AATGATAACGGCGACCACCGAGATCTACACTCTTCCCTACA CGACGCTCTCCGATCT
---------------	--

* index-PCR - праймеры приобретены в Epicentre (праймеры ScriptSeq™ Index PCR Primers)

Пример VIII

Система оценки редактирования генома (GEAS)

Для обнаружения редких геномных изменений использовали секвенирование нового поколения. См. источники 27-30, включенные в данный документ ссылкой в полном объеме. Для того, чтобы сделать возможным широкое применение данного подхода для быстрой оценки эффективности HDR и NHEJ в hiPSCS, было создано программное обеспечение, названное «pipeline», для анализа данных инженерии генома. «Pipeline» интегрирована в одиночный модуль Unix, который использует различные инструменты, такие как R, BLAT и FASTX Toolkit.

Разбивка по штрих-коду: Группы образцов объединяли вместе и секвенировали по технологии спаренных концов с длиной прочтения 150 п.о. на секвенаторе MiSeq (PE150) (Illumina Next Gen Sequencing), а далее разделяли на основе ДНК-штрихкодирования с помощью пакета программ FASTX.

Качество фильтрации: Нуклеотиды с низким качеством последовательности (балл Phred <20) были обрезаны. После обрезки, «риды» короче 80 нуклеотидов были отброшены.

Картирование: BLAT использовали для независимого картирования «ридов», полученных секвенированием спаренных концов, на эталонном геноме и на выходе получали файлы *.psl.

Определение вставок/делеций: вставки/делеции были определены как полноразмерные «риды», содержащие 2 блока совпадений при выравнивании. В расчет принимались только «риды», соответствующие этому паттерну в обоих «ридах» спаренных концов. В качестве контроля качества требовались «риды» вставок/делеций, которые обладали минимальным 70 нуклеотидным совпадением с эталонным геномом, а оба блока должны иметь длину, по меньшей мере, 20 нуклеотидов. Размер и положение вставок/делеций рассчитывали по положению каждого блока в эталонном геноме. Негомологичное соединение концов (NHEJ) оценивали как процент «ридов», содержащих вставки/делеции (см. уравнение 1 ниже). Большинство событий NHEJ были обнаружены в непосредственной близости от участка направленного воздействия.

Эффективность рекомбинации, направляемой гомологией (HDR): Сравнение с паттерном (grep) в окне 12 п.о. центрированном по DSB, использовали для подсчета специфических сигнатур, соответствующих «ридам», содержащим эталонную последовательность, модификации эталонной последовательности (предполагаемые 2 п.о. несовпадения), и «риды», содержащие только 1 п.о. мутацию в пределах предполагаемых 2 п.о. несовпадений (2 п.о. предполагаемые несовпадения) (см. уравнение 1 ниже).

Уравнение 1. Оценка NHEJ и HDR

A = риды идентичные эталону: XXXXXABXXXXX

B = риды, содержащие 2 п.о. несовпадения, которые кодируются ssODN:
XXXXXabXXXXX

C = риды, содержащие только 1 п.о. мутацию в участке-мишени: например,

XXXXXaBXXXXX или XXXXXAbXXXXX

D = риды, содержащие вставки/делеции, как описано выше

$$\text{Эффективность NHEJ} = \left(\frac{B}{A + B + C + D} \right) \times 100\%$$

$$\text{Эффективность HDR} = \left(\frac{B}{A + B + C + D} \right) \times 100\%$$

⁵ Пример IX

Скрининг Генотипа колонизированных hiPSCs

Человеческие клетки iPSC на безфидерных культурах предварительно обрабатывали средой mTesi-1, обогащенной SMC4 (5 мкМ тиазовивин, 1 мкМ CHIR99021, 0,4 мкМ

¹⁰ PD0325901, 2 мкМ SB431542) (см. источник 23, включенный в данный документ ссылкой в полном объеме), по меньшей мере, за 2 часа до сортировки FACS. Культуры диссоциировали с помощью аккутазы (Millipore) и ресуспендировали в среде mTesi-1, обогащенной SMC4 и красителем ToPro-3 для проверки жизнеспособности (Invitrogen) в концентрации 1-2 X10⁷/мл. Живые клетки hiPS сортировали по одной клетке с помощью BD FACSAria II SORP UV (BD Biosciences) со 100 мкм соплом в стерильных

¹⁵ условиях в 96-луночные планшеты, покрытые облученными мышиными эмбриональными фибробластами CF-1 (Global Stem). Каждая лунка содержала клеточную среду для hES (см. источник 31, включенный в данный документ ссылкой в полном объеме) со 100 нг/мл рекомбинантного человеческого основного фактора роста фибробластов (bFGF) (Millipore) с добавлением SMC4 и 5 мкг/мл фибронектина (Sigma).

²⁰ После сортировки планшеты центрифугировали при 70 x g в течение 3 мин. Образование колоний было отмечено через 4 дня после сортировки, а культуральную среду заменили клеточной средой для hES с SMC4. SMC4 может быть удалена из клеточной среды для hES через 8 дней после сортировки.

²⁵ Собирали несколько тысяч клеток через 8 дней после флуоресцентно-активируемой сортировки клеток и (FACS) и добавляли 0,1 мкл фермента-тканевой протеазы prepGEM (ZyGEM) и 1 мкл буфера prepGEM gold (ZyGEM) к 8,9 мкл клеток в среде. Реакции затем добавляли к 40 мкл ПЦР-смеси, содержащей 35,5 мл platinum 1,1X Supermix (Invitrogen), 250 нМ каждого dNTP и 400 нМ праймеров. Реакционные смеси инкубировали при 95°C в течение 3 мин с последующими 30 циклами 95°C, 20 с; 65°C, 30 и 72°C, 20 с. Продукты секвенировали по Сенгеру с использованием любого из ПЦР-праймеров таблицы 5 и анализировали последовательности с помощью DNASTAR (DNASTAR).

Пример X

Иммунноокрашивание и анализы тератом, полученных из hiPSCs

³⁵ Клетки инкубировали в нокаутной среде DMEM/F-12 при 37°C в течение 60 минут с использованием следующих антител: анти-SSEA-4 PE (Millipore) (разбавленное 1: 500); Tra-1-60 (BD Pharmingen) (разбавленное 1: 100). После инкубации клетки промывали три раза нокаутной DMEM/F-12 и фотографировали на Axio Observer Z.1 (ZIESS).

⁴⁰ Для проведения анализа образования тератом, человеческие iPSCs собирали с помощью коллагеназы IV типа (Invitrogen), ресуспендировали клетки в 200 мкл матригеля и вводили внутримышечно в подушечки задних конечностей мышей с нокаутом по Rag2gamma. Тератомы выделяли и фиксировали в формалине через 4-8 недель после инъекции. Тератомы затем анализировали после окрашивания гематоксилином и эозином.

Пример XI

⁴⁵ Ориентация локусов генома в соматических клетках человека и стволовых клетках человека с помощью reTALENS

Согласно некоторым аспектам, TALEs, известные специалистам в данной области техники, модифицированы или перекодированы для устранения повторяющихся

последовательностей. Такие TALEs, подходящие для модификации и применения в способах редактирования генома в вирусных средствах доставки и в различных клеточных линиях и микроорганизмах, описанных в данном документе, описаны в источниках 2, 7-12, включенных в данных документ ссылкой в полном объеме. Было 5 разработано несколько стратегий для сборки массива повторяющихся TALE RVD (см. источники 14 и 32-34, включенные в данных документ ссылкой в полном объеме). Однако после сборки повторяющиеся последовательности TALE остаются нестабильными, что ограничивает широкую применимость этого инструмента, особенно для вирусных средств доставки гена (см. источники 13 и 35 включенные в данных документ ссылкой 10 в полном объеме). Соответственно, один аспект настоящего изобретения относится к TALEs, лишенных повторов, например, полностью лишенных повторов. Такие перекодированные TALE являются эффективными, поскольку позволяют быстрее и проще синтезировать расширенные массивы TALE RVD.

Для устранения повторов, нуклеотидные последовательности массивов TALE RVD

15 эволюционировали вычислительным образом с тем, чтобы минимизировать количество последовательностей-повторов при сохранении аминокислотной композиции. Перекодированные TALE (Re-TALEs), кодирующие 16 tandemных мономеров RVD, распознающих ДНК, плюс конечный полуповтор RVD, лишены любых 12 п.о. повторов (см. Фигуру 5а). Примечательно, что этот уровень перекодирования достаточен для 20 того, чтобы провести амплификацию ПЦР какого-либо конкретного мономера или подраздела из полноразмерного конструкции с re-TALE (см. Фиг. 5б). Усовершенствованный дизайн re-TALEs может быть синтезирован с помощью стандартных технологий синтеза ДНК (см. источник 36, включенный в данный документ ссылкой в полном объеме, без каких дополнительных расходов или процедур, связанных 25 с обогащенными повторами последовательностями. Кроме того, дизайн перекодированных последовательностей позволяет эффективно собирать конструкции с re-TALE с использованием модифицированной изотермической реакции сборки, как описано в способах данного документа и со ссылкой на Фиг. 6.

NGS-данные по редактированию генома статистически анализировали следующим

30 образом. Для анализа специфичности HDR, использовали точный биномиальный тест для вычисления вероятности наблюдения различных количеств «ридов», содержащих 2 п.о. несовпадения. На основании результатов секвенирования 10 п.о. окон до и после участка направленного воздействия, оценивали максимальные скорости замены оснований двух окон (P1 и P2). С помощью нулевой гипотезы о том, что изменения 35 каждой из двух целевых п.о. были независимыми, была вычислена ожидаемая вероятность обнаружения 2 п.о. несоответствия в участке направленного воздействия случайно как произведение этих двух вероятностей ($P1 * P2$). Учитывая набор данных, содержащих N количество общих ридов и n количество HDR-ридов, мы рассчитали p-значение наблюданной эффективности HDR. Для анализа чувствительности HDR, 40 донорные ssODN содержали 2 п.о. несовпадение относительно генома-мишени, что делает вероятной одновременную замену оснований в двух п.о., являющихся мишениями при включении ssODN в геном-мишень. Прочие нецелевые наблюдаемые изменения последовательностей, вероятно, не будут изменяться в это же время. Соответственно, нецелевые изменения были гораздо менее взаимозависимы. Исходя из этих 45 предположений, взаимная информация (MI) была использована для измерения взаимной зависимости одновременных двух изменений пар оснований во всех других парах позиций, и предел обнаружения HDR оценивали как наименьшее HDR, где MI целевого 2 п.о. сайта выше, чем MI всех пар в других позициях. Для данного эксперимента, в

оригинальном файле fastq были идентифицированы риды HDR с целевым 2 п.о. несовпадением и набор файлов fastq с разбавленными эффективностями HDR был смоделирован посредством систематического удаления различных количеств ридов HDR из исходного набора данных. Взаимную информацию (MI) вычисляли между всеми 5 парами положений в пределах 20 п.о. окна с центром на участке направленного воздействия. В этих расчетах вычисляется взаимная информация о композиции оснований между любыми двумя позициями. В отличие от меры специфичности HDR, описанной выше, эта мера не оценивает тенденцию пар позиций изменять какие-либо конкретные пары целевых оснований, а только оценивает их тенденцию изменяться в одно и тоже 10 время (см. Фиг. 8А). В таблице 6 приведены эффективности HDR и NHEJ для направленного воздействия re-TALEN/ssODN на CCR5 и эффективность NHEL для Cas9-гидРНК. Мы запрограммировали наш анализ в R и рассчитали MI при помощи пакета infotheo.

Таблица 6

# участок направленного воздействия	типа клеток	HDR (reTALEN) (%)	NHEJ (reTALEN) (%)	Предел обнаружения HDR на основе информационного анализа	NHEJ (Cas9-гидРНК)	HDR (Cas9-гидРНК)
1	PGP1-iPS	0,06%	0,80%	0,04%	0,58%	0,38%
2	PGP1-iPS	0,48%	0,26%	0,01%	16,02%	3,71%
3	PGP1-iPS	1,71%	0,07%	0,03%	3,44%	3,20%
4	PGP1-iPS	0,02%	1,20%	0,02%*	1,50%	0,14%
5	PGP1-iPS	0,80%	0,04%	0,00%	3,70%	0,39%
6	PGP1-iPS	0,20%	0,73%	0,00%	1,12%	0,49%
7	PGP1-iPS	0,01%	0,15%	0,01%*	1,98%	1,78%
8	PGP1-iPS	0,03%	0,00%	0,00%	1,85%	0,03%
9	PGP1-iPS	1,60%	0,06%	0,00%	0,50%	0,13%
10	PGP1-iPS	0,68%	1,25%	0,01%	8,77%	1,32%
11	PGP1-iPS	0,06%	0,27%	0,00%	0,62%	0,44%
12	PGP1-iPS	1,60%	0,03%	0,04%	0,18%	0,99%
13	PGP1-iPS	0,00%	1,47%	0,00%	0,65%	0,02%
14	PGP1-iPS	0,47%	0,13%	0,02%	2,50%	0,31%
15	PGP1-iPS	0,8	0,14	0,08%	1,50	1,10%

* Группа, в которой предел обнаружения HDR превышает реальное детектируемое значение HDR

Корреляцию между эффективностью редактирования генома и эпигенетическим состоянием урегулировали следующим образом. Коэффициенты корреляции Пирсона рассчитывали для изучения возможных связей между эпигенетическими параметрами (заполненность ДНКазой I HS или нуклеосомами) и эффективностями генной инженерии (HDR, NHEJ). Набор данных по гиперчувствительности к ДНКазе I был загружен из геномного браузера UCSC. ДНКаза I HS из hiPSCs:

/gbdb/hg19/bbi/wgEncodeOpenChromDnaseIpsihi7Sig. bigWig

Для вычисления значения p, наблюдаемую корреляцию сравнивали с моделируемым распределением, которая была построена рандомизацией положения эпигенетического параметра ($N = 100\ 000$). Наблюдаемые корреляции выше, чем 95-й процентиль, или ниже, чем 5-й процентиль моделируемого распределения, считались потенциальными ассоциациями.

Определяли функцию re-TALEN по сравнению с соответствующими не перекодированными TALEN в клетках человека. Линию клеток HEK 293, содержащую репортерную кассету с GFP, и несущую вставку, смещающую рамку считывания, использовали, как описано в источнике 37, включенном в данное описание ссылкой в

полном объеме. Смотрите также Фиг. 1а. Доставка TALENs или reTALENs, направленно воздействующих на последовательность вставки, вместе с беспромоторной донорной конструкцией с GFP, приводит к DSB-индуцированному HDR-восстановлению кассеты с GFP, при этом эффективность восстановления GFP может быть использована для 5 оценки эффективности разрезания нуклеазой. См. источник 38, включенный в данный документ ссылкой в полном объеме. Индуцированное reTALENs восстановление GFP в 1,4% трансфецированных клеток, было аналогично полученному с помощью TALENs (1,2%) (см. Фиг. 1б). Активность reTALENs была протестирована в локусе AAVS1 в PGP1 hiPSCs (см. Фиг. 1с), при этом успешно восстановленные клеточные клоны, 10 содержащие специфические вставки (см. Фиг. 1д, е), подтвердили, что reTALENs активны как в соматических, так и в плюрипотентных клетках человека.

Устранение повторов позволяет получить функциональный лентивирирус, 15 нагруженный re-TALE. В частности, лентивирусные частицы, в которые были упакованы фрагменты, кодирующие re-TALE-2A-GFP, были протестированы на активность re-TALE-TF, кодируемых вирусными частицами, путем трансфекции репортера mCherry в пул клеток 293T, инфицированных lenti-reTALE-2A-GFP. Клетки 293T, 20 трансдудцированные lenti-re-TALE-TF показали 36x активацию экспрессии репортера по сравнению с только отрицательным репортером (см. Фиг. 7а, б, с). Проверяли целостность последовательности re-TALE-TF в лентивирус-инфицированных клетках и обнаружили полноразмерные reTALEs во всех 10 протестированных клонах (см. Фиг. 1д).

Пример XII

Сравнение эффективности ReTALEs и Cas9-гидРНК в hiPSCs с помощью системы оценки редактирования генома (GEAS)

Для сравнения эффективности редактирования re-TALENs относительно Cas9-гидРНК 25 в hiPSCs, была разработана платформа секвенирования следующего поколения (Система оценки Редактирования Генома) для выявления и количественной оценки как NHEJ, так и HDR событий генного редактирования. Пары re-TALEN и Cas9-гидРНК были спроектированы и сконструированы как направленно воздействующие на область, 30 лежащую выше CCR5 (re-TALEN, Cas9-гидРНК пара # 3 в Таблице 3), вместе с 90 нуклеотидным ssODN-донором, идентичным участку направленного воздействия за исключением 2 п.о. несовпадения (см. Фиг. 2а). Нуклеазные конструкции и донорный ssODN трансфецировали в hiPSCs. Для количественной оценки эффективности генного 35 редактирования, через 3 дня после трансфекции проводили глубокое секвенирование с использованием спаренных концевых фрагментов на целевой геномной области. Эффективность HDR измеряли по проценту «ридов», содержащих точное 2 п.о. несовпадение. Эффективность NHEJ измеряли как процент «ридов», несущих вставки/делеции.

Доставка ssODN отдельно в hiPSCs давала минимальные показатели HDR и NHEJ, 40 в то время как доставка re-TALENs и ssODN давала эффективности 1,7% по HDR и 1,2% по NHEJ (см. Фиг. 2б). Введение Cas9-гидРНК с ssODN давало эффективности 1,2% по HDR и 3,4% по NHEJ. Примечательно, что показатель геномных делеций и вставок достиг своего пика в середине спейсерной области между двумя участками связывания reTALENs, но своего пика достиг на 3-4 п.о. выше последовательности Protospacer 45 Associated Motif (PAM) участка направленного воздействия Cas9-гидРНК (см. Фиг. 2б), как и следовало ожидать, поскольку двухцепочечные разрывы происходят в этих областях. В случае re-TALENs наблюдали средний геномный размер делеции 6 п.о. и размер вставки 3 п.о., а в случае Cas9-гидРНК средний размер делеции 7 п.о., вставки

- 1 п.о. (см. Фиг. 2b), в соответствии с паттернами повреждений ДНК обычно образованных NHEJ (см. источник 4, включенный в данный документ ссылкой в полном объеме). Несколько анализов на платформе секвенирования следующего поколения показало, что GEAS может обнаружить уровень обнаружения HDR до 0,007%, который 5 является одновременно высоко воспроизводимым (коэффициент вариации между повторами = $\pm 15\%$ * измеренная эффективность) и в 400X более чувствительным, чем наиболее часто используемые тесты с чувствительной к несовпадениям эндонуклеазой (см. Фиг. 8).

Пары ге-TALEN и Cas9-гидРНК, нацеленные на пятнадцать участков в геномном локусе CCR5, были созданы для определения эффективности редактирования (см. Фиг. 2c, таблица 3). Эти участки были выбраны для того, чтобы представлять широкий спектр чувствительностей ДНКазы I (см. источник 39, включенный в данный документ ссылкой в полном объеме). Нуклеазные конструкции трансфенировали вместе с соответствующими донорными ssODNs (см. таблицу 3) в PGP1 hiPSCs. Через шесть дней 10 после трансфекции профилировали эффективности редактирования генома в этих участках (таблица 6). Для 13 из 15 пар ге-TALEN с донорным ssODN, NHEJ и HDR детектировали на уровнях выше статистического порога обнаружения, со средней эффективностью NHEJ 0,4% и средней эффективностью HDR 0,6% (см. Фиг. 2c). Кроме того, статистически значимую положительную корреляцию ($R^2 = 0,81$) обнаружили 15 между эффективностями HR и NHEJ в одних и тех же целевых локусах ($P < 1 \times 10^{-4}$) (см. Фиг. 9a), что предполагает, что образование DSB, общей вышестоящей стадии как для HDR, так и для NHEJ, является скоростью-лимитирующей стадией для геTALEN-опосредованного редактирования генома.

В противоположность этому, все 15 пар Cas9-гидРНК показали значительные уровни 20 NHEJ и HR, со средней эффективностью NHEJ 3% и средней эффективностью HDR 1,0% (см. Фиг. 2c). Кроме того, положительная корреляция была также обнаружена между эффективностью NHEJ и HDR, вызванных Cas9-гидРНК (см. Фиг. 9b) ($R^2 = 0,52$, $p = 0,003$), что согласуется с наблюдениями для геTALENs. Эффективность NHEJ, 25 достигаемая Cas9-гидРНК была значительно выше, чем та, что достигается с помощью геTALENs (t-тест, спаренные концы, $P = 0,02$). Наблюдалось умеренная, но статистически значимая корреляция между эффективностью NHEJ и температурой плавления целевой последовательности гидРНК (см. Фиг. 9c) ($R^2 = 0,28$, $p = 0,04$), что позволяет 30 предположить, что сила спаривания оснований между гидРНК и его геномной мишенью может объяснить вплоть до 28% вариации в эффективности Cas9-гидРНК-опосредованного образования DSB. Даже если Cas9-гидРНК продуцируют уровни 35 NHEJ в среднем в 7 раз выше, чем соответствующие геTALEN, Cas9-гидРНК достигает уровней HDR (среднее = 1,0%), схожих с уровнями соответствующих геTALENs (среднее = 0,6%). Без связи с научной теорией, эти результаты позволяют предложить либо то, что концентрация ssODN при DSB является ограничивающим фактором для HDR, либо 40 то, что структура геномного разрыва, образованного Cas9-гидРНК не выгодна для эффективного HDR. Никакой корреляции между ДНКазой I HS и эффективностями направленного воздействия на геном не наблюдали ни для одного из способов (см. Фиг. 10).

Пример XIII

45 Оптимизация Дизайна донорных ssODN для HDR

Высоко эффективные в hiPSCs ssODNs были разработаны следующим образом. Был разработан набор донорных ssODNs разных длин (50-170 нуклеотидов), которые все несут одинаковое 2 п.о. несовпадение в середине спайсерной области участков

направленного воздействия пары ге-TALEN #3 для CCR5. Наблюдалась эффективность HDR варьировала в зависимости от длины ssODN, и оптимальная эффективность HDR -1,8% наблюдалась с 90 нуклеотидным ssODN, тогда как более длинные ssODNs имели пониженную эффективность HDR (см. Фиг. 3а). Поскольку более длинные участки гомологии улучшают показатели HDR, когда донорные дНК использовались с нуклеазами (см. источник 40, включенный ссылкой в полном объеме), возможные причины такого результата могут заключаться в том, что ssODNs используются в альтернативном процессе восстановления генома; более длинные ssODNs менее доступны для аппарата восстановления генома; или в том, что более длинные ssODNs привносят отрицательные эффекты, которые нейтрализуют любые улучшения, придаваемые более длинной гомологией, по сравнению с дНК-донорами (см. источник 41, включенный в данный документ ссылкой в полном объеме). Тем не менее, если дело было в любой из первых двух причин, то показатели NHEJ должны быть либо неизменными, либо увеличиваться с более длинными ssODNs, потому что репарация NHEJ не включает донорный ssODN. Тем не менее, наблюдали снижение показателей NHEJ вместе с HDR (см. Фиг. 3а), что позволяет предположить, что более длинные ssODNs демонстрируют компенсирующие эффекты. Возможные гипотезы могут заключаться в том, что более длинные ssODNs являются токсичными для клетки (см. источник 42, включенный в данный документ ссылкой в полном объеме), или в том, что трансфекция длинных ssODNs насыщает машинерию процессирования ДНК, тем самым вызывая уменьшение молярного поглощения ДНК и уменьшая емкость клеток поглощать и экспрессировать плазмиды с ге-TALEN.

Была проведена проверка вариабельности показателя включения несовпадений, которые несут донорные ssODN, в зависимости от дистанции от двухцепочечного разрыва («DSB»). Для этого разработали серию 90 нуклеотидных ssODNs, которые все обладают одинаковым 2. п.о. несовпадением (A) в центре спайсерной области пары ге-TALEN #3. Каждый ssODN также содержала второе 2 п.о. несовпадение (B) на различных расстояниях от центра (см. Фиг. 3б). ssODN, обладающий лишь одним 2 п.о. несовпадением использовали в качестве контроля. Каждый из этих ssODNs вводили в индивидуальном порядке с парой ге-TALEN #3 и анализировали результаты с помощью GEAS. Мы обнаружили, что общий HDR - если измерять по показателю, при котором включение несовпадения A (только A или A + B) - уменьшается по мере того, как несовпадения B располагаются дальше от центра (см. Фиг. 3б, см. Фиг. 11а). Более высокий показатель HDR, наблюдаемый, когда B находится всего лишь в 10 п.о. от A, может отражать меньшую потребность в отжиге ssODN против геномной ДНК в непосредственной близости от разрыва двухцепочечной ДНК.

Для каждого расстояния B от A, часть событий HDR относится к включению только несовпадения A, тогда как другая часть относится к включению как A, так и B несовпадений (см. Фиг. 3б (только A и A + B)). Эти два результата могут быть связаны с трактами генной конверсии (см. источник 43, включенный в данный документ ссылкой в полном объеме) по длине одноцепочечного ДНК-олигонуклеотида, в результате чего включение несовпадений A+B являющегося результатом длинных конверсионных трактов, которые простираются за пределы несовпадения B, а включение только A несовпадения является результатом более коротких путей, которые не достигают B. в соответствии с этой интерпретацией, оценивали распределение трактов генной конверсии в обоих направлениях вдоль ssODN (см. Фиг. 11б). Оцененное распределение подразумевает, что тракты генной конверсии постепенно становятся менее частыми, по мере увеличения их длины, результат, очень похожий на распределения трактов генной

конверсии, наблюдаемых с донорными дЦДНК, но на сильноожатой шкале дистанций в десятки основания для оЦДНК донора относительно сотен оснований для дЦДНК доноров. В соответствии с этим результатом, эксперимент с ssODN, включающим три пары 2 п.о. несовпадений с интервалом 10 нуклеотидов в обе стороны от центрального 5 2 п.о. несовпадения «А», позволил получить паттерн, в котором А отдельно был включен в 85% случаев, с множественными несовпадениями В, включенными в других случаях (см. Фиг. 11c). Хотя количество событий с включением только В было слишком низким для оценки распределения длин путей менее 1 п. о., ясно, что область с коротким путем в пределах 10 п.о. от нуклеазного сайта является преобладающей (см. Фиг. 11,b).

10 Наконец, во всех экспериментах с одиночными несовпадениями В, наблюдалась небольшая часть событий включения только В (0,04% ~ 0,12%), что приблизительно постоянно для всех расстояний от А до В.

Кроме того, был проведен анализ того, как далеко может быть расположен донор ssODN от re-TALEN-индуцированного разрыва двухцепочечной ДНК, при сохранении 15 его включения. Протестировали набор 90-нуклеотидных ssODNs с центральными 2 п.о. несовпадениями, нацеленными на диапазон больших дистанций (-600 п.о. до + 400 п.о.) от re-TALEN-индуцированного участка разрыва дЦДНК. Когда ssODNs совпадает с последовательностью в > 40 п.о., мы наблюдали > 30x более низкие эффективности HDR по сравнению с контрольным ssODN, расположенным по центру над областью 20 разрезания (Фиг. 3c). Наблюдавшийся низкий уровень включения может быть результатом процессов, несвязанных с разрезанием дЦДНК, как видно в экспериментах, в которых геномы изменены отдельно оЦДНК донором, см. источник 42, включенный в данный документ ссылкой в полном объеме. Между тем, низкий уровень HDR, который имеет место, когда ssODN расположен в ~ 40 п.о., может быть связан с комбинацией 25 ослабленной гомологии на содержащей несовпадение стороне разреза дЦДНК, вместе с недостаточной длиной оЦДНК олигонуклеотида на другой стороне дЦДНК разрыва.

Было проведено тестирование дизайна донора ssODNs ДНК для опосредованного Cas9-гидРНК направленного воздействия. Сконструировали Cas9-гидРНК (C2), направленно действующий на локус AAVS 1, и ssODN-доноры различных ориентаций 30 (Ос: комплементарная к гидРНК и On: некомплементарная к гидРНК) и длин (30, 50, 70, 90, 110 нуклеотидов). Os были более эффективны, чем On, а 70-мерный Os достиг оптимального показателя HDR 1,5% (см. Фиг. 3d.) Такую же тенденциозность цепи ssODN детектировали с помощью низкозы из Cas9 (CC: Cas9_D10A), несмотря на тот факт, что эффективности HDR, опосредованные Cc с ssODN были значительно меньше, 35 чем C2 (t-тест, спаренные концы, P = 0,02). (см. Фиг. 12).

Пример XIV

Клональное Выделение Скорректированных Клеток hiPSC

GEAS показало, что re-TALEN пара #3 достигла точного редактирования генома в hiPSCsc эффективностью ~ 1%, уровня, при котором правильно отредактированные 40 клетки обычно можно выделить скринингом клонов. HiPSCs имеют плохую жизнеспособность в виде отдельных клеток. Оптимизированные протоколы, описанные в источнике 23, включенном в данный документ ссылкой в полном объеме, вместе с процедурой сортировки одноклеточной FACS использовали для создания надежной платформы сортировки одиночных hiPSCs и их поддержания, где клоны hiPSC могут 45 быть восстановлены с показателями выживания > 25%. Этот способ был объединен с быстрой и эффективной системой генотипирования, для проведения экстракции хромосомной ДНК и целевой амплификации генома в 1 ч реакциях в одной пробирке, что позволяет осуществить генотипирование hiPSCs в большом масштабе. Вместе эти

способы представляют собой конвейер для надежного получения hiPSCs с отредактированным геномом без отбора.

Чтобы продемонстрировать эту систему (см. Фиг. 4а), PGP1 hiPSCs трансфецировали парой reTALENs и ssODN, предназначенных для направленного воздействия на CCR5

5 в участке #3 (таблица 3). GEAS проводили на части трансфицированных клеток, при этом обнаружили частоту HDR 1,7% (см. Фиг. 4б). Эта информация, наряду с 25% выходом отсортированных клонов единичных клеток, позволяют оценить получение, по меньшей мере, одного правильно отредактированного клона из пяти 96-луночных планшетов с пуассоновской вероятностью 98% (при условии, что $\mu = 0,017 * 0,25 * 96$

10 * 5 * 2). Через шесть дней после трансфекции, hiPSCs сортировали с помощью FACS, а через восемь дней после сортировки, 100 клонов hiPSC подвергли скринингу.

Секвенирование по Сенгеру показало, что 2 из этих 100 не подвергнутых отбору колоний hiPSC содержали гетерозиготный генотип, обладающий 2 п.о. мутацией, введенной донорным ssODN (см. Фиг. 4с). Эффективность направленного воздействия 1% (1% =

15 2/2 * 100, 2monoаллельных скорректированных клона из 100 клеток, подвергнутых скринингу) была сопоставима с данными, полученными при анализе секвенированием следующего поколения (1,7%) (см. Фиг. 4б). Плюрипотентность полученных hiPSCs подтвердили иммуногистохимическое окрашивание SSEA4 и TRA-1-60 (см. Фиг. 4д).

Подвергнутые успешному направленному воздействию клоны hiPSCs были способны 20 образовывать зрелые тератомы с признаками всех трех зародышевых листков (см. Фиг. 4е).

Пример XV

Способ непрерывного редактирования генома клетки

Согласно некоторым аспектам, предлагается способ редактирования генома в

25 клетках, включающий человеческую клетку, например, человеческую стволовую клетку, где клетка генетически модифицирована для включения нуклеиновой кислоты, кодирующую фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени, и который расщепляет ДНК-мишень сайт-специфическим образом. Такой фермент включает направляемый РНК ДНК-

30 связывающий белок, такой как РНК-направляемый ДНК-связывающий белок системы CRISPR II типа. Типичным ферментом является Cas9. Согласно этому аспекту, клетка экспрессирует фермент, а направляющая РНК вносится в клетку из среды, окружающей клетку. Направляющая РНК и фермент образуют комплекс колокализации на ДНК-мишени, где указанный фермент разрезает ДНК. Необязательно, донорная нуклеиновая

35 кислота может присутствовать для вставки в ДНК в месте разреза, например, путем негомологичного соединения концов или гомологичной рекомбинации. Согласно одному аспекту, нуклеиновая кислота, кодирующая фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишени сайт-специфическим образом, такой как Cas9, находится под управлением

40 промотора, так что нуклеиновая кислота может быть активирована и подавлена. Такие промоторы, хорошо известны специалистам в данной области техники. Одним примерным промотором является промотор, индуцируемый DOX. Согласно одному аспекту, клетка генетически модифицирована обратимой вставкой в ее геном

45 нуклеиновой кислоты, кодирующую фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом. После вставки нуклеиновая кислота может быть удалена с помощью реагтива, такого как транспозаза. Таким образом, нуклеиновая кислота может быть легко удалена после применения.

Согласно одному аспекту, предлагается система непрерывного редактирования генома человеческих индуцированных плюрипотентных стволовых клеток (hiPSCs) с помощью системы CRISPR. В соответствии с типичным аспектом способ включает применение линии hiPSC с Cas9, обратимо вставленным в геном (Cas9-hiPSCs); и гидРНК, которые были модифицированы из их нативной формы с тем, чтобы позволить им переход из среды, окружающей клетки, в клетки для применения вместе с Cas9. Такие гидРНК обрабатывали фосфатазой, для удаления фосфатных групп. Редактирование генома в клетки осуществляется с помощью Cas9 путем обогащения обработанной фосфатазой гидРНК среды для культивирования тканей. Такой подход позволяет 5 редактировать геном HiPSCs без «шрамов» с эффективностями вплоть до 50% при однодневных обработках, которые в 2-10 раз более эффективны, чем наилучшие 10 эффективности, известные к настоящему времени. Кроме того, способ прост в использовании и имеет значительно меньшую клеточную токсичность. Воплощения 15 настоящего изобретения включают одночное редактирование hiPSCs для биологических исследований и терапевтических применений, мультиплексное редактирование hiPSCs для биологических исследований и терапевтических применений, направленную эволюцию hiPSCs и скрининг hiPSCs по фенотипу и клеток, производных от них.

В соответствии с некоторыми аспектами, другие клеточные линии и организмы, описанные в данном документе, могут быть использованы в дополнение к стволовым 20 клеткам. Например, описанный в данном документе способ может быть использован для клеток животных, таких как мышиные или крысиные клетки, с получением мышиных и крысиных клеток со стабильной интеграцией Cas9, и может быть проведено тканеспецифичное редактирование генома местным введением обработанной фосфатазой направляющей РНК из среды, окружающей клетки. Более того, другие производные 25 Cas9 могут быть вставлены во множество клеточных линий и микроорганизмов, и могут быть проведены геномные манипуляции направленного воздействия, такие как внесение одноцепочечных разрывов, специфичное к последовательности, активация генов, супрессия и эпигенетическая модификация.

Аспекты настоящего описания направлены на получение стабильных hiPSCs с 30 вставленным в геном Cas9. Аспекты настоящего описания направлены на модификацию РНК, так чтобы позволить ей проникать в клетку через клеточную стенку и колокализоваться с Cas9, при этом избегая иммунного ответа клетки. Такая модифицированная направляющая РНК может достичь оптимальных эффективностей трансфекции с минимальной токсичностью. Аспекты настоящего раскрытия направлены 35 на оптимизированное редактирование генома в Cas9-hiPSCs с использованием гидРНК, обработанных фосфатазой. Аспекты настоящего изобретения включают устранение Cas9 из hiPSCs для достижения редактирования генома без «шрамов», где нуклеиновая кислота, кодирующая Cas9, была обратимо помещена в клеточный геном. Аспекты настоящего изобретения включают биомедицинскую инженерию с использованием 40 hiPSCs с вставленным в геном Cas9 в для создания искомых генетических мутаций. Такие подвергнутые инженерии hiPSCs поддерживают плюрипотентность и могут успешно дифференцироваться в различные типы клеток, в том числе кардиомиоциты, которые полностью воспроизводят фенотип клеточных линий пациента.

Аспекты настоящего изобретения включают библиотеки направляющих РНК, 45 обработанных фосфатазой, для мультиплексного редактирования генома. Аспекты настоящего изобретения включают получение библиотеки клеточных линий РГР, в которых каждая клеточная линия несет несколько обозначенных мутаций в геном, которые служат в качестве ресурса для скрининга лекарственных средств. Аспекты

настоящего изобретения включают получение клеточных линий PGP1 со всеми ретроэлементами, подвергнутыми штрихкодированию различными последовательностями, для отслеживания расположений и активности этого элемента.

Пример XVI

Получение стабильных hiPSCs со вставленным в геном Cas9

Конструкт с Cas9, экспрессируемый под DOX-индукцируемым промотором, помещали в вектор PiggyBac, который может быть вставлен в геном и удален из него с помощью транспозазы PiggyBac. Реакция ПЦР подтверждала стабильное введение вектора (см. Фиг 14). Индуцируемую экспрессию Cas9 определяли с помощью количественной ПЦР в реальном времени. Уровень мРНК Cas9 увеличился в 1000x после 8 часов 1 мкг/мл DOX в культуральной среде и упал до нормального уровня через ~20 часов после удаления DOX (См. Фиг. 15).

Согласно одному аспекту, редактирование генома на основе системы Cas9-hiPSC обходит процедуру трансфекции плазмида с Cas9/РНК, большой конструкции обычно с эффективностью трансфекции в hiPSCs <1%. Настоящая система Cas9-hiPSC может служить в качестве платформы для осуществления высоко эффективной геномной инженерии стволовых клеток человека. Кроме того, кассета Cas9, введенная в hiPSCs, с помощью системы PiggyBac, может быть легко удалена из генома при введении транспозазы.

Пример XVII

Направляющая РНК, обработанная фосфатазой

Для того, чтобы сделать возможным непрерывное редактирование генома Cas9-hiPSCs, были получены серии модифицированных РНК, кодирующими направляющую РНК, которые вносились в культуральную среду для Cas9-iPS в комплексе с липосомой. Нативная РНК, обработанная фосфатазой, без какого-либо кэпирования достигает оптимального значения HDR 13%, что более чем в 30X больше, чем значения эффективности, сообщаемые ранее для 5'Cap-Mod РНК (см. Фиг. 16).

Согласно одному аспекту, направляющая РНК физически прикреплена к донорной ДНК. Таким образом, предлагается способ объединения опосредованного Cas9 разрезания в геноме и ssODN-опосредованной HDR, что приводит к стимуляции последовательность-специфического редактирования генома. Направляющая РНК, связанная с донорным ssODN в оптимизированной концентрации, достигает 44% HDR и 2% неспецифического NHEJ (см. Фиг. 17). Следует отметить, что эта процедура не вызывает видимой токсичности, которая наблюдается в случае нуклеофекции и электропорации.

Согласно одному аспекту, в настоящем изобретении предлагается *in vitro* сконструированная РНК-структура, кодирующая направляющую РНК, которая достигает высокой эффективности трансфекции, эффективности редактирования генома в сочетании со вставленным в геном Cas9. Кроме того, настоящее изобретение обеспечивает химерный конструкт направляющая РНК-ДНК для объединения события разрезания генома с реакцией направляемой гомологией рекомбинации.

Пример XVIII

Удаление из hiPSCs обратимо вставленного Cas9

Для достижения геномного редактирования без «шрамов» согласно некоторым аспектам, кассету с Cas9 вставляли в геном клеток hiPSC с помощью обратимого вектора. Соответственно, кассета Cas9 обратимо встраивается в геном клеток hiPSC с помощью вектора PiggyBac. Кассета с Cas9 удаляли из hiPSCs с отредактированным геномом трансфекцией клетки плазмидой, кодирующей транспозазу. Соответственно, аспекты

настоящего изобретения включают применение обратимого вектора, который известен специалистам в данной области техники. Обратимый вектор является вектором, который может быть, например, вставлен в геном, и затем удален с помощью соответствующего фермента для удаления вектора. Такие векторы и соответствующие ферменты для

5 удаления вектора известны специалистам в данной области техники. Проводили скрининг на колонизированных iPSC-клетках, и выделяли колонии, лишенные кассеты с Cas9, что было подтверждено реакцией ПЦР. Соответственно, настоящее изобретение обеспечивает способ редактирования генома, который не затрагивает остальную часть генома, при наличии в клетке перманентно вставленной кассеты с Cas9.

10 Пример XIX

Редактирование генома в клетках iPSC1

Исследование патогенеза кардиомиопатии исторически было затруднено из-за отсутствия подходящих модельных систем. Дифференцировка кардиомиоцитов из индуцированных плюрипотентных стволовых клеток, полученных из пациентов (iPSCs)

15 считается одним из перспективных путей преодоления данного препятствия, и в настоящее время начинают появляться сообщения о моделировании кардиомиопатии на основе iPSCs. Тем не менее, реализация этой перспективы требует подходов, которые преодолевают генетическую гетерогенность линий iPSC, полученных из пациентов.

Клеточные линии Cas9-iPGP1 и обработанную фосфатазой направляющую РНК, 20 связанную с ДНК, использовали для получения трех линий iPSC, которые являются изогенными, за исключением последовательности экзона 6 TAZ, который был идентифицирован как несущий однонуклеотидную делецию у пациентов с синдромом Барта. Однорундная трансфекция РНК дает ~30% эффективность HDR.

Модифицированные клетки Cas9-iPGP1 с искомыми мутациями выращивали до уровня 25 колоний (см. Фиг. 18), после чего клеточные линии дифференцировали в кардиомиоциты. Кардиомиоциты полученные из сконструированных Cas9-iPGP1, полностью восстановили кардиолипин, митохондриальный дефицит и дефицит АТФ, наблюдавшиеся в iPSCs, которые были получены из пациентов, и в неонатальной крысиной модели с нокаутом TAZ (см. Фиг. 19). Соответственно, предлагаются способы коррекции мутаций, 30 вызывающих заболевания, в плюрипотентных клетках с последующей дифференцировкой клеток в искомые типы клеток.

Пример XX

Материалы и методы

1. Получение стабильных клеточных линий iPSC/ES с PiggyBacCas9, индуцируемым

35 DOX

1. После достижения клетками 70% конфлюэнтности культуру предварительно обрабатывали в течение ночи ингибитором ROCK Y27632 в конечной концентрации 10 мКМ.

2. На следующий день готовили раствор для нуклеофекции путем объединения 82 40 мкл раствора для нуклеофекции человеческих стволовых клеток и 18 мкл добавки 1 в стерильной 1,5 мл пробирке Eppendorf. Хорошо перемешивали. Инкубировали раствор при 37°C в течение 5 минут.

3. Аспирировали mTeSR1; осторожно промывали клетки 2 мл DPBS на лунку шестилуночного планшета.

45 4. Аспирировали DPBS, добавляли 2 мл/лунку версена, и оставляли культуру в инкубаторе при температуре 37°C, пока они не сгруппируются вместе и не потеряют адгезию, но до отсоединения. Это требует 3-7 мин.

5. Аккуратно аспирировали версен и добавляли mTeSR1. Добавляли 1 мл mTeSR1 и

сбивали клетки с помощью 1000 мкл микропипетки медленным потоком mTeSR1.

6. Собирали отсоединившиеся клетки, мягко разбивали их в одноклеточную суспензию, количественно оценивали с помощью гематоцитометра, и корректировали плотность клеток до 1 миллиона клеток на мл.

5 7. Добавляли 1 мл суспензии клеток в 1,5 мл пробирку Eppendorf и центрифугировали при 1100 оборотов в минуту в течение 5 мин в настольной центрифуге.

8. Ресуспендировали клетки в 100 мкл раствора Nucleofector для нуклеофекции человеческих стволовых клеток из стадии 2.

10 9. Переносили клетки к кювету Nucleofector с помощью пипетки на 1 мл. Добавляли в суспензию клеток в кювете 1 мкг плазмиды с транспозазой и 5 мкг плазмид PB Cas9. Смешивали клетки и ДНК осторожным втряхиванием.

10 10. Помещали кювету в Nucleofector. Выбирали программу B-016 и подвергали клетки нуклеофекции нажатием кнопки X.

11. Добавляли 500ul среды mTeSR1 с ингибитором ROCK в кювету после нуклеофекции.

15 12. Асперировали подвергнутые нуклеофекции клетки из кюветы с помощью предоставленной пластиковой пастеровской пипетки. Переносили клетки покапельно в покрытые матригелем лунки 6 луночного планшета со средом mTeSR1 и ингибитором ROCK. Инкубировали клетки при 37°C в течение ночи.

20 13. Заменяли среду в mTesr1 на следующий день и через 72 часа после трансфекции, добавляли пуромицин в конечной концентрации 1 мкг/мл. Линия выводилась в течение 7 дней.

2. Выделение РНК

1. Готовили ДНК-матрицу, гдеT7-промотор выше последовательности, кодирующей направляющую РНК.

25 2. Очищали ДНК с помощью Mega Clear Purification и нормализовали ее концентрацию.

30 3. Готовили пользовательские смеси NTPS для получения различных направляющих РНК.

	# 1 Смесь нативной РНК	[Конечная] (мМ)
30	GTP	7,5
	АТФ	7,5
	СТР	7,5
	УТР	7,5
	Общий объем	°
	# 2 Смесь Кэпированной нативной РНК	[Конечная] (мМ)
35	структурный аналог 3'-О-Me-m7G Cap (NEB)	6
	GTP	1,5
	ATP	7,5
	СТР	7,5
	УТР	7,5
	Общий объем	°
	#3 Модифицированная смесь РНК	[Конечная] (мМ)
40	GTP	7,5
	ATP	7,5
	5-Me-СТР (Tri-Link)	7,5
	Pseudo-УТР (Tri-Link)	7,5
	Общий объем	°
	#4 смесь кэпированной/модифицированной РНК	[Конечная] (мМ)
45	структурный аналог 3'-О-Me-m7G Cap (NEB)	6
	GTP	1,5
	ATP	7,5
	5-Me-СТР (Tri-Link)	7,5

Псевдо-UTP (Tri-Link)	7,5
Общий объем	◦

4. Готовили смесь для транскрипции *in vitro* при комнатной температуре.

Amt (мкл)	
Пользовательские NTP (* Добавить об./IVT rxn как указано выше) на льду	Н.Д.
Продукт ПЦР (100 нг/мкл) = всего 1600 нг ([Окончательная конц.] = 40 нг/мкл)	16
Буфер x10 (набор MEGAscript из Ambion) @ RT	4
Фермент T7 (набор MEGAscript из Ambion)	4

5 5. Инкубировали в течение 4 часов (3-6 часов ОК) при 37°C (амплификатор).

10 6. Добавляли 2мкл Turbo ДНКазы (набор MEGAscript из Ambion) к каждому образцу.

Тщательно перемешивали и инкубировали при 37°C в течение 15'.

15 7. Очищали обработанную ДНКазой реакцию с использованием MegaClear от Ambion в соответствии с инструкциями изготовителя.

8. Очищали РНК с помощью MEGAclear (очищенная РНК может храниться при -80 в течение нескольких месяцев).

15 9. Удаляли фосфатные группы, с целью избежать иммунной реакции Toll2 в клетке-хозяине.

Обработка РНК фосфатазой	IX	12
Для каждого образца РНК	~ 100 мкл	НД
10x буфер для антарктической фосфатазы	11 мкл	132
Антарктическая фосфатаза	2 мкл	24

20 Аккуратно перемешивали образец и инкубировали при 37 С в течение 30 секунд (30 секунд - 1 ч ок)

25 3. Трансфекция РНК

1. Рассевали 10K-20K клеток на 48 лунок без антибиотиков. Для трансфекции клетки должны быть на 30-50% конфлюэнтны.

2. Заменяли клеточную среду с B18R (200 нг/мл), DOX (1 мкг/мл), пуромицином (2 мкг/мл), по меньшей мере, за два часа до трансфекции.

30 3. Готовили реагент для трансфекции, содержащий направляющую РНК (0,5 мкг ~ 2 мкг), донорную ДНК (0,5 мкг ~ 2 мкг) и RNAiMax, инкубировали смесь при температуре rm в течение 15 минут и переносили в камеру к клеткам.

4. Рассеивание одиночных человеческих iPS клеток рассевали и сбор одиночных клонов

35 5. Через 4 дня индукции DOX и 1 дня отмены DOX, асперировали среду, отмывали осторожно с помощью DPBS, добавляли 2 мл/лунку версена, и помещали культуру обратно в инкубатор при температуре 37°C до тех пор, пока клетки не группировались вместе и становились непрочно прикрепленными, но не отсоединенными. Это требовало 3-7 мин.

40 6. Аккуратно аспирировали версен и добавляли mTeSR1. Добавляли 1 мл mTeSR1 и отсоединяли клетки с помощью 1000 мкл микропипетки медленным потоком mTeSR1.

7. Собирали отсоединенные клетки, мягко разбивали их в одноклеточную суспензию, количественно оценивали с помощью гематоцитометра, и корректировали плотность клеток до 100 тыс. клеток на мл.

45 8. Рассевали клетки на покрытые матригелем 10 см чашки с mTeSR1 плюс ингибитор ROCK при плотности клеток 50K, 100K и 400K на 10 см чашку.

9. Скрининг клонов, образованных из одиночных клеток

10 1. Через 12 дней культивирования в 10 см чашке, клоны становились достаточно

большими, чтобы быть идентифицированными невооруженным глазом, с возможностью мечения цветным маркером. Не давали клонам становиться слишком большими, и слипаться друг с другом.

2. Помещали 10 см чашку в ламинар и с помощью пипетки P20 (установленной на

5 10 мкл) с наконечниками с фильтрами аспирировали 10 мкл среды из каждой лунки 24 луночного планшета. Отбирали клоны клон путем соскребания клона и переноса в каждую лунку 24 луночного планшета. Для каждого клона использовали отдельный наконечник с фильтром.

3. Через 4-5 дней клоны внутри одной лунки 24-луночного планшета становились

10 достаточно большими для разделения.

4. Аспирировали среду и промывали 2 мл/лунку DPBS.

5. Аспирировали DPBS, заменяли 250 мкл/лунку диспазы (0,1 Ед/мл) и инкубировали клетки с диспазой при 37°C в течение 7 минут.

6. Заменяли диспазу 2 мл DPBS.

15 7. Добавляли 250 мкл mTeSR1. С помощью скребка для клеток снимали и собирали клетки.

8. Переносили 125 мкл суспензии клеток в лунку 24-луночного планшета, покрытого матригелем.

9. Переносили 125 мкл суспензии клеток в 1,5 мл пробирку Эппendorф для экстракции 20 геномной ДНК.

6. Скрининг клонов

1. Центрифугировали пробирку из стадии 7.7

2. Аспирировали среду и добавляли 250 мкл буфера для лизиса на лунку (10 mM + TrispH7.5 + (или + 8,0), 10 mM ЭДТА, 10 mM

25 3. NaCl + 10% ДСН, 40 мкг/мл протеиназы K + (добавленной свежей перед использованием буфера).

4. Выдерживали при 55 в течение ночи.

5. Высаживали ДНК добавлением 250 мкл изопропанола.

6. Центрифугировали в течение 30 минут на максимальной скорости. Промывали 30 70% этианолом.

7. Аккуратно удаляли этианол. Сушили на воздухе в течение 5 мин.

8. Ресуспендировали гДНК в 100-200 мкл dH2O.

9. Амплифицировали с помощью ПЦР целевую геномную область с помощью специфических праймеров.

35 10. Секвенировали по Сенгеру продукт ПЦР с помощью соответствующего праймера.

11. Анализировали данные о последовательностях по Сэнгеру и экспансию целевых клонов.

7. Удаление вектора PiggyBac

1. Повторяли стадии 2.1-2.9

40 2. Переносили клетки в кювету Nucleofector с помощью 1 мл наконечника для пипетки.

Добавляли 2 мкг плазмида с транспозазой в суспензию клеток в кювете. Смешивали клетки и ДНК осторожным покачиванием.

3. Повторяли стадию 2.10-2.11

45 4. Асперировали нуклеофицированные клетки из кюветы с помощью предоставленной пастеровской пластиковой пипетки и переносили клетки покапельно в 10 см чашки

покрытые матригелеем со средой mTeSR1 и ингибитором ROCK. Инкубировали клетки при 37°C в течение ночи.

5. На следующий день заменяли среду mTesr1 и заменяли среду каждый день в течение

4 последующих дней.

6. После того, как клоны стал достаточно большими забирали 20-50 клонов и рассевали в 24 луночный планшет.

7. Генотипировали клоны с помощью праймеров на вектор PB Cas9 PiggyBac и размножали отрицательные клоны.

Список литературы

Литературные источники обозначены в описании их номером и включены в описание как если бы они были полностью изложены в нем. Каждый из следующих источников включен в данное описание ссылкой в полном объеме.

10. 1. Carroll, D. (2011) Genome engineering with zinc-finger nucleases. *Genetics*, 188, 773-82.
2. Wood, A.J., Lo, T. -W., Zeitler, B., Pickle, C.S., Ralston, E.J., Lee, A.H., Amora, R., Miller, J.C., Leung, E., Meng, X., et al. (2011) Targeted genome editing across species using ZFNs and TALENs. *Science* (New York, N. Y.), 333, 307.
3. Perez-Pinera, P., Ousterout, D.G. and Gersbach, C. A. (2012) Advances in targeted genome editing. *Current opinion in chemical biology*, 16, 268-77.
4. Symington, L.S. and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annual review of genetics*, 45, 247-71.
5. Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D. and Holmes, M.C. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435, 646-51.
6. Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* (New York, N. Y.), 326, 1509-12.
7. Cell, P., Replacement, K.S., Talens, A., Type, A., Collection, C., Ccl-, A. and Quickextract, E. Genetic engineering of human pluripotent cells using TALE nucleases.
8. Mussolino, C., Morbitzer, R., Lutge, F., Dannemann, N., Lahaye, T. and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic acids research*, 39, 9283-93.
9. Ding, Q., Lee, Y., Schaefer, E.A.K., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T., et al. (2013) Resource A TALEN Genome-Editing System for Generating Human Stem Cell-Based Disease Models.
10. Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology*, 29, 731-4.
11. Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krag Ii, R.G., Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y. H., et al. (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature*, 490, 114-118.
12. Miller, J.C., Tan, S., Qiao, G., Barlow, K. a, Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011) A TALE nuclease architecture for efficient genome editing.
40. *Nature biotechnology*, 29, 143-8.
13. Holkers, M., Maggio, L., Liu, J., Janssen, J.M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T. and Goncalves, M. a F. V (2012) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic acids research*, 10. 1093/nar/gks1446.
45. 14. Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. a, Sander, J.D. and Joung, J.K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30, 460-465.
15. Qiu, P., Shandilya, H., D'Alessio, J.M., O'Connor, K., Durocher, J. and Gerard, G. F. (2004) Mutation detection using Surveyor nuclease. *BioTechniques*, 36, 702-7.

16. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science* (New York, N.Y.), 339, 823-6.
17. Ding, Q., Regan, S. N., Xia, Y., Oostrom, L.A., Cowan, C.A. and Musunura, K. (2013) Enhanced Efficiency of Human Pluripotent Stem Cell Genome Editing through Replacing TALENs with CRISPRs. *Cell Stem Cell*, 12, 393-394.
18. 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* (New York, N. Y.), 339, 819-23.
19. 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. *Nature biotechnology*, 31, 230-232.
20. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J. -RJ. and Joung, J. K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology*, 31, 227-229.
21. Chen, F., Praett-Miller, S.M., Huang, Y., Gjoka, M., Duda, K., Taunton, J., Collingwood, T.N., Frodin, M. and Davis, G.D. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature methods*, 8, 753-5.
22. Soldner, F., Laganiere, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.L., Myers, R.H., Lindquist, S., et al. (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, 146, 318-31.
23. Valamehr, B., Abujarour, R., Robinson, M., Le, T., Robbins, D., Shoemaker, D. and Flynn, P. (2012) A novel platform to enable the high-throughput derivation and characterization of feeder-free human iPSCs. *Scientific reports*, 2, 213.
24. Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G. and Zhang, F. (2012) A transcription activator-like effector toolbox for genome engineering. *Nature protocols*, 7, 171-92.
25. Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Iii, C. A. H., Smith, H. O. and America, N. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. 6, 12-16.
26. Zou, J., Maeder, M.L., Mali, P., Praett-Miller, S.M., Thibodeau-Beganny, S., Chou, B. -K., Chen, G., Ye, Z., Park, I. -H., Daley, G. Q., et al. (2009) Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell stem cell*, 5, 97-110.
27. Perez, E.E., Wang, J., Miller, J. C, Jouvenot, Y., Kim, K. a, Liu, O., Wang, N., Lee, G., Bartsevich, V.V, Lee, Y. -L., et al. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature biotechnology*, 26, 808-16.
28. Bhakta, M.S., Henry, I.M., OusteroutD. G., Das, K.T., Lockwood, S.H., Meckler, J.F., Wallen, M.C, Zykovich, A., Yu, Y., Leo, H., et al. (2013) Highly active zinc-finger nucleases by extended modular assembly. *Genome research*, 10. 1101/gr. 143693. 112.
29. Kim, E., Kim, S., Kim, D.H., Choi, B.-S., Choi, I.-Y. and Kim, J.-S. (2012) Precision genome engineering with programmable DNA-nicking enzymes. *Genome research*, 22, 1327-33.
30. Gupta, A., Meng, X., Zhu, L.J., Lawson, N. D. and Wolfe, S. a (2011) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic acids research*, 39, 381-92.
31. Park, I.-H., Lerou, P.H., Zhao, R., Huo, H. and Daley, G.Q. (2008) Generation of human-induced pluripotent stem cells. *Природа протоколов*, 3, 1180-6.
32. Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C, Baller, J.A., Somia, N.V, Bogdanove, A.J. and Voytas, D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*, 39, e82.
33. Briggs, A.W., Rios, X., Chari, R., Yang, L., Zhang, F., Mali, P. and Church, G.M. (2012)

- Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic acids research*, 10. 1093/nar/gks624.
34. Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M. and Arlotta, P. (2011) LETTERs Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. 29, 149-154.
- 5 35. Pathak, V.K. and Temin, H.M. (1990) Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 6019-23.
- 10 36. Tian, J., Ma, K. and Saaem, I. (2009) Advancing high-throughput gene synthesis technology. *Molecular bioSystems*, 5, 714-22.
37. Zou, J., Mali, P., Huang, X., Dowey, S. N. and Cheng, L. (2011) Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*, 118, 4599-608.
- 15 38. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., Dicarlo, J. E., Norville, J. E. and Church, G. M. (2013) RNA-Guided Human Genome.
39. Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S. and Crawford, G.E. (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell*, 132, 311-22.
- 20 40. Orlando, S.J., Santiago, Y., DeKelver, R.C., Freyvert, Y., Boydston, E. a, Moehle, E. a, Choi, V.M., Gopalan, S.M., Lou, J.F., Li, J., et al. (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic acids research*, 38, e152.
41. Wang, Z., Zhou, Z.-J., Liu, D.-P. and Huang, J.-D. (2008) Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. *Oligonucleotides*, 18, 21-32.
- 25 42. Rios, X., Briggs, A.W., Christodoulou, D., Gorham, J.M., Seidman, J. G. and Church, G. M. (2012) Stable gene targeting in human cells using single-strand oligonucleotides with modified bases. *PloS one*, 7, e36697.
- 30 43. Elliott, B., Richardson, C, Winderbaum, J., Jac, A., Jasin, M. and Nickoloff, J.A. C.A. (1998) Gene Conversion Tracts from Double-Strand Break Repair in Mammalian Cells Gene Conversion Tracts from Double-Strand Break Repair in Mammalian Cells. 18.
44. Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.-L., Kim, K. a, Ando, D., Urnov, F.D., Galli, C, Gregory, P.D., et al. (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature biotechnology*, 25, 1298-306.
- 35 45. 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* (New York, N. Y.), 337, 816-21.
- 40 46. Shrivastav, M., De Haro, L.P. and NickoloffJ. a (2008) Regulation of DNA double-strand break repair pathway choice. *Cell research*, 18, 134-47.
47. Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 1156-60.
- 45 48. Mimitou, E.P. and Symington, L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature*, 455, 7701.
49. Doyon, Y., Choi, V.M., Xia, D.F., Vo, T.D., Gregory, P.D. and Holmes, M.C. (2010) Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. *Nature methods*,

7,459-60.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Church, George

Yang, Luhan

5 Cargol, Marc

Yang, Joyce

<120> Genome Engineering

<130> 010498.00517

<140> PCT/US14/048140

10 <141> 2014-06-30

<150> US 61/858,866

<151> 2013-07-26

<160> 224

<170> PatentIn version 3.5

15 <210> 1

<211> 15

<212> DNA

<213> Artificial

<220>

20 <223> Oligonucleotide sequence

<400> 1

tggggcaaggc ttcttg 15

<210> 2

<211> 102

25 <212> DNA

<213> Artificial

<220>

<223> TALE template

<220>

30 <221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 2

ctaaccaggc agcaggtcggt ggcaatcgcc tccnnnnnng gcggaaaaca ggcattggaa 60

35 acagtacagc ggctgctgcc ggtgctgtgc caagcgcacg ga 102

<210> 3

<211> 102

<212> DNA

<213> Artificial

40 <220>

<223> PCR primer

<220>

<221> misc_feature

<222> (34)..(39)

45 <223> n is a, c, g, or t

<400> 3

ctaaccctg aacaggtagt cgctatacgct tcannnnnng ggggcaagca agcacttgag 60

accgttcaac gactcctgcc agtgctctgc caagccatg ga 102

```

<210> 4
<211> 102
<212> DNA
<213> Artificial
5 <220>
<223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
10 <223> n is a, c, g, or t
<400> 4
ttgactccgg agcaagtcgt cgcgatcgcg agcnnnnnng ggggaaagca ggctggaa 60
actgttcaga gactgctgcc tgtacttgtt caggcgcata gt 102
<210> 5
15 <211> 102
<212> DNA
<213> Artificial
<220>
<223> PCR primer
20 <220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
<400> 5
25 ctcacccccc aacaggttgt cgcaatagca agtnnnnnng gcggtaagca agccctagag 60
actgtgcaac gcctgctccc cgtgctgtgt caggctcacg gt 102
<210> 6
<211> 102
<212> DNA
30 <213> Artificial
<220>
<223> PCR primer
<220>
<221> misc_feature
35 <222> (34)..(39)
<223> n is a, c, g, or t
<400> 6
ctgacacacctg aacaagttgt cgcgatagcc agtnnnnnng ggggaaaaca agctctagaa 60
acggttcaaa ggttgttgcc cgttctgtgc caagcacatg gg 102
40 <210> 7
<211> 102
<212> DNA
<213> Artificial
<220>
45 <223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)

```

```

<223> n is a, c, g, or t
<400> 7
ttaacaccccg aacaagtagt agcgatagcg tcannnnnng gggtaaaca ggctttggag 60
acgtacagc ggattattgcc ggtcctctgc caggcccacg ga 102
5 <210> 8
<211> 102
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
15 <400> 8
cttacgccag aacaggtggt tgcaattgcc tccnnnnnng gcgggaaaca agcgttggaa 60
actgtgcaga gactccttcc tgtttgtgt caagcccacg gc 102
<210> 9
<211> 102
20 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<220>
25 <221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
<400> 9
ttgacgcctg agcaggttgt ggccatcgct agcnnnnnng gagggaaagca ggctttgaa 60
30 accgtacagc gacttctccc agtttgtgc caagtcacg gg 102
<210> 10
<211> 102
<212> DNA
<213> Artificial
35 <220>
<223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
40 <223> n is a, c, g, or t
<400> 10
ctaaccaccccg agcaagtagt tgccatagca agcnnnnnng gagaaaaca ggcattagaa 60
acagttcagc gcttgctccc ggtactctgt caggcacacg gt 102
<210> 11
45 <211> 102
<212> DNA
<213> Artificial
<220>

```

<223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
5 <223> n is a, c, g, or t
<400> 11
ctaaccccg aacaggtcgt agccattgct tccnnnnnng gcggcaaaca ggcttagag 60
acagtccaga ggctttgcc tgtttatgc caggcacatg gc 102
<210> 12
10 <211> 102
<212> DNA
<213> Artificial
<220>
<223> PCR primer
15 <220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
<400> 12
20 ctcaccccg agcaggtcgt tgccatcgcc agtnnnnnng gcggaaagca agctctcgaa 60
acagtacaac ggctgttgcc agtcctatgt caagctcatg ga 102
<210> 13
<211> 102
<212> DNA
25 <213> Artificial
<220>
<223> PCR primer
<220>
<221> misc_feature
30 <222> (34)..(39)
<223> n is a, c, g, or t
<400> 13
ctgacgccc agcaggtagt ggcaatcgca tctnnnnnng gaggtaaaca agcactcgag 60
actgtccaaa gattgttacc cgtactatgc caagcgcatt gt 102
35 <210> 14
<211> 102
<212> DNA
<213> Artificial
<220>
40 <223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
45 <400> 14
ttaaccccg agcaagttgt ggctattgca tctnnnnnng gtggcaaaca agccttggag 60
acagtgcac gattactgcc tgtttatgt caggccccatg gc 102
<210> 15

<211> 102
<212> DNA
<213> Artificial
<220>
5 <223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
10 <400> 15
cttactcctg agcaagtcgt agctatcgcc agcnnnnnng gtggaaaca ggccctggaa 60
accgtacaac gtctcctccc agtactttgt caagcacacg gg 102
<210> 16
<211> 102
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<220>
20 <221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t
<400> 16
ttgacacccgg aacaagtggt ggcgattgcg tccnnnnnng gaggaagca ggcactggag 60
25 accgtccaac ggcttcttcc ggttcttgc caggctcatg gg 102
<210> 17
<211> 102
<212> DNA
<213> Artificial
30 <220>
<223> PCR primer
<220>
<221> misc_feature
<222> (34)..(39)
35 <223> n is a, c, g, or t
<400> 17
ctcacgccag agcaggtggt agcaatagcg tcgnnnnnng gtggtaagca agcgcttgaa 60
acgggtccagc gtcttctgcc ggtgttgtgc caggcgcacg ga 102
<210> 18
40 <211> 102
<212> DNA
<213> Artificial
<220>
<223> PCR primer
45 <220>
<221> misc_feature
<222> (34)..(39)
<223> n is a, c, g, or t

<400> 18
 ctcacaccag aacaagtggt tgctattgct agtnnnnnng gtggaaagca ggccctcgag 60
 acgggtgcaga gtttacttcc cgtcctctgt caagcgcacg gc 102
 <210> 19
 5 <211> 60
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR primer
 10 <220>
 <221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 19
 15 ctcactccag agcaagtggt tgcgatcgct tcannnnnng gtggaaagacc tgccctggaa 60
 <210> 20
 <211> 5714
 <212> DNA
 <213> Artificial
 20 <220>
 <223> Oligonucleotide sequence encoding TALEN backbone
 <400> 20
 atgtcgcgga cccggctccc ttccccaccc gcacccagcc cagcgtttc ggccgactcg 60
 ttctcagacc tgcttaggca gttcgacccc tcactgtta acacatcggt gttcgactcc 120
 25 cttcctccgt ttggggcgca ccatacgagg gcggccaccc gggagtggga tgaggtgcag 180
 tcgggattga gagctcgcgga tgcaccaccc ccaaccatgc ggggtggccgt caccgctgcc 240
 cgaccgcccga gggcgaagcc cgacccaagg cggagggcag cgcaaccgtc cgacgcaagc 300
 cccgcagcgc aagtagatt gagaactttg gatatattcac agcagcagca ggaaaagatc 360
 aagcccaaag tgaggtcgac agtcgcgcag catcacgaag cgctgggtggg tcatgggttt 420
 30 acacatgccc acatcgtagc cttgtcgac caccctgcag cccttggcac ggtcgccgtc 480
 aagtaccagg acatgattgc ggcgttgccc gaagccacac atgaggcgat cgtcggtgtg 540
 gggaaacagt ggagcggagc ccgagcgctt gaggccctgt tgacggcgc gggagagctg 600
 agagggcctc cccttcagct ggacacgggc cagttgctga agatcgcgaa gcggggagga 660
 gtcacggcggt tcgaggcggt gcacgcgtgg cgcaatgcgc tcacggagc acccctcaac 720
 35 agttcacgct gacagagacc gcggccgcatt taggcacccccc aggcttaca ctatgtctt 780
 ccggctcgta taatgtgtgg attttgagtt agatccgtc gagatttca ggagctaagg 840
 aagctaaaat ggagaaaaaaa atcactggat ataccaccgt tgatatatcc caatggcatc 900
 gtaaagaaca ttttgaggca tttcagtcag ttgctcaatg tacctataac cagaccgttc 960
 agctggatatacggccttt taaaagaccg taaagaaaaaa taagcacaag ttttatccgg 1020
 40 cctttattca cattcttgcc cgccctgatga atgctcatcc ggaattccgt atggcaatga 1080
 aagacgggtga gctgggtata tgggatagtg ttcacccttg ttacaccgtt ttccatgagc 1140
 aaactgaaac gtttcatcg ctctggagtg aataccacga cgattccgg cagtttctac 1200
 acatatatttc gcaagatgtg gcgtgttacg gtgaaaacct ggcctatttc cctaaagggt 1260
 ttattgagaa tatgttttc gtctcagcca atccctgggt gagttcacc agtttgatt 1320
 45 taaacgtggc caatatggac aacttcttcg ccccccgtttt caccatgggc aaatattata 1380
 cgcaaggcgca caaggtgctg atgcccgtgg cgattcaggt tcatcgatcc gtttgtatg 1440
 gcttccatgt cggcagaatg cttaatgaat tacaacagta ctgcgtatgag tggcagggcg 1500
 gggcgtaaag atctggatcc ggcttactaa aagccagata acagtatgcg tatttgcgcg 1560

ctgatttttgcggataaga atatataactg atatgtatacccgaaagtatgtcaaaaagag 1620
 gtatgtatg aagcagcgta ttacagtgc acgttgacacgc gacagctatc agttgctcaa 1680
 ggcataatatgtcaaatatctccggctgt gtaaggacaa ccatgcagaa tgaagccgt 1740
 cgtctgcgtg ccgaacgctg gaaagcggaa aatcaggaag ggatggctga ggtcgcccgg 1800
 5 tttattgaaa tgaacggctc ttttgcgtac gagaacagggg gctggtaaa tgcatgtt 1860
 ggtttacacc tataaaagag agagccgtt tcgtctgtt gtggatgtac agagtatat 1920
 tattgacacg cccgggac ggtatggat cccccctggcc agtgcacgtc tgctgtcaga 1980
 taaagtctcc cgtgaactt acccggtggt gcatatcggg gatgaaagct ggcgcatgtat 2040
 gaccaccat atggccagtg tgccggctc cgttatcggg gaagaagtgg ctgatctcag 2100
 10 ccaccgcgaa aatgacatca aaaacgcctt taacctgtatgttctgggaa tataaatgtc 2160
 aggctccctt atacacagcc agtctgcagg tcgacggctc cgctcttcga aggttacttc 2220
 ccgtccctgt tcaagcgcac ggcctcactc cagagcaagt ggttgcgatc gcttcaaaca 2280
 acgggtggaa acctgcctg gaatcaatcg tggcccagct ttcgaggccg gacccgcgc 2340
 tggccgcact cactaatgtatcatcttgcgtatcgctggcctgcctggcggac 2400
 15 tggatgcggtaaagaagggg ctccgcacg cgcctgcatt gattaagcgg accaacagaa 2460
 ggattcccgaa gaggacatca catcgagtgg caggttccca actcgtgaag agtgaacttg 2520
 aggagaaaaaa gtcggagctg cggcacaaat taaaatacgt accgcataac tacatcgaac 2580
 ttatcgaaat tgcttaggaac tcgactcaag acagaatcct tgagatgaag gtaatggagt 2640
 tctttatgaa ggttatgaa taccgaggga agcatctcgg tggatcacga aaacccgacg 2700
 20 gagcaatcta tacgggtggg agcccgattt attacggatg gatcgtcgac acgaaagcct 2760
 acagcggtgg gtacaatctt cccatcgggc aggcatgatg gatcacaatgt tatgtcgaag 2820
 aaaatcagac caggaacaaa cacatcaatc caaatgatgt gtggaaatgt tatccttcat 2880
 cagtgaccga gtttaagttt ttgttgtct ctggcattt caaaggcaac tataaggccc 2940
 agctcacacg gttgaatcac attacgaact gcaatggtgc ggtttgtcc gtagaggaac 3000
 25 tgctcattgg tggagaaatg atcaaagcgg gaaactctgac actggaaagaa gtcagacgca 3060
 agtttaacaa tggcgagatc aatttccgca agctaaaatg gaaaaaaa tcactggata 3120
 taccaccgtt gatataatccc aatggcatcg taaagaacat tttgaggcat ttcagtcagt 3180
 tgctcaatgt acctataacc agaccgttca gctggatatt acggccttt taaagaccgt 3240
 aaagaaaaat aagcacaagt tttatccggc ctttatttcac attcttgccc gcctgatgaa 3300
 30 tgctcatccg gaattccgta tggcaatgaa agacggtgag ctggtgatatt gggatagtgt 3360
 tcacccttgt tacaccgttt tccatgagca aactgaaacg ttttcatcg tctggagtga 3420
 ataccacgac gatttccggc agtttctaca catatattcg caagatgtgg cgtgttacgg 3480
 tggaaacctg gccttattcc ctaaagggtt tattgagaat atgttttgc tctcagccaa 3540
 tccctgggtg agtttccacca gtttgattt aaacgtggcc aatatggaca acttcttcgc 3600
 35 ccccgtttc accatgggca aatattatac gcaaggcgac aaggtgctga tgccgctggc 3660
 gattcaggtt catcatgccg tttgtgatgg cttccatgtc ggcagaatgc ttaatgaatt 3720
 acaacagttac tgcgtatgatg ggcaggccgg ggcgtaaaga tctggatccg gcttactaaa 3780
 agccagataa cagtatgcgt atttgcgcgc tgattttgc ggtataagaa tatataactga 3840
 tatgtataacc cgaagttatgt caaaaagagg tatgctatga agcagcgtat tacagtgaca 3900
 40 gttgacagcg acagctatca gttgctcaag gcatatatac tgtcaatatc tccggctctgg 3960
 taagcacaac catgcagaat gaagccgcgtc gtctgcgtgc cgaacgcgtgg aaagcggaaa 4020
 atcaggaagg gatggctgag gtcgcccgtt ttattgaaat gaacggctct tttgctgacg 4080
 agaacagggg ctgggtaaat gcagtttaag gtttacacct ataaaagaga gagccgttat 4140
 cgtctgtttg tggatgtaca gagtgatatt attgacacgc ccggcgacg gatggtgatc 4200
 45 cccctggcca gtgcacgtct gctgtcagat aaagtctccc gtgaacttta cccgggtgg 4260
 catatcgggg atgaaagctg ggcgcataatgc accaccgata tggccagtgt gccggctcc 4320
 gttatcgggg aagaagtggc tgatctcagc caccgcgaaa atgacatcaa aaacgcatt 4380
 aacctgatgt tctggaaat ataaatgtca ggctccctt tacacagcca gtctgcaggt 4440

cgacggcttc gctttcgaa ggttacttcc cgtcctctgt caagcgcacg gcctcactcc 4500
 agagcaagtg gttgcgatcg cttcaaaca cggtggaaaga cctgccctgg aatcaatcgt 4560
 ggcggcgtt tcgaggccgg accccgcgct ggccgcactc actaatgatc atctttagtc 4620
 gctggcctgc ctcggcggac gaccgcctt ggatgcggtg aagaaggggc tcccgacgc 4680
 5 gcctgcattt attaagcggg ccaacagaag gattcccgag aggacatagc cccaagaaga 4740
 agagaaagggt ggaggccagc ggttccggac gggctgacgc attggacgt tttgatctgg 4800
 atatgctggg aagtgacgcc ctcgatgatt ttgaccttga catgcttggt tcggatgcc 4860
 ttgatgactt tgacctcgac atgctcgca gtgacgcct tcatgatttc gacctggaca 4920
 tgctgattaa ctctagaggc agtggagagg gcagaggaag tctgctaaca tgcggtgacg 4980
 10 tcgaggagaa tcctggccca gtgagcaagg gcgaggagct gttcacccgg gtgggtcccc 5040
 tcctggtcga gctggacggc gacgtaaacg gccacaagtt cagcgtgtcc ggcgaggggcg 5100
 agggcgatgc cacctacggc aagctgaccc tgaagttcat ctgcaccacc ggcaagctgc 5160
 ccgtgcctg gcccaccctc gtgaccaccc tgacctacgg cgtgcagtgc ttcaagccgct 5220
 accccgacca catgaagcag cacgacttct tcaagtccgc catgcccggaa ggctacgtcc 5280
 15 aggagcgcac catcttcttc aaggacgacg gcaactacaa gacccgcgc gaggtgaagt 5340
 tcgaggcgca caccctggtg aaccgcacg agctgaaggg catcgacttc aaggaggacg 5400
 gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc tatatcatgg 5460
 ccgacaagca gaagaacggc atcaaggta acttcaagat ccggccacaac atcgaggacg 5520
 gcagcgtgca gctcgccgac cactaccaggc agaacaccccatcgac ggccccgtgc 5580
 20 tgctgcccga caaccactac ctgagcaccc agtccgcctt gagcaaagac cccaacgaga 5640
 agcgcgatca catggtcctg ctggagttcg tgaccgcccgc cgggatcaact ctcggcatgg 5700
 acgagctgta caag 5714
 <210> 21
 <211> 3465
 25 <212> DNA
 <213> Artificial
 <220>
 <223> Oligonucleotide sequence encoding TALE backbone sequence
 <400> 21
 30 atgtcgccga cccggctccc ttccccaccc gcacccagcc cagcgtttc ggccgactcg 60
 ttctcagacc tgcttaggca gttcgacccccc tcactgtttt acacatcggtt gttcgactcc 120
 cttcctccgt ttggggcgca ccatacggag gcccgcaccg gggagtgggta tgaggtgcag 180
 tcgggattga gagctcgccga tgcaccaccc ccaaccatgc gggtgccgtt caccgctgcc 240
 cgaccgcccga gggcgaagcc cgcaccaagg cggaggccag cgcaaccgtc cgacgcaagc 300
 35 cccgcagcgc aagtagattt gagaactttt ggtatattcac agcagcagca ggaaaagatc 360
 aagcccaaag tgaggtcgac agtcgcgcag catcacgaag cgctgggtgg tcatgggttt 420
 acacatgccc acatcgtagc cttgtcgccag caccctgcag cccttggcac ggtcgccgtc 480
 aagtaccagg acatgattgc ggcttgccg gaagccacac atgaggcgat cgtcggtgt 540
 gggaaacagt ggagcggagc ccgagcgctt gaggccctgt tgacggtcgc gggagagctg 600
 40 agagggccctc cccttcagct ggacacgggc cagttgtca agatcgcaaa gcggggagga 660
 gtcacggccgg tcgaggccgg gcacgcgtgg cgcaatgcgc tcacgggagc acccctcaac 720
 agttcacgct gacagagacc gcccgcac taggcacccccc aggcttaca ctttatgctt 780
 ccggctcgta taatgtgtgg attttgagtt aggatccgtc gagatttca ggagctaaagg 840
 aagctaaaat ggagaaaaaaa atcactggat ataccaccgt tcatatatcc caatggcatc 900
 45 gtaaagaaca ttttgaggca tttcagtcag ttgctcaatg tacctataac cagaccgttc 960
 agctggatatac cttttttttt ttaaagaccg taaagaaaaaa taagcacaag ttttatccgg 1020
 cttttattca cattcttgcc cgcctgatga atgctcatcc ggaattccgt atggcaatga 1080
 aagacgggtga gctgggtgata tgggatagtgt ttcacccttg ttacaccgtt ttccatgagc 1140

aaactgaaac gtttcatcg ctctggagtg aataccacga cgattccgg cagtttctac 1200
 acatatattc gcaagatgtg gcgtgttacg gtaaaaacct ggcctatttc cctaaagggt 1260
 ttattgagaa tatgttttc gtctcagcca atcccgggt gagttcacc agtttgatt 1320
 taaacgtggc caatatggac aacttctcg cccccgttt caccatgggc aaatattata 1380
 5 cgcaaggcga caaggtgctg atgccgctgg cgattcaggt tcatcatgcc gtttgtatg 1440
 gcttccatgt cggcagaatg cttaatgaat tacaacagta ctgcgatgag tggcaggcgc 1500
 gggcgtaaag atctggatcc ggcttactaa aagccagata acagtatgcg tatttgcgcg 1560
 ctgattttg cggataaga atatatactg atatgtatac ccgaagtgatg tcaaaaagag 1620
 gtatgctatg aagcagcgt a tacagtgac agttgacagc gacagctatc agttgctcaa 1680
 10 ggcataatag atgtcaaatat ctccggctcg gtaagcacaa ccatgcagaa tgaagccgt 1740
 cgtctcgctg cccaaacgctg gaaagcggaa aatcaggaag ggtggctga ggtcgccgg 1800
 tttattgaaa tgaacggctc ttttgctgac gagaacaggg gctggtgaaa tgcagttaa 1860
 ggtttacacc tataaaagag agagccgtt tcgtctgtt gtggatgtac agagtgatat 1920
 tattgacacg cccggggcgcac ggatggtgat cccctggcc agtgcacgtc tgctgtcaga 1980
 15 taaagtctcc cgtgaacttt accccgggt gcatatcggg gatgaaagct ggccatgtat 2040
 gaccaccgat atggccagtg tgccggctc cgttatcggg gaagaagtgg ctgatctcag 2100
 ccaccgcgaa aatgacatca aaaacgccat taacctgatg ttctgggaa tataaatgtc 2160
 aggctccctt atacacagcc agtctgcagg tcgacggctc cgctcttcga aggttacttc 2220
 ccgtccctcg tcaagcgcac ggcctcactc cagagcaagt ggttgcgatc gcttcaaaca 2280
 20 acggtgaaag acctgcccctg gaatcaatcg tggcccagct ttcgaggccg gacccgcgc 2340
 tggccgcact cactaatgat catttgttag cgctggctc cctcggcggc cgacccgcct 2400
 tggatgcggcgt gaagaagggg ctcccgacg cgcctgcatt gattaagcgg accaacagaa 2460
 ggattcccgaa gaggacatag ccccaagaag aagagaaagg tggaggccag cggttccgga 2520
 cgggctgacg cattggacgaa tttgatctg gatatgctgg gaagtgcgc cctcgatgtat 2580
 25 tttgaccttg acatgcttg ttccggatgcc cttgatgact ttgacctcgatc catgctcgcc 2640
 agtgacgccc ttgatgattt cgacctggac atgctgatta actctagagg cagtggagag 2700
 ggcagaggaa gtctgctaac atgcgggtac gtcgaggaga atcctggccc agtgagcaag 2760
 ggcgaggcgcgt tgccaccgg ggtggtgccc atcctggctg agctggacgg cgacgtaaac 2820
 ggccacaagt tcagcgtgtc cggcgaggc gaggcgcgtt ccacctacgg caagctgacc 2880
 30 ctgaagttca tctgcaccac cggcaagctg cccgtgcctt ggcccaccct cgtgaccacc 2940
 ctgacccatcg gcgtgcagtg cttcagccgc taccggcacc acatgaagca gcacgacttc 3000
 ttcaagtccg ccatgcccga aggctacgtc caggagcgc ccatcttctt caaggacgac 3060
 ggcaactaca agacccgcgc cgaggtaag ttccggccg acaccctggt gaaccgcac 3120
 gagctgaagg gcatcgactt caaggaggac ggcaacatcc tggggcacaa gctggagttac 3180
 35 aactacaaca gccacaacgt ctatcatg gccgacaagc agaagaacgg catcaaggtg 3240
 aacttcaaga tccgcccaca catcgaggac ggcagcgtgc agctcgccga ccactaccag 3300
 cagaacaccc ccatcggcga cggcccccgtg ctgctgccc acaaccacta cctgagcacc 3360
 cagtccggccc tgagcaaaga ccccaacgag aagcgcgc acatggctt gctggagttc 3420
 gtgaccggccg cgggatcac tctcggcatg gacgagctgt acaag 3465
 40 <210> 22
 <211> 1368
 <212> PRT
 <213> Streptococcus pyogenes
 <400> 22
 45 Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
 1 5 10 15
 Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

RU 2764757 C2

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45
 Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60
 5 Lys Arg Thr Ala Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80
 Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95
 Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 10 100 105 110
 His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125
 His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140
 15 Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160
 Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175
 Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 20 180 185 190
 Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205
 Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220
 25 Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 225 230 235 240
 Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255
 Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 30 260 265 270
 Asp Asp Leu Asp Asn Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300
 35 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 40 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 45 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 5 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480
 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 10 485 490 495
 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510
 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525
 15 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540
 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560
 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 20 565 570 575
 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605
 25 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 30 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685
 35 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720
 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 40 725 730 735
 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750
 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765
 45 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780
 Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800

RU 2764757 C2

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830
5 Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
10 865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910
15 Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
20 945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990
25 Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
30 1025 1030 1035
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065
35 Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
1085 1090 1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
40 1100 1105 1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140
45 Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185
 Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200
 5 Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215
 Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230
 Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 10 1235 1240 1245
 Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260
 His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275
 15 Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290
 Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305
 Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 20 1310 1315 1320
 Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335
 Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350
 25 Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365
 <210> 23
 <211> 82
 <212> DNA
 30 <213> Artificial
 <220>
 <223> Oligonucleotide sequence
 <220>
 <221> misc_feature
 35 <222> (41)..(41)
 <223> n is a, c, g, or t
 <400> 23
 ttcttggc ttatatatctt gtggaaagga cgaaacacccg ngtttagag ctagaaatag 60
 caagttaaaa taaggctagt cc 82
 40 <210> 24
 <211> 1692
 <212> DNA
 <213> Artificial
 <220>
 45 <223> Oligonucleotide sequence encoding re-TALE
 <400> 24
 ctaaccctg aacaggttgt cgctatacg tcaaatacg gggcaagca agcactttag 60
 accgttcaac gactcctgcc agtgctctgc caagccatg gattgactcc ggagcaagtc 120

gtcgcgatcg cgagcaacgg cggggggaag caggcgctgg aaactgttca gagactgctg 180
 cctgtacttt gtcaggcgca tggtctcacc cccgaacagg ttgtcgcaat agcaagtaat 240
 ataggcggta agcaagccct agagactgtg caacgcctgc tccccgtgct gtgtcaggct 300
 cacggctcta 5 cacctgaaca agttgtcgct atagccagtc acgacggggg aaaacaagct 360
 cttagaaacgg ttcaaagggtt gttcccgtt ctgtgccaag cacatgggtt aacacccgaa 420
 caagtagtag cgatagcgctc aaataacggg ggtaaacacagg ctttgagac ggtacagcg 480
 ttattgccgg tcctctgcca ggcccacgga cttacgcccag aacaggtgg 540
 tccaacatcg gcgggaaaca agcgttgaa actgtgcaga gactcctcc tgaaaaatgtgt 600
 caagcccacg gcttgacgccc tgagcagggtt gtggccatcg ctagccacga cggagggaaag 660
 10 caggctcttg aaaccgtaca gcgacttctc ccagttttgt gccaaagctca cgggctaacc 720
 cccgagcaag tagttgccat agcaagcaac ggaggaggaa aacaggcatt agaaacagtt 780
 cagcgcttgc tcccgtact ctgtcaggca cacggtctaa ctccggaaaca ggtcgtagcc 840
 attgcttccc atgatggcgg caaacaggcg ctagagacag tccagaggct cttgcctgtg 900
 ttatgccagg cacatggctt caccggag caggtcggtt ccacatcgccag taatatcgcc 960
 15 ggaaagcaag ctctcgaaaac agtacaacgg ctgttgccag tcctatgtca agctcatgga 1020
 ctgacgcccc agcaggtgtt ggcaatcgca tctcacatgtt gaggtaaaca agcactcgag 1080
 actgtccaaa gattgttacc cgtactatgc caagcgcgtt gtttaacccc agagcaagtt 1140
 gtggctattt catctaacgg cggtggcaaa caagccttgg agacagtgcac acgattactg 1200
 cctgtcttat gtcaggcccc tggccttact cctgagcaag tcgttagctat cgccagcaac 1260
 20 ataggtggaa aacaggccct ggaaaccgta caacgtctcc tcccgtact ttgtcaagca 1320
 cacgggttga caccggaaaca agtggtggcg attgcgtcca acggcggagg caagcaggca 1380
 ctggagaccg tccaacggct tcttccgggtt ctttgccagg ctcatggct cacggccagag 1440
 caggtggtag caatagcgctc gaacatcggtt ggtaagcaag cgcttggaaac ggtccagcgt 1500
 cttctgcccgg tggtggccca ggcgcacgga ctcacaccag aacaagtgg 560 tgctattgct 1560
 25 agtaacaacg gtggaaagca ggccctcgag acgggtgcaga ggttacttcc cgtcctctgt 1620
 caagcgcacg gcctcactcc agagcaagtg gttgcgatcg cttcaaaca 1680
 cctgccctgg aa 1692
 <210> 25
 <211> 237
 30 <212> DNA
 <213> Artificial
 <220>
 <223> Template sequence
 <220>
 35 <221> misc_feature
 <222> (64)..(69)
 <223> n is a, c, g, or t
 <220>
 <221> misc_feature
 40 <222> (166)..(171)
 <223> n is a, c, g, or t
 <400> 25
 cgcaatgcgc tcacgggagc acccctcaac ctaacccctg aacaggtgtt cgctatacg 60
 tcannnnnng ggggcaagca agcacttgag accgttcaac gactcctgcc agtgctctgc 120
 45 caagcccatg gattgactcc ggagcaagtc gtcgcgatcg cgagcnnnnn ngggggggaaag 180
 cagggcgtgg aaactgttca gagactgctg cctgtacttt gtcaggcgca tggtctc 237
 <210> 26
 <211> 240

```

<212> DNA
<213> Artificial
<220>
<223> Template sequence
5 <220>
<221> misc_feature
<222> (67)..(72)
<223> n is a, c, g, or t
<220>
10 <221> misc_feature
<222> (169)..(174)
<223> n is a, c, g, or t
<400> 26
agactgctgc ctgtactttg tcaggcgcat ggtctcaccc ccgaacaggt tgtcgcaata 60
15 gcaagtnnnn nnggcggtaa gcaagcccta gagactgtgc aacgcctgct ccccgtgctg 120
tgtcaggctc acggtctgac acctgaacaa gttgtcgcga tagccagtnn nnnnggggga 180
aaacaagctc tagaaacggt tcaaagggtt ttgcccgttc tgtgccaagc acatgggtta 240
<210> 27
<211> 232
20 <212> DNA
<213> Artificial
<220>
<223> Template sequence
<220>
25 <221> misc_feature
<222> (59)..(64)
<223> n is a, c, g, or t
<220>
<221> misc_feature
30 <222> (161)..(166)
<223> n is a, c, g, or t
<400> 27
tgcgctcacf ggagcacccc tcaacctcac ccccgaacag gttgtcgcaa tagcaagtnn 60
nnnnngcggtaa aagcaagccc tagagactgt gcaacgcctg ctccccgtgc tgtgtcaggc 120
35 tcacggctcg acacctgaac aagttgtcgc gatagccagt nnnnnngggg gaaaacaagc 180
tctagaaaacg gttcaaagggt tggccgt tctgtgccaa gcacatgggt ta 232
<210> 28
<211> 246
<212> DNA
40 <213> Artificial
<220>
<223> Template sequence
<220>
<221> misc_feature
45 <222> (67)..(72)
<223> n is a, c, g, or t
<220>
<221> misc_feature

```

```

<222> (169)..(174)
<223> n is a, c, g, or t
<400> 28
    aggttggc ccgttctgtg ccaagcacat gggtaaacac ccgaacaagt agtagcgata 60
    5      gcgtcannnn nnngggtaa acaggcttg gagacggtac agcggttatt gccggccctc 120
        tgccaggccc acggacttac gccagaacag gtggttgcaa ttgcctccnn nnnnggcggg 180
        aaacaagcgt tggaaactgt gcagagactc cttcctgtt tgtgtcaagc ccacggctt 240
        acgcct 246
    <210> 29
    10     <211> 240
    <212> DNA
    <213> Artificial
    <220>
    <223> Template sequence
    15     <220>
    <221> misc_feature
    <222> (67)..(72)
    <223> n is a, c, g, or t
    <220>
    20     <221> misc_feature
    <222> (169)..(174)
    <223> n is a, c, g, or t
    <400> 29
    agactccttc ctgttttgtg tcaagccac ggcttgacgc ctgagcagggt tgtggccatc 60
    25      gctagcnnnn nnngagggaa gcaggctttt gaaaccgtac agcgacttct cccagtttg 120
        tgccaagctc acgggctaac ccccgagcaa gttagtgcctt tagcaagcnn nnnnggagga 180
        aaacaggcat tagaaacagt tcagcgcttgc ctcccggtac tctgtcaggc acacggctta 240
    <210> 30
    <211> 240
    30     <212> DNA
    <213> Artificial
    <220>
    <223> Template sequence
    <220>
    35     <221> misc_feature
    <222> (67)..(72)
    <223> n is a, c, g, or t
    <220>
    <221> misc_feature
    40     <222> (169)..(174)
    <223> n is a, c, g, or t
    <400> 30
    cgcttgctcc cggtaactctg tcagggcacac ggtctaactc cggaacagggt cgtagccatt 60
    gcttcnnnn nnngcggcaa acaggcgctt gagaccgtcc agaggctttt gcctgtgtta 120
    45      tgccaggcac atggcctcac cccggagcag gtcgttgccctt tcgcccgtt nnnnggcgg 180
        aagcaagctc tcgaaacagt acaacggctg ttgccagttcc tatgtcaagc tcatggactg 240
    <210> 31
    <211> 240

```

```

<212> DNA
<213> Artificial
<220>
<223> Template sequence
5 <220>
<221> misc_feature
<222> (67)..(72)
<223> n is a, c, g, or t
<220>
10 <221> misc_feature
<222> (169)..(174)
<223> n is a, c, g, or t
<400> 31
cggctgttgc cagtcctatg tcaagctcat ggactgacgc ccgagcaggt agtggcaatc 60
15 gcatctnnnn nnggaggtaa acaagcactc gagactgtcc aaagattgtt acccgtacta 120
tgccaagcgc atggtttaac cccagagcaa gttgtggcta ttgcatctnn nnnnggtggc 180
aaacaagcct tggagaccgt gcaacgatta ctgcctgtct tatgtcaggc ccatggcctt 240
<210> 32
<211> 240
20 <212> DNA
<213> Artificial
<220>
<223> Template sequence
<220>
25 <221> misc_feature
<222> (67)..(72)
<223> n is a, c, g, or t
<220>
<221> misc_feature
30 <222> (169)..(174)
<223> n is a, c, g, or t
<400> 32
cgattactgc ctgtctttag tcagggccat ggccttactc ctgagcaggt ggtcgctatc 60
gccagcnnnn nngggggcaa gcaaggactg gaaacagtcc agcgttgct tccagtactt 120
35 tgtcaggcgc atggattgac accggaacaa gtgggtggcta tagcctcann nnnnggagga 180
aagcaggcgc tggaaaccgt ccaacgtctt ttaccggcgc tttgccaggc gcacgggctc 240
<210> 33
<211> 240
<212> DNA
40 <213> Artificial
<220>
<223> Template sequence
<220>
<221> misc_feature
45 <222> (67)..(72)
<223> n is a, c, g, or t
<220>
<221> misc_feature

```

<222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 33
 cgattactgc ctgtctttag tcagggccat ggccttactc ctgagcaagt cgtagctatc 60
 5 gccagcnnnn nnggtggaa acaggccctg gaaaccgtac aacgtctcct cccagtactt 120
 tgtcaagcac acgggttgac accggaacaa gtgggtggcga ttgcgtccnn nnnnggaggc 180
 aagcaggcac tggagaccgt ccaacggctt cttccggttc tttgccaggc tcatgggctc 240
 <210> 34
 <211> 240
 10 <212> DNA
 <213> Artificial
 <220>
 <223> Template sequence
 <220>
 15 <221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220>
 <221> misc_feature
 20 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 34
 cggcttcttc cggttctttg ccaggctcat gggctcacgc cagagcaggt ggtagcaata 60
 gcgtcgnnnn nnggtggtaa gcaagcgctt gaaacggtcc agcgtcttct gccggtgttg 120
 25 tgccaggcgc acggactcac accagaacaa gtggttgcta ttgctagtnn nnnnggtgga 180
 aagcaggccc tcgagacggt gcagaggtta cttcccggtcc tctgtcaagc gcacggcctc 240
 <210> 35
 <211> 49
 <212> DNA
 30 <213> Artificial
 <220>
 <223> Primer sequence
 <400> 35
 cgcaatgcgc tcacgggagc acccctcaac ctaaccctg aacaggtag 49
 35 <210> 36
 <211> 46
 <212> DNA
 <213> Artificial
 <220>
 40 <223> Primer sequence
 <400> 36
 gagaccatgc gcctgacaaa gtacaggcag cagtctctga acagtt 46
 <210> 37
 <211> 33
 45 <212> DNA
 <213> Artificial
 <220>
 <223> Primer sequence

<400> 37
tggcgcaatg cgctcacggg agcaccctc aac 33
<210> 38
<211> 49
5 <212> DNA
<213> Artificial
<220>
<223> Primer sequence
<400> 38
10 agactgctgc ctgtactttg tcaggcgcat ggtctcaccc ccgaacagg 49
<210> 39
<211> 46
<212> DNA
<213> Artificial
15 <220>
<223> Primer sequence
<400> 39
taacccatgt gcttggcaca gaacgggcaa caaccttga accgtt 46
<210> 40
20 <211> 48
<212> DNA
<213> Artificial
<220>
<223> Primer sequence
25 <400> 40
aggttgtgc ccgttctgtg ccaagcacat gggtaaacac ccgaacaa 48
<210> 41
<211> 52
<212> DNA
30 <213> Artificial
<220>
<223> Primer sequence
<400> 41
aggcgtcaag ccgtgggctt gacacaaaac aggaaggagt ctctgcacag tt 52
35 <210> 42
<211> 45
<212> DNA
<213> Artificial
<220>
40 <223> Primer sequence
<400> 42
agactccttc ctgttttgtg tcaagccac ggcttgacgc ctgag 45
<210> 43
<211> 40
45 <212> DNA
<213> Artificial
<220>
<223> Primer sequence

<400> 43
tagaccgtgt gcctgacaga gtaccggag caagcgctga 40
<210> 44
<211> 39
5 <212> DNA
<213> Artificial
<220>
<223> Primer sequence
<400> 44
10 cgcttgctcc cggtaactctg tcagggcacac ggtctaact 39
<210> 45
<211> 40
<212> DNA
<213> Artificial
15 <220>
<223> Primer sequence
<400> 45
cagtccatga gcttgacata ggactggcaa cagccgttgt 40
<210> 46
20 <211> 39
<212> DNA
<213> Artificial
<220>
<223> Primer sequence
25 <400> 46
cggctgttgc cagtcctatg tcaagctcat ggactgacg 39
<210> 47
<211> 40
<212> DNA
30 <213> Artificial
<220>
<223> Primer sequence
<400> 47
aaggccatgg gcctgacata agacaggcag taatcgttgc 40
35 <210> 48
<211> 39
<212> DNA
<213> Artificial
<220>
40 <223> Primer sequence
<400> 48
cgattactgc ctgtcttatg tcaggccat ggccttact 39
<210> 49
<211> 43
45 <212> DNA
<213> Artificial
<220>
<223> Primer sequence

<400> 49
gagcccggtgc gcctggcaaa gcaccggtaa aagacggttg acg 43
<210> 50
<211> 50
5 <212> DNA
<213> Artificial
<220>
<223> Primer sequence
<400> 50
10 cgattactgc ctgtcttatg tcaggccat ggccttactc ctgagcaagt 50
<210> 51
<211> 40
<212> DNA
<213> Artificial
15 <220>
<223> Primer sequence
<400> 51
gagcccatga gcctggcaaa gaaccggaag aagccgttgg 40
<210> 52
20 <211> 51
<212> DNA
<213> Artificial
<220>
<223> Primer sequence
25 <400> 52
cggcttcttc cggttcttg ccaggctcat gggctcacgc cagagcaggt g 51
<210> 53
<211> 40
<212> DNA
30 <213> Artificial
<220>
<223> Primer sequence
<400> 53
gaggccgtgc gcttgacaga ggacgggaag taacctctgc 40
35 <210> 54
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 54
tccccacttt ctgtgaa 18
<210> 55
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 55
taaccactca ggacaggg 18
<210> 56
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 56
10 cactttcttg tgaatcctt 19
<210> 57
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 57
tcacacagca agtcagca 18
<210> 58
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 58
tagcggagca ggctcgga 18
<210> 59
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 59
tgggctagcg gagcaggct 19
35 <210> 60
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 60
tacccagacg agaaagct 18
<210> 61
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 61
tcagactgcc aagcttga 18
<210> 62
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 62
10 acccagacga gaaagctga 19
<210> 63
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 63
tcttgtggct cgggagta 18
<210> 64
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 64
tattgtcagc agagctga 18
<210> 65
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 65
agagggcatc ttgtggctc 19
35 <210> 66
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 66
ttgagatttt cagatgtc 18
<210> 67
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 67
tatacagtca tatcaagc 18
<210> 68
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 68
10 atcaagctct cttggcggt 19
<210> 69
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 69
ttcagataga ttatatct 18
<210> 70
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 70
tgccagatac ataggtgg 18
<210> 71
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 71
gcttcagata gattatatac 19
35 <210> 72
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 72
ttatactgtc tatatgat 18
<210> 73
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 73
tcagctttc tggccaga 18
<210> 74
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 74
10 acggatgtct cagctttc 19
<210> 75
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 75
tggccagaag agctgaga 18
<210> 76
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 76
ttaccgggga gagtttct 18
<210> 77
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 77
ccggggagag tttcttgta 19
35 <210> 78
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 78
tttgcagaga gatgagtc 18
<210> 79
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 79
ttagcagaag ataagatt 18
<210> 80
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 80
10 gaaatcttat cttctgcta 19
<210> 81
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 81
tataagacta aactaccc 18
<210> 82
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 82
tcgtctgccca ccacagat 18
<210> 83
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 83
aatgcatgac attcatctg 19
35 <210> 84
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 84
taaaaacagtt tgcattca 18
<210> 85
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 85
tataaagtcc tagaatgt 18
<210> 86
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 86
10 aacagtttgc attcatgga 19
<210> 87
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 87
tggccatctc tgacctgt 18
<210> 88
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 88
tagtgagccc agaagggg 18
<210> 89
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 89
ccagaagggg acagtaaga 19
35 <210> 90
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 90
taggtacctg gctgtcgt 18
<210> 91
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 91
tgaccgtcct ggcttta 18
<210> 92
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 92
10 ctgacaatcg ataggtacc 19
<210> 93
<211> 18
<212> DNA
<213> Artificial
15 <220>
<223> Targeting sequence
<400> 93
tgtcatggtc atctgcta 18
<210> 94
20 <211> 18
<212> DNA
<213> Artificial
<220>
<223> Targeting sequence
25 <400> 94
tcgacacccga agcagagt 18
<210> 95
<211> 19
<212> DNA
30 <213> Artificial
<220>
<223> Targeting sequence
<400> 95
acaccgaagc agagtttt 19
35 <210> 96
<211> 18
<212> DNA
<213> Artificial
<220>
40 <223> Targeting sequence
<400> 96
tgcccccgcg aggccaca 18
<210> 97
<211> 18
45 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence

<400> 97
tctggaaagg tt gaacaccc 18
<210> 98
<211> 19
5 <212> DNA
<213> Artificial
<220>
<223> Targeting sequence
<400> 98
10 ggaagttgaa cacccttgc 19
<210> 99
<211> 90
<212> DNA
<213> Artificial
15 <220>
<223> Target sequence
<400> 99
ctactgtcat tcagggcaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
tggcagtctg actacagagg ccactggctt 90
20 <210> 100
<211> 90
<212> DNA
<213> Artificial
<220>
25 <223> Target sequence
<400> 100
ctactgtcat tcagcccaat accctaacga gaaagctgag ggtataacag gtttcaagct 60
tggcagtctg actacagagg ccactggctt 90
<210> 101
30 <211> 90
<212> DNA
<213> Artificial
<220>
<223> Target sequence
35 <400> 101
ctactgtcat tcagcccaat acccagacga gaaaagttag ggtataacag gtttcaagct 60
tggcagtctg actacagagg ccactggctt 90
<210> 102
<211> 90
40 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 102
45 ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
tggcagtctg actacagagg ccactggctt 90
<210> 103
<211> 90

<212> DNA
<213> Artificial
<220>
<223> Target sequence
5 <400> 103
ctactgtcat tcagccaat acccagacga gaaagctgag ggtataacag gttttagct 60
tggcagtctg actacagagg ccactggctt 90
<210> 104
<211> 90
10 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 104
15 ctactgtcat tcagccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
tggctctcg actacagagg ccactggctt 90
<210> 105
<211> 90
<212> DNA
20 <213> Artificial
<220>
<223> Target sequence
<400> 105
ctactgtcat tcagccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
25 tggcagtctg actatgtgagg ccactggctt 90
<210> 106
<211> 90
<212> DNA
<213> Artificial
30 <220>
<223> Target sequence
<400> 106
cactttatat ttccctgctt aaacagtccc ccgagggtgg gtgcggaaaa ggctctacac 60
ttgttatcat tccctctcca ccacaggcat 90
35 <210> 107
<211> 90
<212> DNA
<213> Artificial
<220>
40 <223> Target sequence
<400> 107
tttgtatttg ggtttttta aacacctccac tctacagtta agaattctaa ggcacagagc 60
ttcaataatt tggtcagagc caagtagcag 90
<210> 108
45 <211> 90
<212> DNA
<213> Artificial
<220>

<223> Target sequence
<400> 108
ggaggttaaa cccaggcaga tgactgcagt tcttaatcaa tgcccattga attgcacata 60
tgggatgaac tagaacattt tctcgatgtat 90
5 <210> 109
<211> 90
<212> DNA
<213> Artificial
<220>
10 <223> Target sequence
<400> 109
ctcgatgatt cgctgtcatt gttatgatta tgttactgag ctctactgta gcacagacat 60
atgtccctat atggggcggg ggtgggggtg 90
<210> 110
15 <211> 90
<212> DNA
<213> Artificial
<220>
<223> Target sequence
20 <400> 110
ggtgtcttga tcgctggct atttctatac ttttctggct tttcggaaagc agtcatttct 60
ttctattctc caagcaccag caatttagctt 90
<210> 111
<211> 90
25 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 111
30 gcttctagtt tgctgaaact aatctgctat agacagagac tccgacgaac caattttatt 60
aggatttgat caaataaaact ctctctgaca 90
<210> 112
<211> 90
<212> DNA
35 <213> Artificial
<220>
<223> Target sequence
<400> 112
gaaagagtaa ctaagagttt gatgttact gagtgcatag tatgcactag atgctggccg 60
40 tggatgcctc atagaatcct cccaacaact 90
<210> 113
<211> 90
<212> DNA
<213> Artificial
45 <220>
<223> Target sequence
<400> 113
gctagatgct ggccgtggat gcctcataga atcctccaa caaccgatga aatgactact 60

gtcattcagc ccaataccca gacgagaaag 90
<210> 114
<211> 90
<212> DNA
5 <213> Artificial
<220>
<223> Target sequence
<400> 114
acaggttca agcttggcag tctgactaca gaggccactg gctttacccc tgggttagtc 60
10 tgcctctgta ggattggggg cacgtaattt 90
<210> 115
<211> 90
<212> DNA
<213> Artificial
15 <220>
<223> Target sequence
<400> 115
ttagtctgcc tctgtaggat tgggggcacg taatttgct gtttaaggc tcatttgct 60
tcttagagat cacaagccaa agcttttat 90
20 <210> 116
<211> 90
<212> DNA
<213> Artificial
<220>
25 <223> Target sequence
<400> 116
ggaagccag agggcatctt gtggctcggg agtagctctc tgctaccttc tcagctctgc 60
tgacaatact tgagatttc agatgtcacc 90
<210> 117
30 <211> 90
<212> DNA
<213> Artificial
<220>
<223> Target sequence
35 <400> 117
tcagctctgc tgacaatact tgagatttc agatgtcacc aaccagcaag agagcttgat 60
atgactgtat atagtatagt cataaagaac 90
<210> 118
<211> 90
40 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 118
45 cataaagaac ctgaacttga ccatatactt atgtcatgtg gaaatcttct catagcttca 60
gatagattat atctggagtg aagaatcctg 90
<210> 119
<211> 90

<212> DNA
<213> Artificial
<220>
<223> Target sequence
5 <400> 119
gtggaaaatt tctcatagct tcagatagat tataatcgga gtgagcaatc ctgccaccta 60
tgtatctggc atagtgttag tcctcataaa 90
<210> 120
<211> 90
10 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 120
15 ggtttaagg gcaacaaaaat agtgaacaga gtgaaaatcc ccacctagat cctgggtcca 60
gaaaaagatg ggaaacctgt ttagctcacc 90
<210> 121
<211> 30
<212> DNA
20 <213> Artificial
<220>
<223> Target sequence
<400> 121
ggccactagg gacaaaattg gtgacagaaa 30
25 <210> 122
<211> 50
<212> DNA
<213> Artificial
<220>
30 <223> Target sequence
<400> 122
ccccacagtgg ggccactagg gacaaaattg gtgacagaaa agccccatcc 50
<210> 123
<211> 70
35 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 123
40 tcccctccac cccacagtgg ggccactagg gacaaaattg gtgacagaaa agccccatcc 60
ttaggcctcc 70
<210> 124
<211> 90
<212> DNA
45 <213> Artificial
<220>
<223> Target sequence
<400> 124

cttttatctg tccccctccac cccacagtgg ggccactagg gacaaaattg gtgacagaaa 60
agccccatcc ttaggcctcc tccttcctag 90
<210> 125
<211> 110
5 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 125

gttctggta cttttatctg tccccctccac cccacagtgg ggccactagg gacaaaattg 60
gtgacagaaa agccccatcc ttaggcctcc tccttcctag tctcctgata 110
<210> 126
<211> 30
<212> DNA
15 <213> Artificial
<220>
<223> Target sequence
<400> 126
tttctgtcac caatggtgtc cctagtgccc 30
20 <210> 127
<211> 50
<212> DNA
<213> Artificial
<220>
25 <223> Target sequence
<400> 127
ggatggggct tttctgtcac caatggtgtc cctagtgccc ccactgtggg 50
<210> 128
<211> 70
30 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 128

ggaggcctaa ggatggggct tttctgtcac caatggtgtc cctagtgccc ccactgtggg 60
gtggagggga 70
<210> 129
<211> 90
<212> DNA
35 <213> Artificial
<220>
<223> Target sequence
<400> 129
ctaggaagga ggaggcctaa ggatggggct tttctgtcac caatggtgtc cctagtgccc 60
40 ccactgtggg gtggagggga cagataaaag 90
<210> 130
<211> 110
<212> DNA

<213> Artificial
<220>
<223> Target sequence
<400> 130

5 tatcaggaga ctaggaagga ggaggcctaa ggatggggct tttctgtcac caatggtgac 60
 cctagtggcc ccactgtggg gtggagggga cagataaaag tacccagaac 110
<210> 131
<211> 70
<212> DNA

10 <213> Artificial
<220>
<223> Target sequence
<400> 131
 ttctagtaac cactcaggac aggggggttc agcccaaaaa ttcacaagaa agtggggacc 60
15 catggaaat 70
<210> 132
<211> 70
<212> DNA
<213> Artificial

20 <220>
<223> Target sequence
<400> 132
 cagcaagtca gcagcacagc gtgtgtgact ccgagggtgc tccgctagcc cacattgcc 60
 tctgggggtg 70

25 <210> 133
<211> 70
<212> DNA
<213> Artificial
<220>

30 <223> Target sequence
<400> 133
 gtcagactgc caagcttcaa acctgtctta ccctctactt tctcgctgg gtattgggct 60
 gaatgacagt 70
<210> 134

35 <211> 70
<212> DNA
<213> Artificial
<220>
<223> Target sequence

40 <400> 134
 cagagctgag aagacagcag agagctactc ccgaaggcaca agatgccctc tgggcttccg 60
 tgaccttggc 70
<210> 135
<211> 70

45 <212> DNA
<213> Artificial
<220>
<223> Target sequence

<400> 135
ctgacaatac ttgagatttt cagatgtcac caacgaccaa gagagcttga tatgactgta 60
tatagtata 70
<210> 136
5 <211> 70
<212> DNA
<213> Artificial
<220>
<223> Target sequence
10 <400> 136
cagatacata ggtggcagga ttcttcactc cagacttaat ctatctgaag ctatgagaaa 60
ttttccacat 70
<210> 137
<211> 70
15 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 137
20 tataatgattt atttgcacag ctcatctggc cagataagct gagacatccg ttcccctaca 60
agaaaactctc 70
<210> 138
<211> 70
<212> DNA
25 <213> Artificial
<220>
<223> Target sequence
<400> 138
atctggccag aagagctgag acatccgttc cccttgaaga aactctcccc ggtaagtaac 60
30 ctctcagctg 70
<210> 139
<211> 70
<212> DNA
<213> Artificial
35 <220>
<223> Target sequence
<400> 139
aggcatctca ctggagaggg ttttagttctc cttaagagaa gataagattt caagaggaa 60
gctaagactc 70
40 <210> 140
<211> 70
<212> DNA
<213> Artificial
<220>
45 <223> Target sequence
<400> 140
ataatataat aaaaaatgtt tcgtctgcc a ccactaatga atgtcatgca ttctggtag 60
tttagtctta 70

<210> 141
<211> 70
<212> DNA
<213> Artificial
5 <220>
<223> Target sequence
<400> 141
tttataaaagt cctagaatgt atttagttgc cctcggtgaa tgcaaactgt tttatacatac 60
aatagggttt 70
10 <210> 142
<211> 70
<212> DNA
<213> Artificial
<220>
15 <223> Target sequence
<400> 142
gctcaacctg gccatctctg acctgtttt ccttcccact gtcccccact gggctcacta 60
tgctgccgcc 70
<210> 143
20 <211> 70
<212> DNA
<213> Artificial
<220>
<223> Target sequence
25 <400> 143
ttttaaagca aacacagcat ggacgacagc caggctccta tcgattgtca ggaggatgtat 60
gaagaagatt 70
<210> 144
<211> 70
30 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 144
35 <210> 145
gcttgtcatg gtcatctgct actcggaat cctaattact ctgcttcgggt gtcgaaatga 60
gaagaagagg 70
<211> 70
<212> DNA
40 <213> Artificial
<220>
<223> Target sequence
<400> 145
atactgcccc cgcgaggcca cattggcaaa ccagcttggg tggtaactt ccagacttgg 60
45 ccatggagaa 70
<210> 146
<211> 90
<212> DNA

<213> Artificial
<220>
<223> Target sequence
<400> 146

5 ctgaagaatt tcccatgggt ccccacttcc ttgtaatcc ttggagtgaa cccccctgtc 60
ctgagtggtt actagaacac acctctggac 90
<210> 147
<211> 90
<212> DNA

10 <213> Artificial
<220>
<223> Target sequence
<400> 147

tggaaagtatc ttgccgaggt cacacagcaa gtcagcagca cagccagtgt gactccgagc 60
15 ctgctccgct agcccacatt gccctctggg 90
<210> 148
<211> 90
<212> DNA
<213> Artificial

20 <220>
<223> Target sequence
<400> 148

ctactgtcat tcagccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
tggcagtctg actacagagg ccactggctt 90

25 <210> 149
<211> 90
<212> DNA
<213> Artificial
<220>

30 <223> Target sequence
<400> 149

ggaagccag agggcatctt gtggctcggt agtagctctc tgctaccttc tcagctctgc 60
tgacaatact tgagatttc agatgtcacc 90
<210> 150

35 <211> 90
<212> DNA
<213> Artificial
<220>

<223> Target sequence

40 <400> 150

tcagctctgc tgacaatact tgagatttc agatgtcacc aacgccaag agagcttgat 60
atgactgtat atagtatagt cataaagaac 90
<210> 151
<211> 90

45 <212> DNA
<213> Artificial
<220>

<223> Target sequence

<400> 151
 gtggaaaatt tctcatagct tcagatagat tatatatcgga gtgagcaatc ctgccaccta 60
 tgtatctggc atagtgttag tcctcataaa 90
 <210> 152
 5 <211> 90
 <212> DNA
 <213> Artificial
 <220>
 <223> Target sequence
 10 <400> 152
 gaaacacgcat ttcctacttt tatactgtct atatgattga tttggtcagc tcatctggcc 60
 agaagagctg agacatccgt tcccctacaa 90
 <210> 153
 <211> 90
 15 <212> DNA
 <213> Artificial
 <220>
 <223> Target sequence
 <400> 153
 20 ttgatttgca cagctcatct ggccagaaga gctgagacat ccgtatccct acaagaaact 60
 ctccccggta agtaacctct cagctgctt 90
 <210> 154
 <211> 90
 <212> DNA
 25 <213> Artificial
 <220>
 <223> Target sequence
 <400> 154
 ggagaggggtt tagttctcct tagcagaaga taagattca agatgagagc taagactcat 60
 30 ctctctgcaa atctttcttt tgagaggtaa 90
 <210> 155
 <211> 90
 <212> DNA
 <213> Artificial
 35 <220>
 <223> Target sequence
 <400> 155
 taatataata aaaaatgttt cgtctgccac cacagatgaa tgcgtcgat tctggtagt 60
 ttagtcttat aaccagctgt cttgcctagt 90
 40 <210> 156
 <211> 90
 <212> DNA
 <213> Artificial
 <220>
 45 <223> Target sequence
 <400> 156
 taaaaaacct attgatgtat aaaacagttt gcattcatgg agggtgacta aatacattct 60
 aggactttat aaaagatcac ttttattta 90

<210> 157
<211> 90
<212> DNA
<213> Artificial
5 <220>
<223> Target sequence
<400> 157
gacatctacc tgctcaacct ggccatctct gacctgttt tcctatttac tgtccccc 60
tgggctcact atgctgccgc ccagtgaaac 90
10 <210> 158
<211> 90
<212> DNA
<213> Artificial
<220>
15 <223> Target sequence
<400> 158
tcatccctcct gacaatcgat aggtacctgg ctgtcgcca tgctacgttt gctttaaaag 60
ccaggacggt cacccttggg gtggtgacaa 90
<210> 159
20 <211> 90
<212> DNA
<213> Artificial
<220>
<223> Target sequence
25 <400> 159
ggctggtcct gccgctgctt gtcatggta tctgctactc gggagaccta aaaactctgc 60
ttcggtgtcg aaatgagaag aagaggcaca 90
<210> 160
<211> 90
30 <212> DNA
<213> Artificial
<220>
<223> Target sequence
<400> 160
35 <210> 161
ggcaaggcctt gggtcataact gccccgcga ggccacattt gcaagttagc aagggtgttc 60
aacttccaga cttggccatg gagaagacat 90
<211> 59
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 161
acactctttc cttacacgac gctcttccga tctcggtttt ttgcagtgtg cgttactcc 59
45 <210> 162
<211> 59
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 162
acactcttcc cttacacgac gctcttccga tctacatcgt ttgcagtgtg cgttactcc 59
5 <210> 163
<211> 59
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 163
acactcttcc cttacacgac gctcttccga tctgcctaatttgcagtgtg cgttactcc 59
<210> 164
<211> 59
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 164
20 acactcttcc cttacacgac gctcttccga tcttggtcat ttgcagtgtg cgttactcc 59
<210> 165
<211> 54
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 165
ctcggcatttc ctgctgaacc gctcttccga tctccaagca actaagtcac agca 54
<210> 166
30 <211> 58
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 166
acactcttcc cttacacgac gctcttccga tctcgtgata tgaggaaatg gaagcttg 58
<210> 167
<211> 58
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 167
acactcttcc cttacacgac gctcttccga tctacatcga tgaggaaatg gaagcttg 58
45 <210> 168
<211> 58
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 168
acactcttc cctacacgac gctcttccga tctgcctaaa tgaggaaatg gaagcttg 58
5 <210> 169
<211> 58
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 169
acactcttc cctacacgac gctcttccga tcttggtaaa tgaggaaatg gaagcttg 58
<210> 170
<211> 51
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 170
20 ctcggcattc ctgctgaacc gctcttccga tctcattagg gtattggagg a 51
<210> 171
<211> 58
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 171
acactcttc cctacacgac gctcttccga tctcggtata atcctcccaa caactcat 58
<210> 172
30 <211> 58
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 172
acactcttc cctacacgac gctcttccga tctacatcga atcctcccaa caactcat 58
<210> 173
<211> 58
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 173
acactcttc cctacacgac gctcttccga tctgcctaaa atcctcccaa caactcat 58
45 <210> 174
<211> 58
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 174
acactcttc cctacacgac gctcttccga tcttggtcaa atcctccaa caactcat 58
5 <210> 175
<211> 52
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 175
ctcggcattc ctgctgaacc gctcttccga tctcccaatc ctacagaggc ag 52
<210> 176
<211> 58
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 176
20 acactcttc cctacacgac gctcttccga tctcgtgata agccaaagct ttttattc 58
<210> 177
<211> 58
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 177
acactcttc cctacacgac gctcttccga tctacatcga agccaaagct ttttattc 58
<210> 178
30 <211> 58
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 178
acactcttc cctacacgac gctcttccga tctgcctaaa agccaaagct ttttattc 58
<210> 179
<211> 58
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 179
acactcttc cctacacgac gctcttccga tcttggtcaa agccaaagct ttttattc 58
45 <210> 180
<211> 53
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 180
acactcttcc cctacacgac gctcttccga tctaagccaa agcttttat tct 53
5 <210> 181
<211> 59
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 181
acactcttcc cctacacgac gctcttccga tctcgtgata tcttgtggct cgggagtag 59
<210> 182
<211> 59
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 182
20 acactcttcc cctacacgac gctcttccga tctacatcga tcttgtggct cgggagtag 59
<210> 183
<211> 53
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 183
ctcggcatttc ctgctgaacc gctcttccga tcttggcagg attttcaact cca 53
<210> 184
30 <211> 59
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 184
acactcttcc cctacacgac gctcttccga tctcgtgatc tattttgttg cccttcaaa 59
<210> 185
<211> 59
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 185
acactcttcc cctacacgac gctcttccga tctacatcgc tattttgttg cccttcaaa 59
45 <210> 186
<211> 55
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 186
ctcggcattc ctgctgaacc gctcttccga tctaacctga acttgaccat atact 55
5 <210> 187
<211> 59
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 187
acactcttgc cttacacgac gctcttccga tctcggtatc agctgagagg ttacttacc 59
<210> 188
<211> 59
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 188
20 acactcttgc cttacacgac gctcttccga tctacatcgc agctgagagg ttacttacc 59
<210> 189
<211> 53
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 189
ctcggcattc ctgctgaacc gctcttccga tctaattgatt taactccacc ctc 53
<210> 190
30 <211> 60
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 190
acactcttgc cttacacgac gctcttccga tctcggtata ctccaccctc cttcaaaaga 60
<210> 191
<211> 60
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 191
acactcttgc cttacacgac gctcttccga tctacatcga ctccaccctc cttcaaaaga 60
45 <210> 192
<211> 52
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 192
ctcggcattc ctgctgaacc gctcttccga tcttggtgtt tgccaaatgt ct 52
5 <210> 193
<211> 59
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 193
acactctttc cctacacgac gctcttccga tctcgatgc ggcacatatt cagaaggca 59
<210> 194
<211> 59
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 194
20 acactctttc cctacacgac gctcttccga tctacatcg ggcacatatt cagaaggca 59
<210> 195
<211> 56
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 195
ctcggcattc ctgctgaacc gctcttccga tcttagtggaa gactttaag ggagca 56
<210> 196
30 <211> 59
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 196
acactctttc cctacacgac gctcttccga tctcgatgc acaattaaga gttgtcata 59
<210> 197
<211> 59
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 197
acactctttc cctacacgac gctcttccga tctacatcg acaattaaga gttgtcata 59
45 <210> 198
<211> 53
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 198
ctcggcattc ctgctgaacc gctcttccga tctctcagct agagcagctg aac 53
5 <210> 199
<211> 51
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 199
ctcggcattc ctgctgaacc gctcttccga tctgacactt gataatccat c 51
<210> 200
<211> 60
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 200
20 acactcttc cctacacgac gctcttccga tctacatcgtaaatgttagac atctatgttag 60
<210> 201
<211> 60
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 201
acactcttc cctacacgac gctcttccga tctcgtgatt caatgttagac atctatgttag 60
<210> 202
30 <211> 59
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 202
acactcttc cctacacgac gctcttccga tctcgtgata ctgcaaaagg ctgaagagc 59
<210> 203
<211> 59
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 203
acactcttc cctacacgac gctcttccga tctacatcgtaaatgttagac atctatgttag 59
45 <210> 204
<211> 59
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 204
acactcttcc cttacacgac gctcttccga tctgcctaaa ctgcaaaagg ctgaagagc 59
5 <210> 205
<211> 59
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 205
acactcttcc cttacacgac gctcttccga tcttggtcaa ctgcaaaagg ctgaagagc 59
<210> 206
<211> 57
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 206
20 ctcggcattc ctgctgaacc gctcttccga tctgcctata aaatagagcc ctgtcaa 57
<210> 207
<211> 60
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 207
acactcttcc cttacacgac gctcttccga tctcggtatc tctatttat aggcttcttc 60
<210> 208
30 <211> 60
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 208
acactcttcc cttacacgac gctcttccga tctacatcgc tctatttat aggcttcttc 60
<210> 209
<211> 52
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 209
ctcggcattc ctgctgaacc gctcttccga tctagccacc acccaagtga tc 52
45 <210> 210
<211> 60
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 210
acactcttcc cctacacgac gctcttccga tctacatcgt tccagacatt aaagatagtc 60
5 <210> 211
<211> 60
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 211
acactcttcc cctacacgac gctcttccga tctcgtgatt tccagacatt aaagatagtc 60
<210> 212
<211> 54
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 212
20 ctcggcattc ctgctgaacc gctcttccga tctaattcatg atggtaaga taag 54
<210> 213
<211> 60
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 213
acactcttcc cctacacgac gctcttccga tctcgtgatc cggcagagac aaacattaaa 60
<210> 214
30 <211> 54
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <400> 214
acactcttcc cctacacgac gctcttccga tctccggcag agacaaacat taaa 54
<210> 215
<211> 53
<212> DNA
40 <213> Artificial
<220>
<223> PCR primer
<400> 215
ctcggcattc ctgctgaacc gctcttccga tcttagctagg aagccatggc aag 53
45 <210> 216
<211> 47
<212> DNA
<213> Artificial

<220>
<223> PCR primer
<400> 216
aatgataacgg cgaccaccga gatctacact cttccctac acgacgc 47
5 <210> 217
<211> 50
<212> DNA
<213> Artificial
<220>
10 <223> PCR primer
<400> 217
caagcagaag acggcatacg agatcggtct cggcattcct gctgaaccgc 50
<210> 218
<211> 58
15 <212> DNA
<213> Artificial
<220>
<223> PCR primer
<400> 218
20 acactcttc cctacacgac gctcttccga tctagtgcatagtatgtgct agatgctg 58
<210> 219
<211> 59
<212> DNA
<213> Artificial
25 <220>
<223> PCR primer
<400> 219
gtgactggag ttcagacgtg tgctttccg atcttgatct ctaagaaggc aaatgagac 59
<210> 220
30 <211> 64
<212> DNA
<213> Artificial
<220>
<223> PCR primer
35 <220>
<221> misc_feature
<222> (25)..(30)
<223> n is a, c, g, or t
<400> 220
40 caagcagaag acggcatacg agatnnnnnn gtgactggag ttcagacgtg tgctttccg 60
atct 64
<210> 221
<211> 58
<212> DNA
45 <213> Artificial
<220>
<223> PCR primer
<400> 221

aatgatacgg cgaccaccga gatctacact cttccctac acgacgctct tccgatct 58
 <210> 222
 <211> 34
 <212> PRT
 5 <213> Artificial
 <220>
 <223> Deduced TALE template amino acid sequence
 <220>
 <221> misc_feature
 10 <222> (12)..(13)
 <223> Xaa can be any naturally occurring amino acid
 <400> 222
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys
 1 5 10 15
 15 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 20 25 30
 His Gly
 <210> 223
 <211> 246
 20 <212> DNA
 <213> Artificial
 <220>
 <223> TALE dimer nucleotide sequence
 <400> 223
 25 cgattactgc ctgtctttag tcaggccat ggccttactc ctgagcaagt cgtagctatc 60
 gccagcaaca taggtggaa acaggccctg gaaaccgtac aacgtctcct cccagtactt 120
 tgtcaagcac acgggttjac accggaacaa gtggtggcga ttgcgtccc ccatggaggc 180
 aagcaggcac tggagaccgt ccaacggctt cttccggc 240
 acgcca 246
 30 <210> 224
 <211> 82
 <212> PRT
 <213> Artificial
 <220>
 35 <223> Deduced TALE dimer amino acid sequence
 <400> 224
 Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
 1 5 10 15
 Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
 40 20 25 30
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
 35 40 45
 Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
 50 55 60
 45 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
 65 70 75 80
 Thr Pro

(57) Формула изобретения

1. Способ изменения ДНК-мишени в стволовой клетке, экспрессирующий фермент Cas из системы CRISPR типа II, который формирует комплекс колокализации с направляющей РНК, комплементарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом, включающий
 5 введение в указанную стволовую клетку направляющую РНК, комплементарную ДНК-мишени, где РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени,
 10 введение в ту же стволовую клетку донорной последовательности нуклеиновой кислоты,
 где направляющая РНК и фермент Cas колокализуются с ДНК-мишенью, фермент Cas расщепляет ДНК-мишень и донорная последовательность нуклеиновой кислоты встраивается в ДНК-мишень, с получением измененной ДНК в указанной стволовой 15 клетке,
 где стволовая клетка не является клеткой зародышевой линии человека или эмбриональной клеткой человека.
2. Способ по п. 1, где фермент Cas из системы CRISPR типа II представляет собой Cas9.
- 20 3. Способ по п. 1, где длина направляющей РНК составляет от около 10 до около 500 нуклеотидов.
4. Способ по п. 1, где длина направляющей РНК составляет от около 20 до около 100 нуклеотидов.
- 25 5. Способ по п. 1, где направляющая РНК представляет собой слияние crRNA-tracrRNA.
6. Способ по п. 1, где ДНК представляет собой геномную ДНК, митохондриальную ДНК, вирусную ДНК или экзогенную ДНК.
7. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается в результате рекомбинации.
- 30 8. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается путем гомологичной рекомбинации.
9. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается негомологичным соединением концов.
10. Способ по п. 1, где фермент Cas и направляющая РНК вводятся в стволовую 35 клетку в виде чужеродной нуклеиновой кислоты, которая экспрессируется стволовой клеткой.
11. Способ по п. 1, дополнительно включающий введение множества донорных нуклеиновых кислот и множества направляющих РНК, чтобы осуществить множество изменений в ДНК в клетке.
- 40 12. Способ по п. 1, где после введения измененной ДНК в стволовую клетку, нуклеиновая кислота, кодирующая фермент Cas, удаляется из генома стволовой клетки.
13. Способ по п. 1, где направляющая РНК и донорные нуклеотидные последовательности экспрессируются как связанная последовательность нуклеиновой кислоты.
- 45 14. Стволовая клетка для РНК-направленного основанного на Cas редактирования генома, где стволовая клетка включает первую чужеродную нуклеиновую кислоту, кодирующую фермент Cas из системы CRISPR типа II, находящуюся под контролем индуцируемого промотора, где фермент Cas образует комплекс колокализации с РНК,

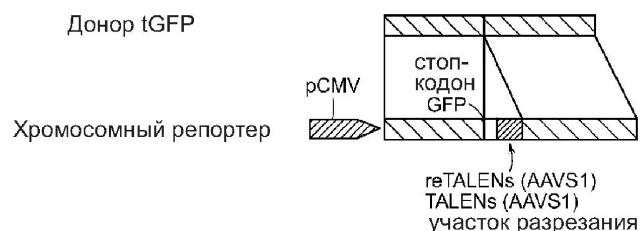
комплémentарной ДНК-мишени и который расщепляет ДНК-мишень сайт-специфическим образом,

где стволовая клетка не является клеткой зародышевой линии человека или эмбриональной клеткой человека.

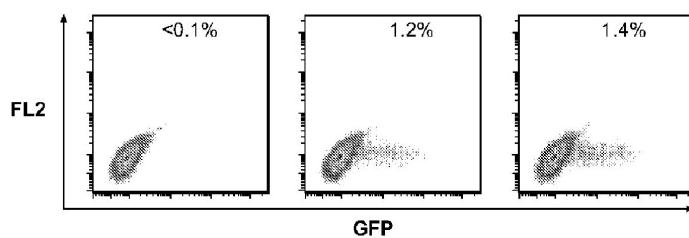
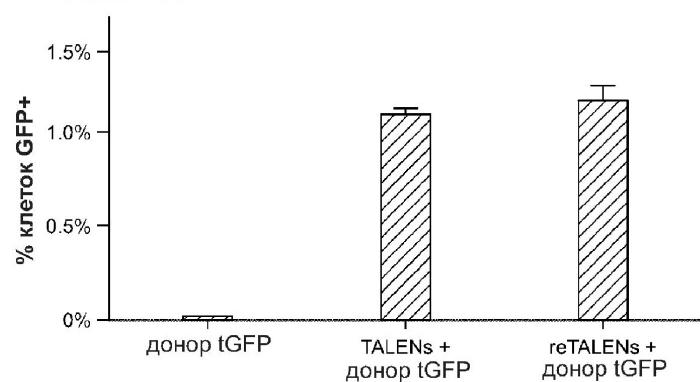
- 5 15. Стволовая клетка по п. 14, дополнительно включающая вторую чужеродную нуклеиновую кислоту, кодирующую направляющую РНК, комплементарную ДНК-мишени, где направляющая РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени.
- 10 16. Стволовая клетка по п. 14, отличающаяся тем, что первая чужеродная нуклеиновая кислота удаляется из геномной ДНК стволовой клетки с помощью транспозазы.
17. Стволовая клетка по п. 14, отличающаяся тем, что фермент Cas является Cas9.
18. Стволовая клетка по п. 14, отличающаяся тем, что направляющая РНК представляет собой слияние crRNA-tracrRNA.
19. Стволовая клетка по п. 14, отличающаяся тем, что ДНК-мишень представляет собой геномную ДНК, митохондриальную ДНК, вирусную ДНК или экзогенную ДНК.
- 20 20. Стволовая клетка по п. 14, отличающаяся тем, что дополнительно включает донорную последовательность нуклеиновой кислоты.
21. Способ по п. 1, где направляющая РНК вводится в стволовую клетку из среды, окружающей стволовую клетку.
22. Способ по п. 21, в котором направляющая РНК включает структуру 5' Cap.
23. Способ по п. 21, где направляющая РНК утратила фосфатные группы.
24. Способ по п. 1, где направляющая РНК непрерывно поступает в стволовую клетку из среды, окружающей стволовую клетку.
- 25 25. Способ по п. 1, где стволовая клетка генетически модифицирована путем встраивания в геном клетки удаляемых ферментом вектора или кассеты, содержащих нуклеиновую кислоту, кодирующую фермент Cas.
26. Способ по п. 25, где вектор представляет собой вектор PiggyBac.
27. Способ по п. 25, где нуклеиновая кислота, кодирующая фермент Cas, удаляется с помощью фермента.
- 30 28. Способ по п. 25, где удаляемые ферментом вектор или кассета, содержащие нуклеиновую кислоту, кодирующую фермент Cas, удаляются с помощью фермента.
29. Способ по п. 25, где удаляемые ферментом вектор или кассета, содержащие нуклеиновую кислоту, кодирующую фермент Cas, встраиваются с помощью фермента.
- 35 30. Способ по п. 25, где удаляемые ферментом вектор или кассета, содержащие нуклеиновую кислоту, кодирующую фермент Cas, удаляются с помощью транспозазы.
31. Способ по п. 25, где нуклеиновая кислота, кодирующая фермент Cas, является индуцибелной.
32. Способ по п. 25, где нуклеиновая кислота, кодирующая фермент Cas, находится под влиянием промотора, индуцируемого dox.
- 40 33. Способ по п. 25, где нуклеиновая кислота, кодирующая фермент Cas, индуцируется с помощью dox.
34. Стволовая клетка по п. 14, также содержащая направляющую РНК, комплементарную ДНК-мишени, где указанная направляющая РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени.
- 45 35. Стволовая клетка по п. 14, также содержащая направляющую РНК, комплементарную ДНК-мишени и донорную последовательность нуклеиновой кислоты, где направляющая РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени.

1/40

Фиг. 1А

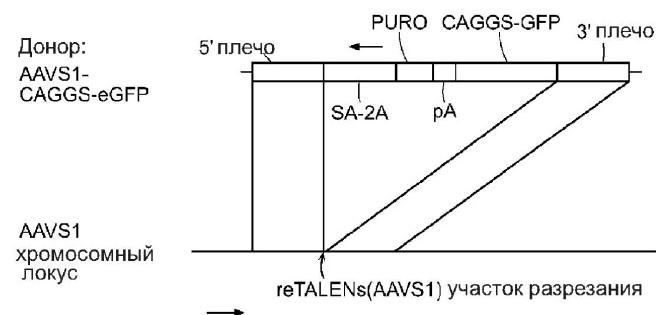


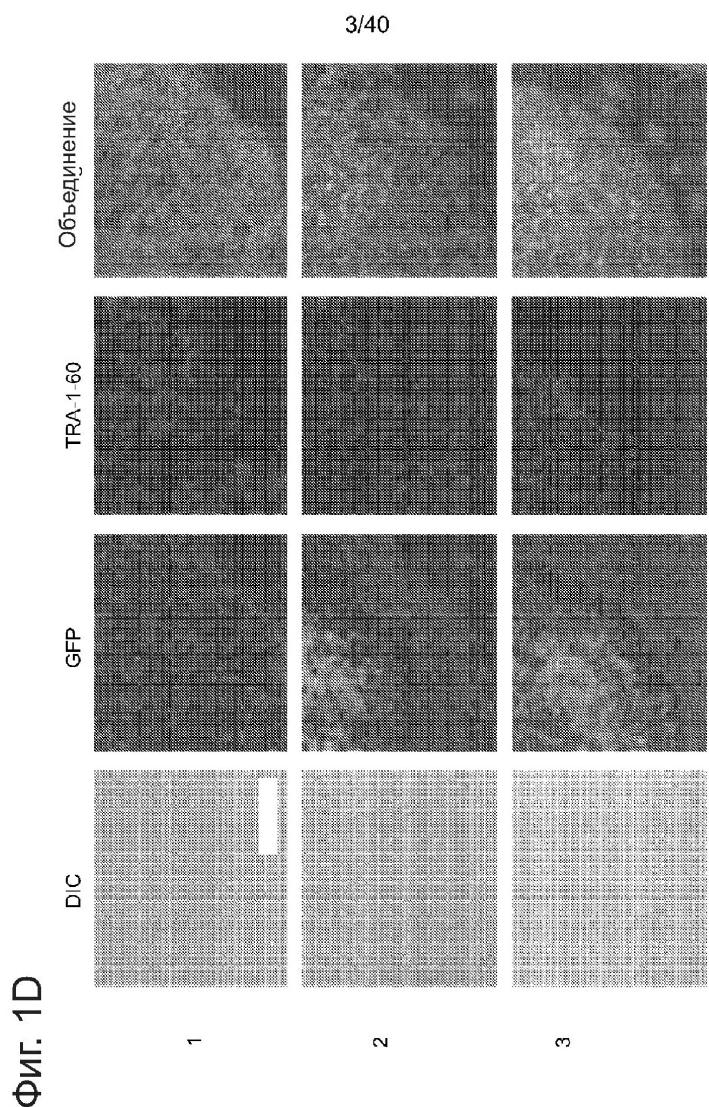
Фиг. 1В



2/40

Фиг. 1С

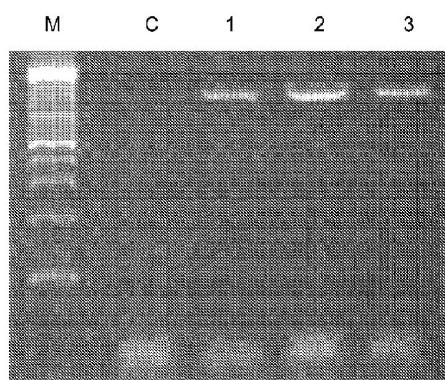




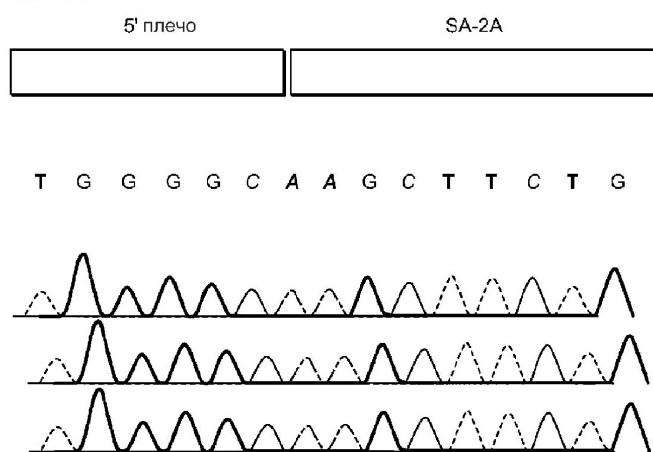
Фиг. 1D

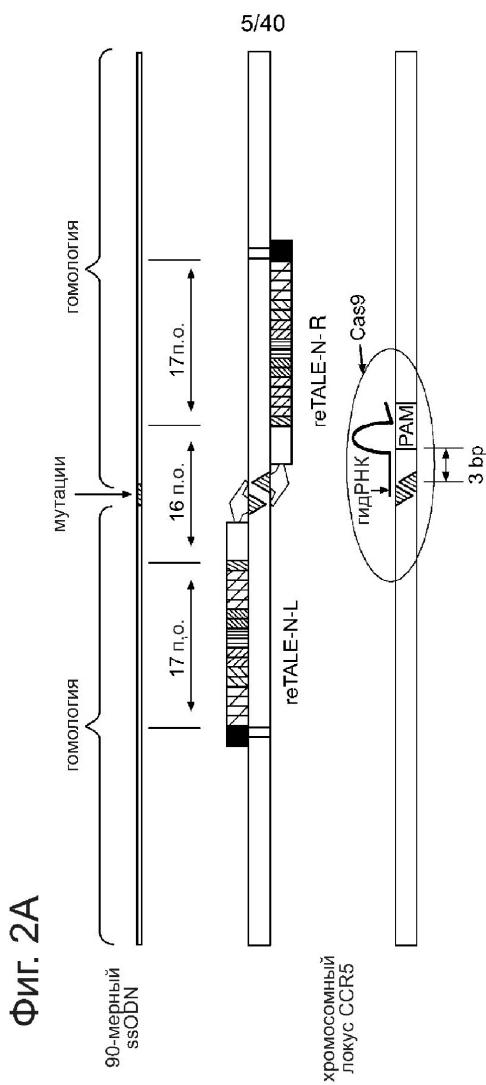
4/40

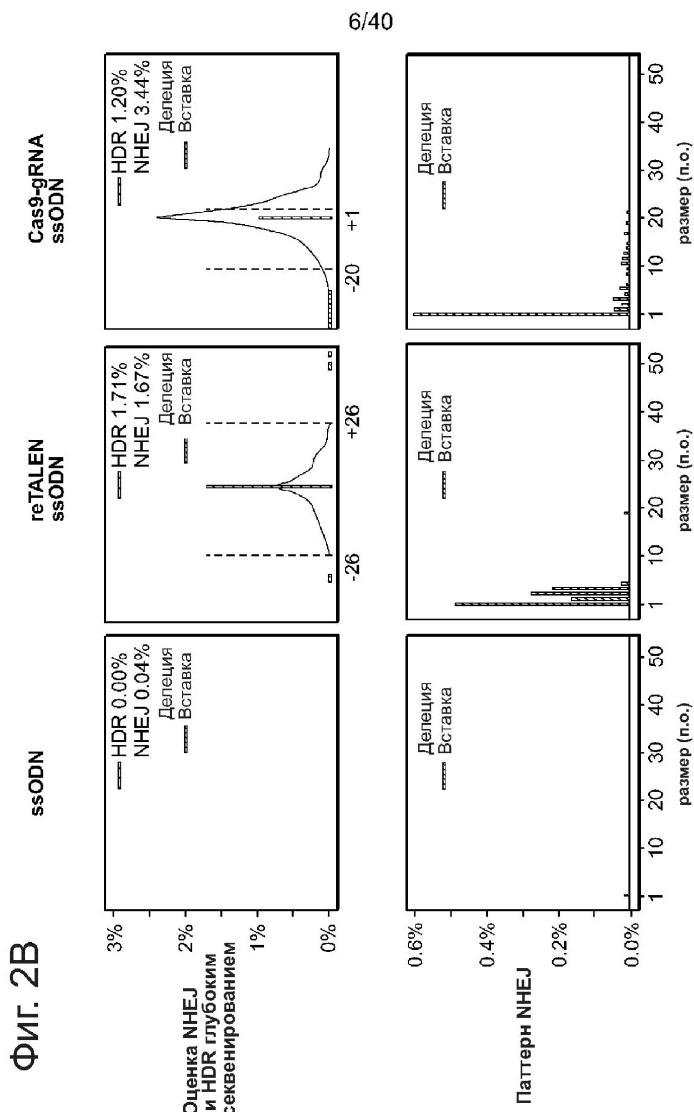
Фиг. 1Е



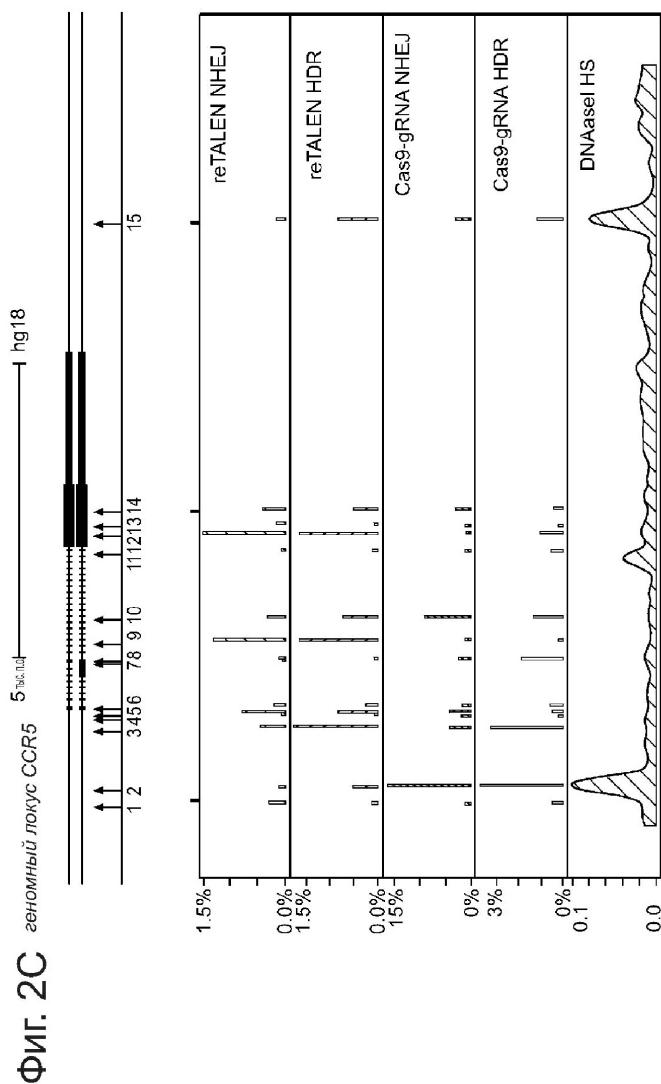
Фиг. 1F





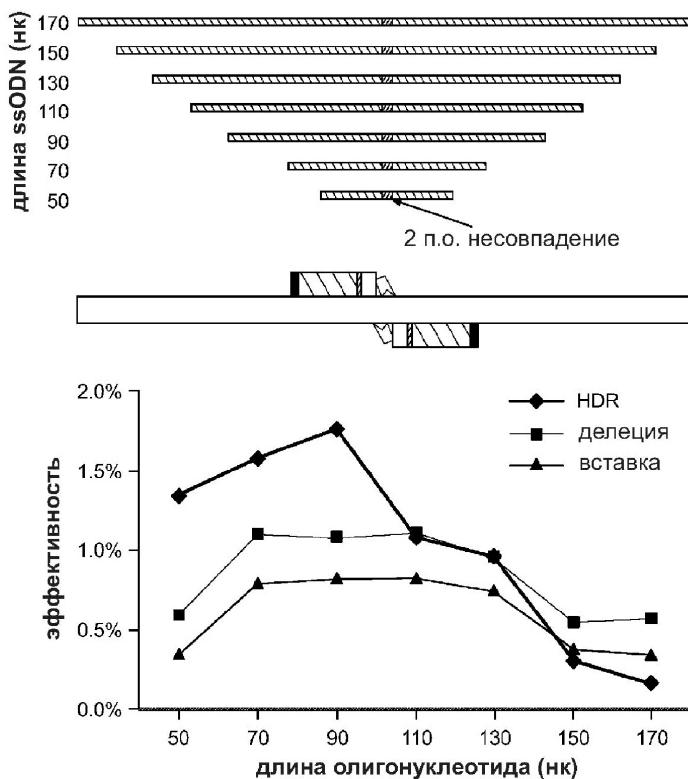


7/40



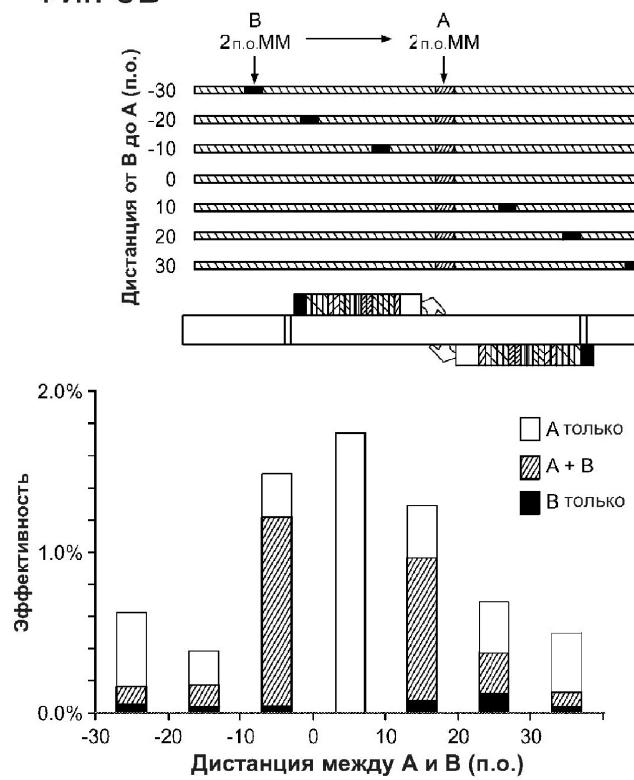
8/40

Фиг. 3А



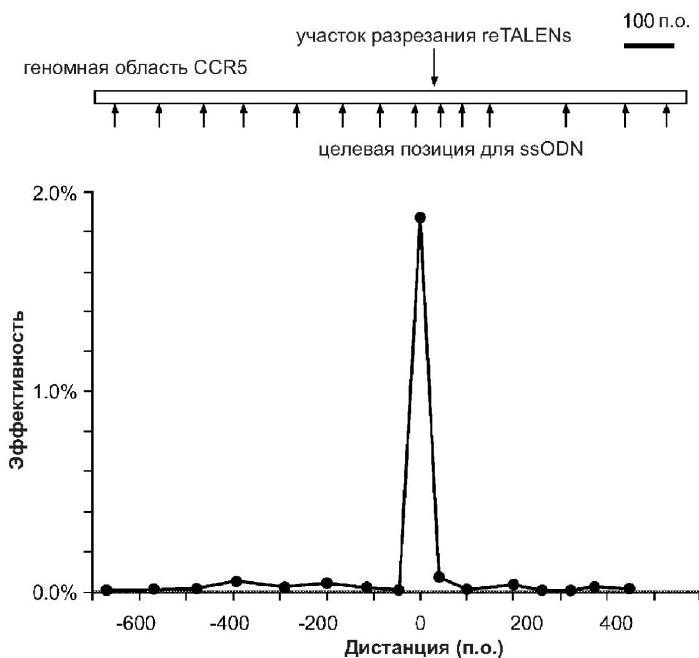
9/40

Фиг. 3В



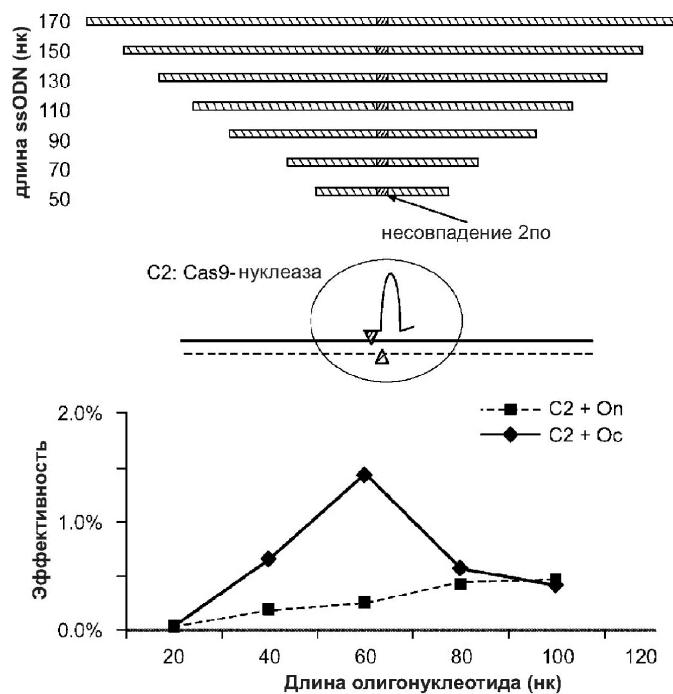
10/40

Фиг. 3С

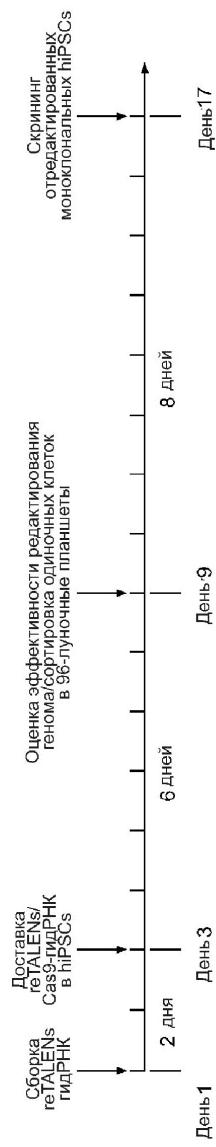


11/40

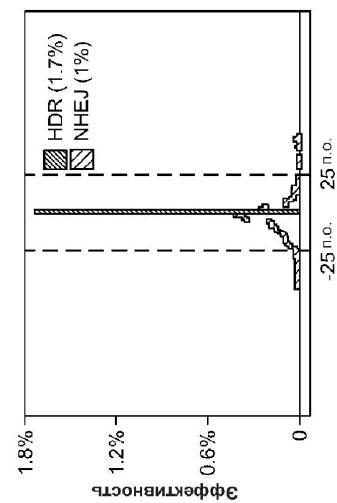
ФИГ. 3Д



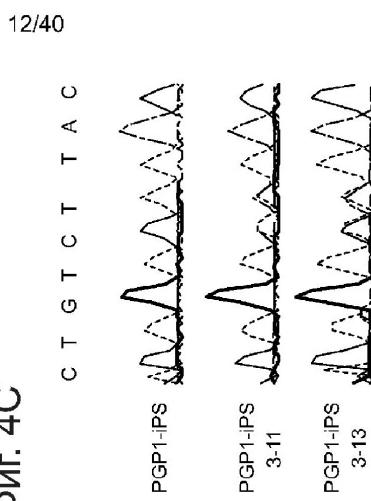
ФИГ. 4А

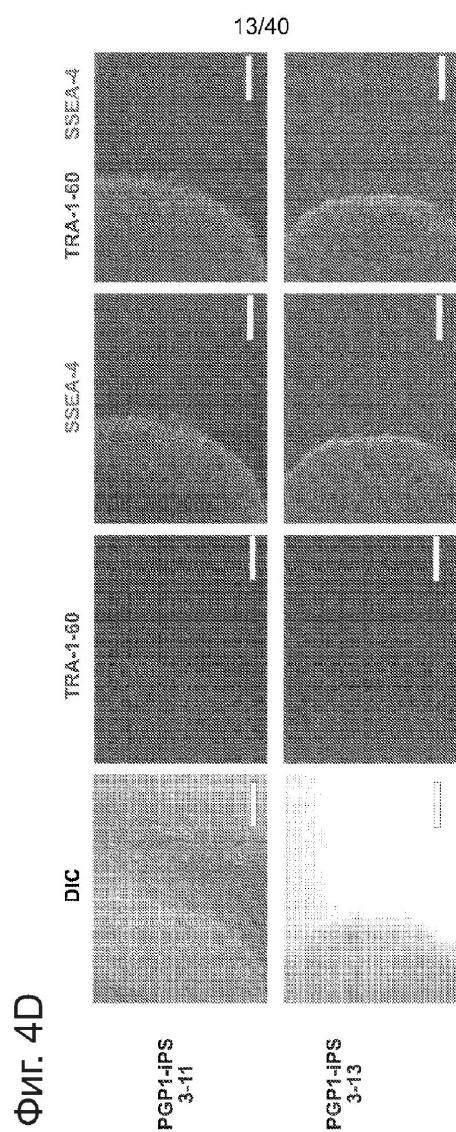


ФИГ. 4В

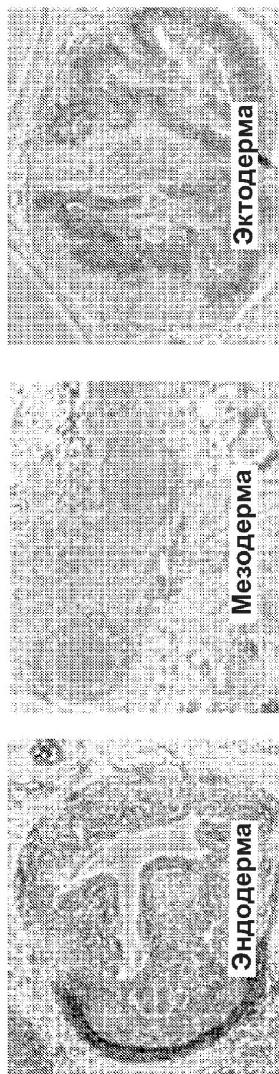


ФИГ. 4С

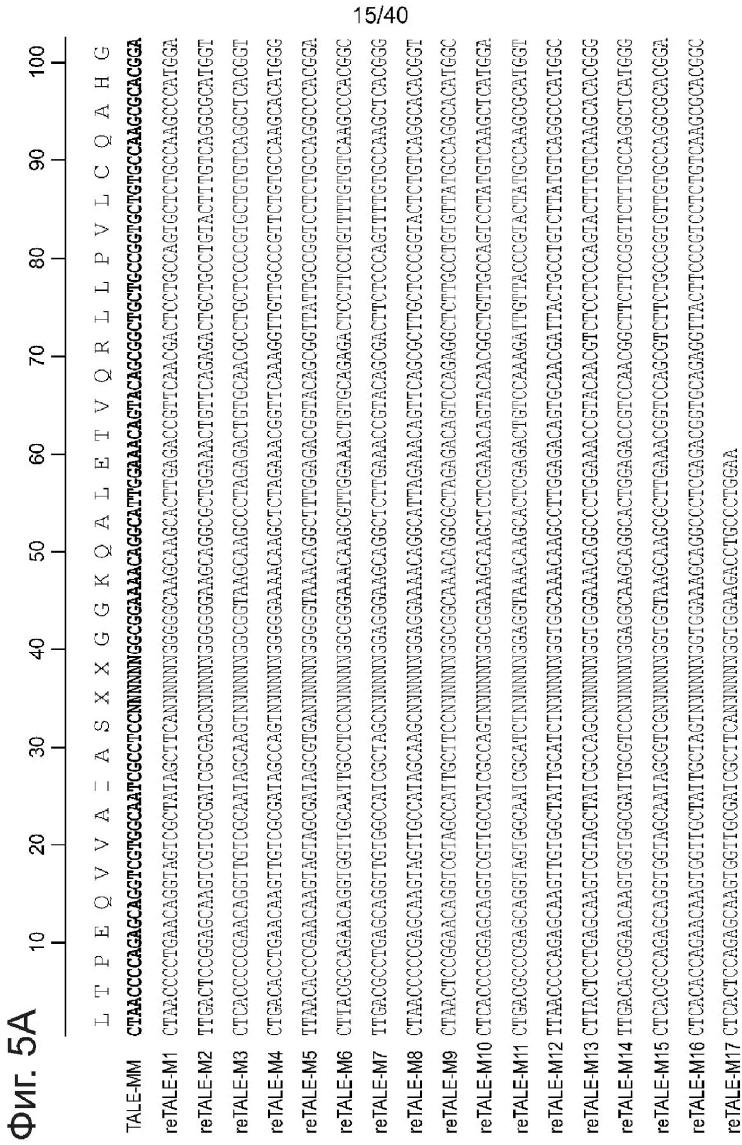




14/40

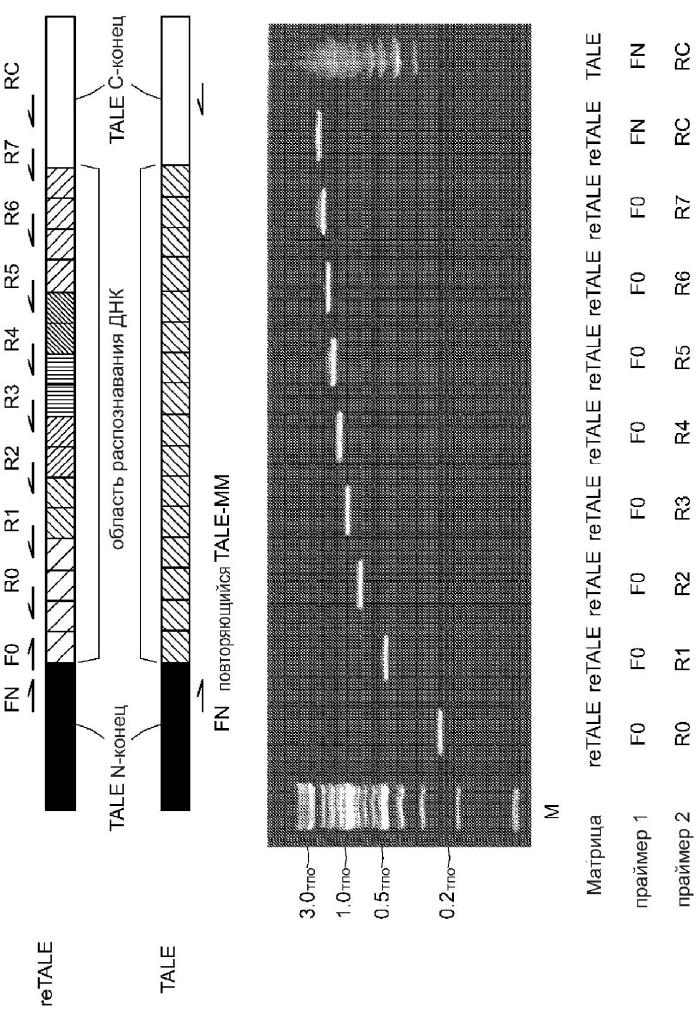


Фиг. 4Е

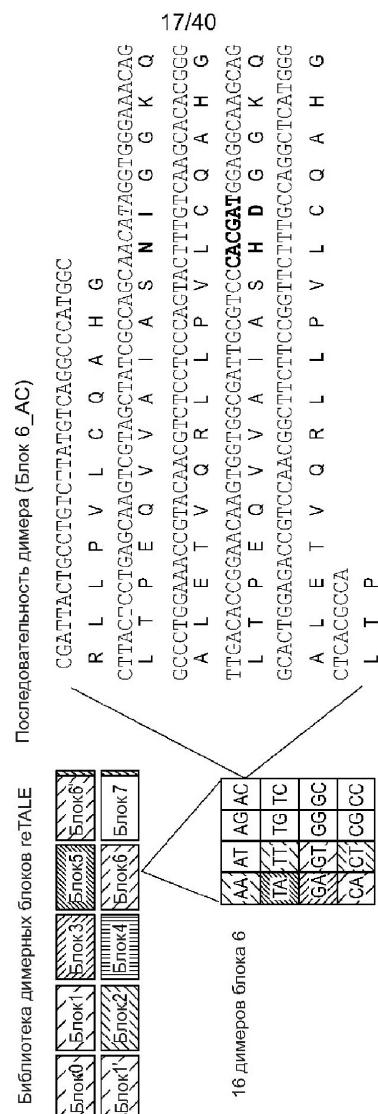


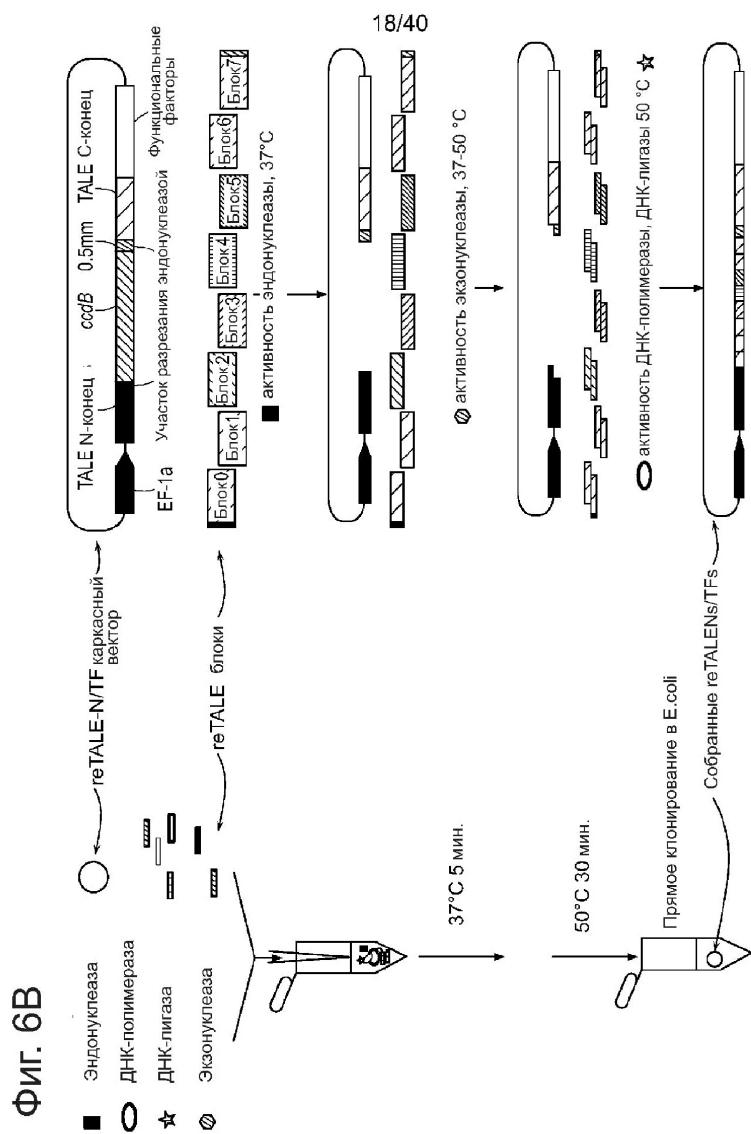
16/40

Фиг. 5В

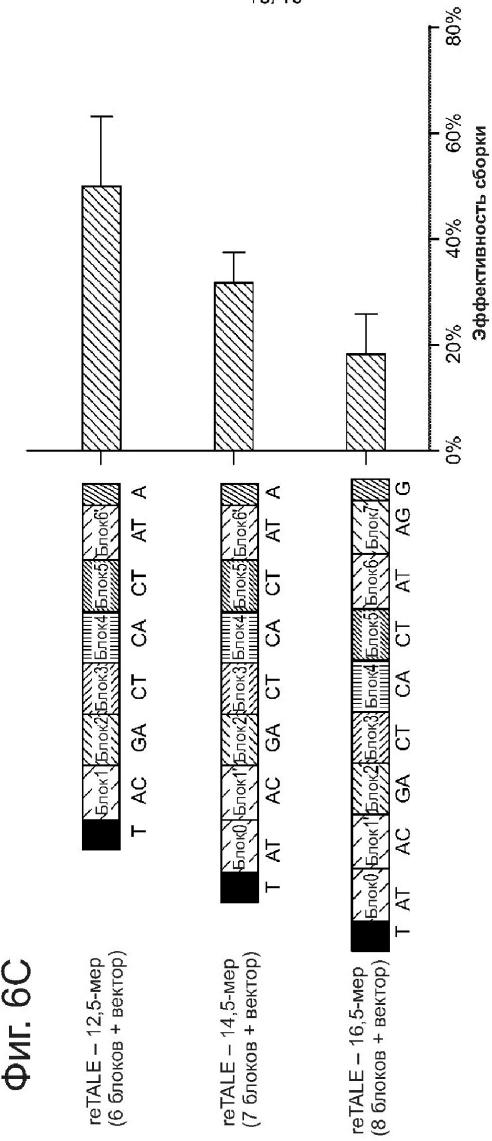


Фиг. 6А



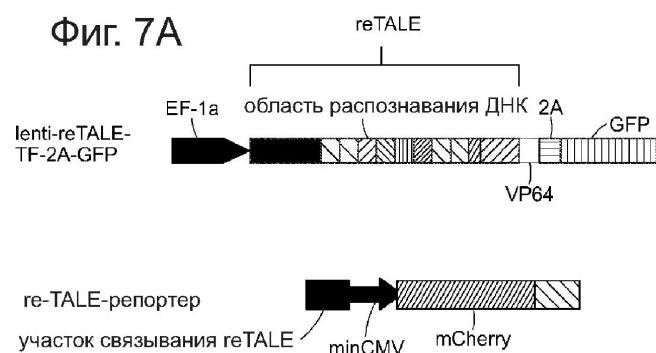


19/40

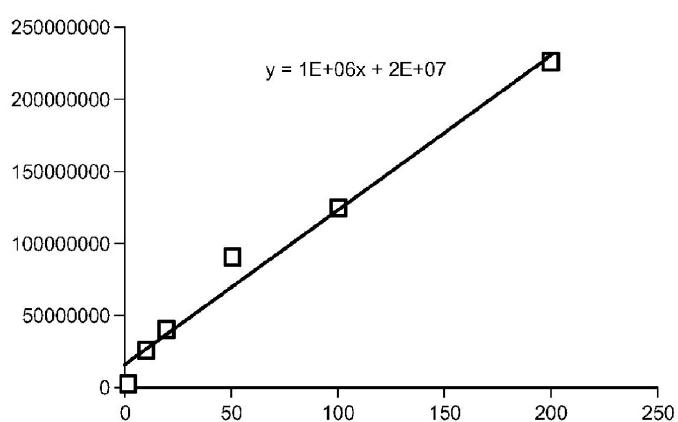


20/40

Фиг. 7А

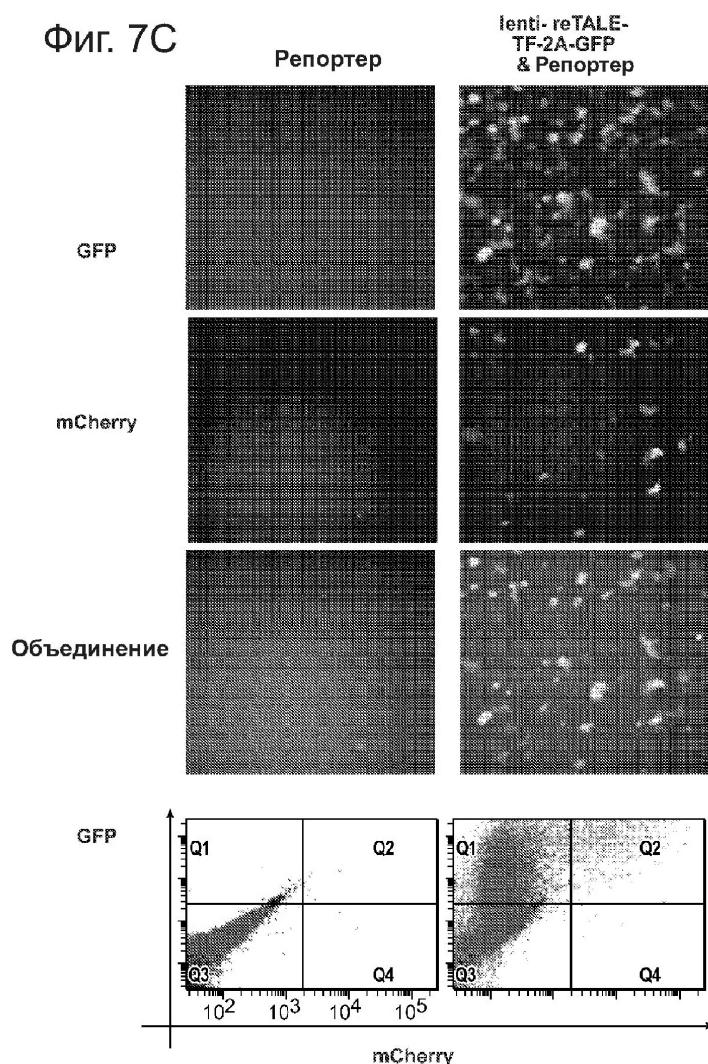


Фиг. 7В



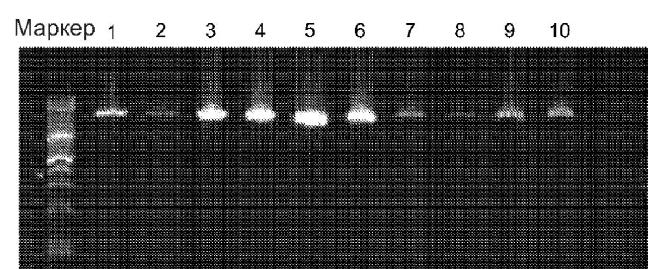
21/40

ФИГ. 7С



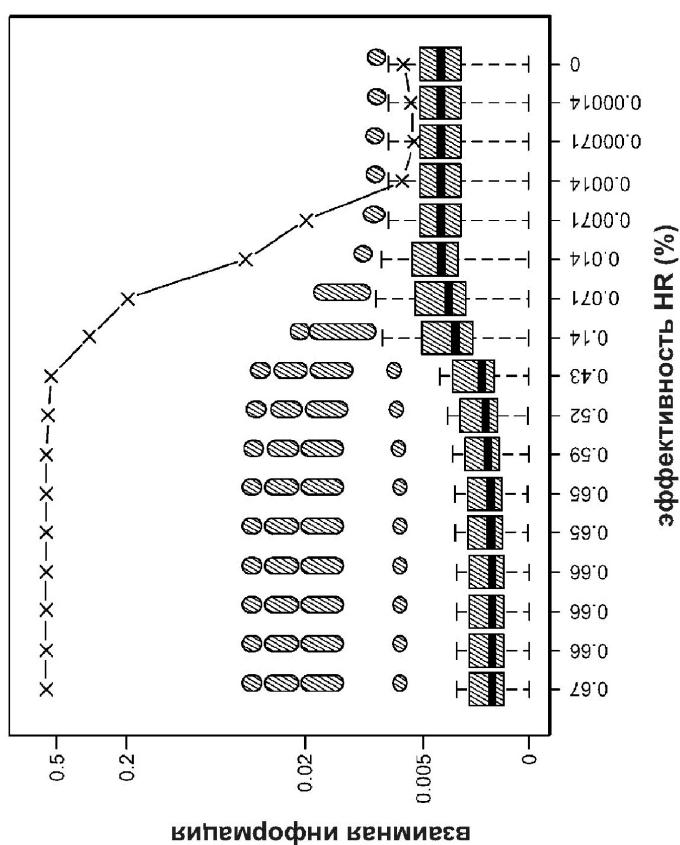
22/40

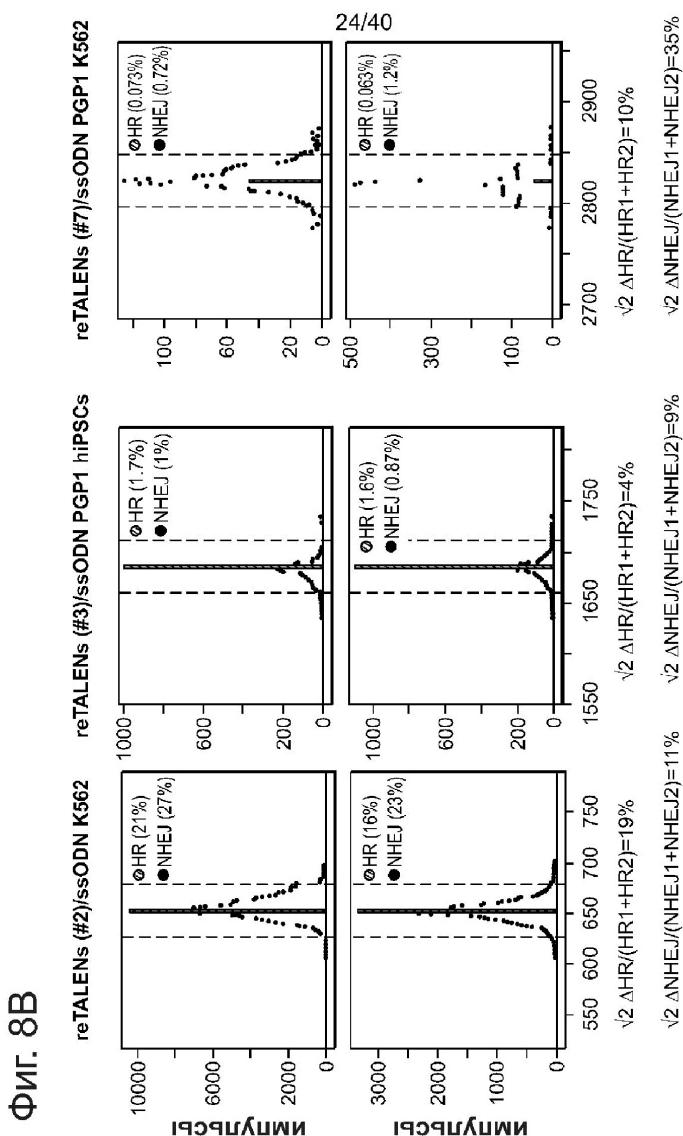
Фиг. 7D



23/40

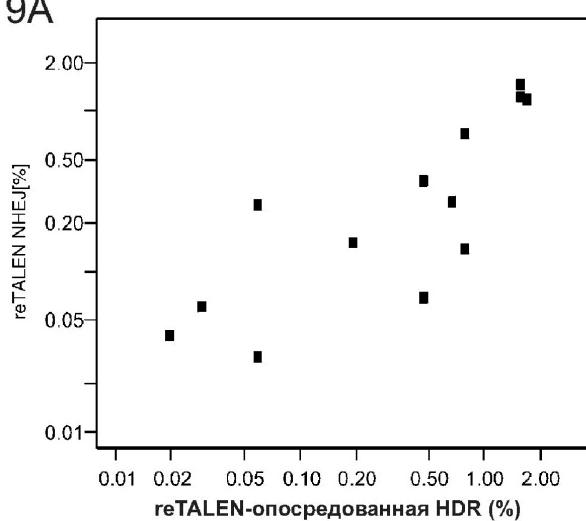
ФИГ. 8А



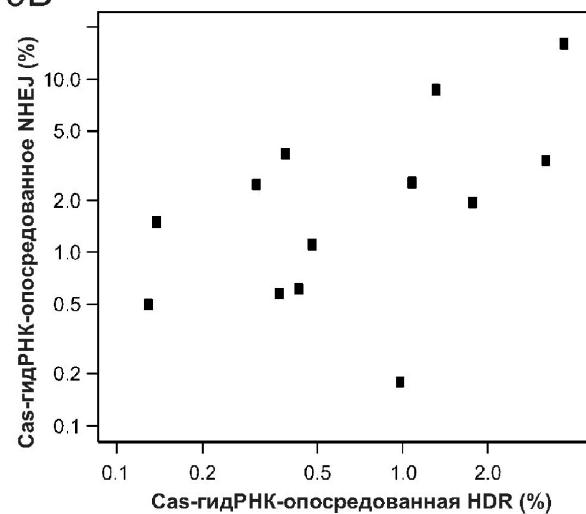


25/40

ФИГ. 9А

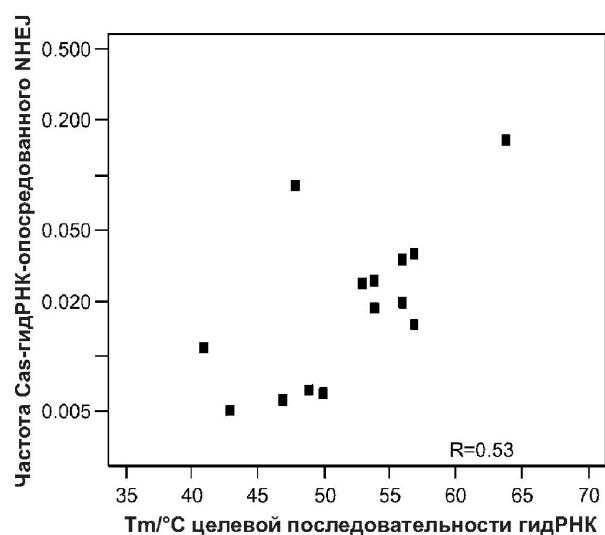


ФИГ. 9В



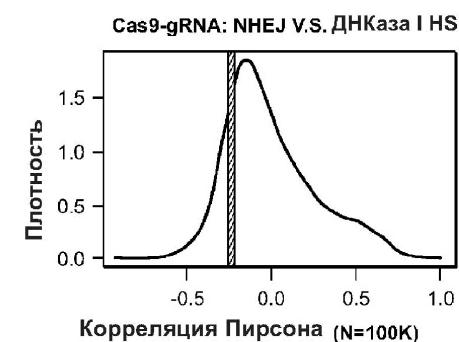
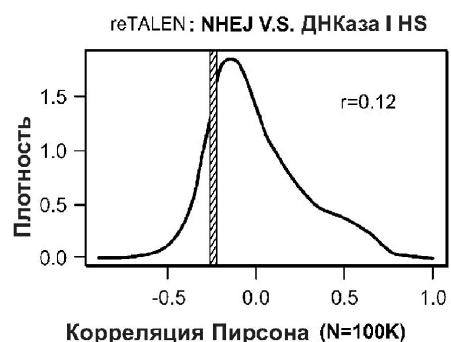
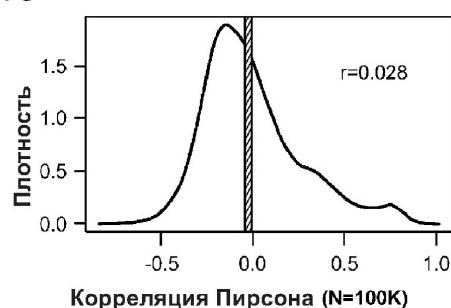
26/40

Фиг. 9С



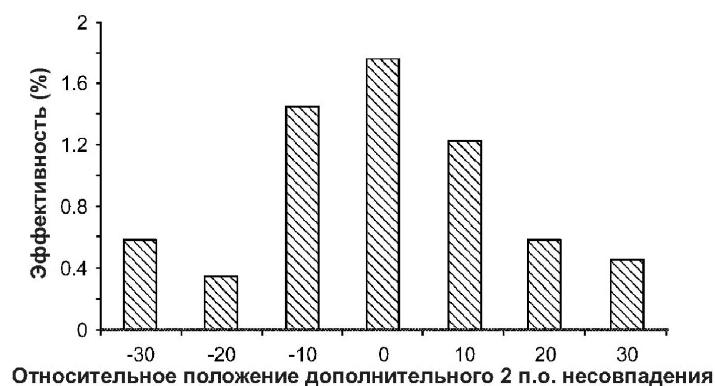
27/40

ФИГ. 10 reTALEN: HDR V.S. ДНКаза I HS

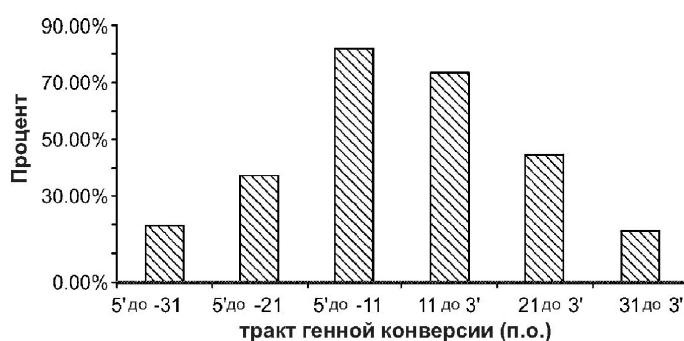


28/40

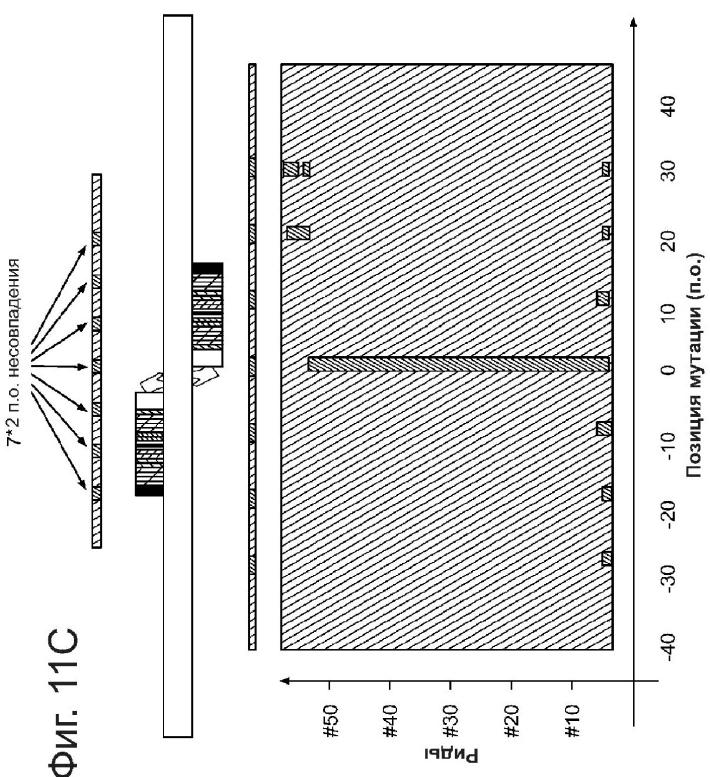
Фиг. 11А



Фиг. 11В

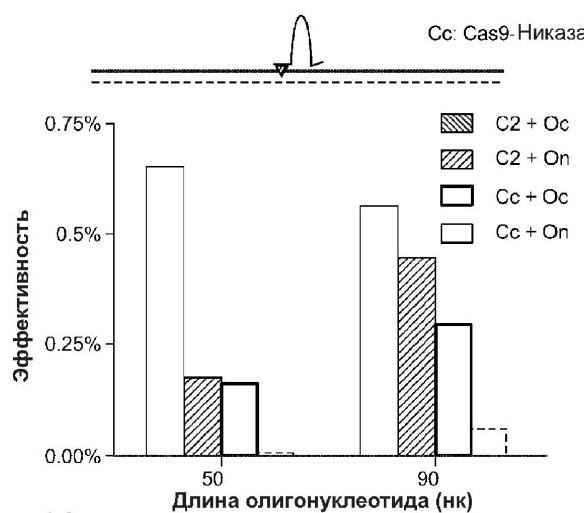


29/40

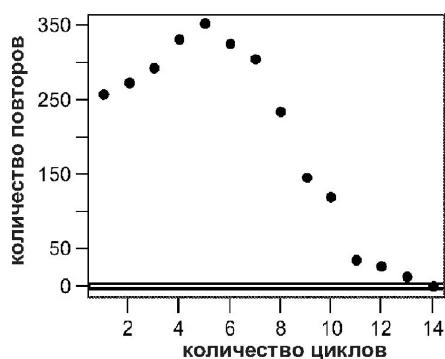


30/40

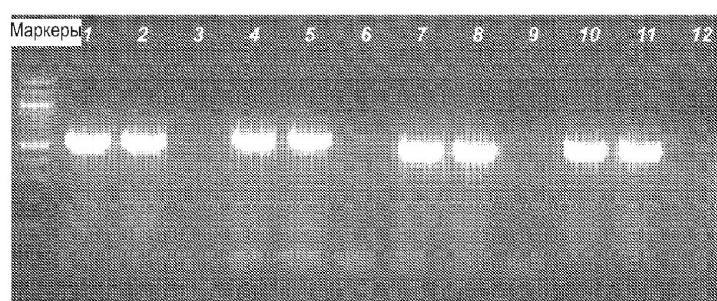
ФИГ. 12



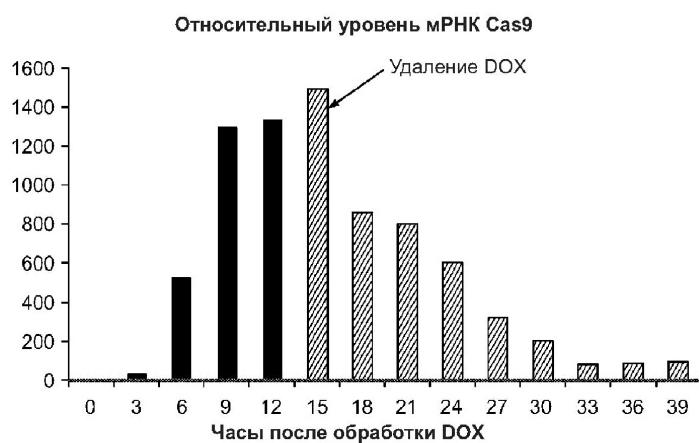
ФИГ. 13



31/40



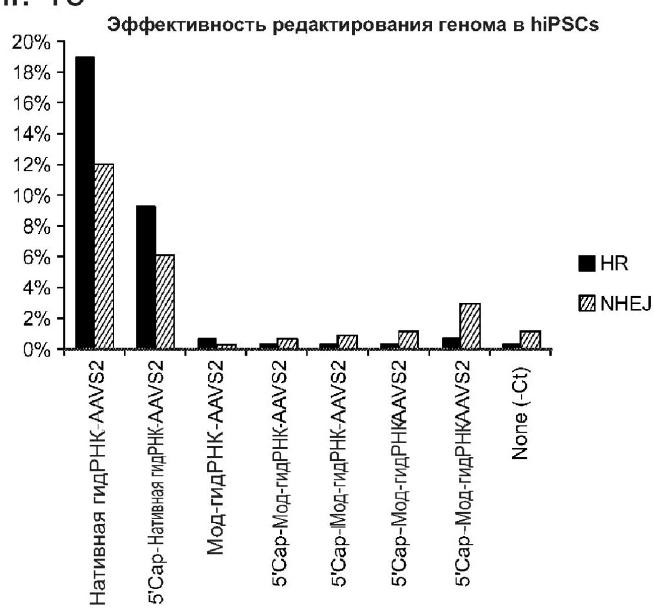
Фиг. 14



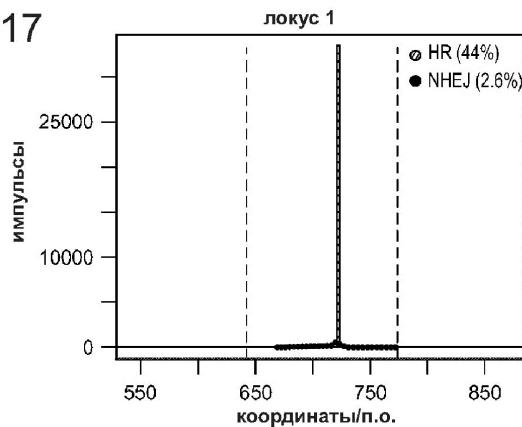
Фиг. 15

32/40

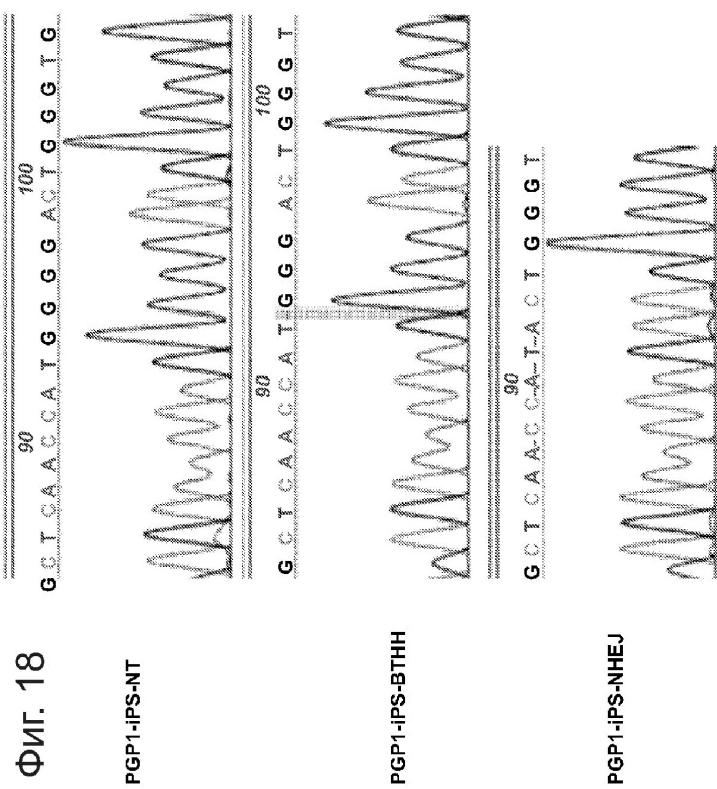
Фиг. 16



Фиг. 17

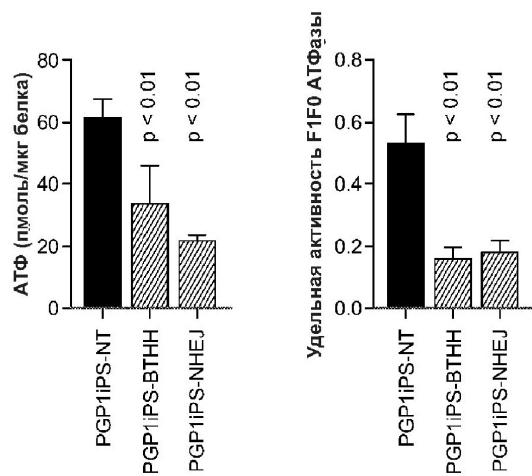


33/40



34/40

ФИГ. 19



ФИГ. 20А

re-TALEN-каркасная последовательность
 (филогетовый re-TALE-N; красный участок *SapI*, зеленый: 0,5 мономера;
 синий: re-TALE-N; оранжевый: FokI)

```

ATGTCGCGGACCCGGCTCCCTCCCCCACCGAACCCAGGCCAGGTTCGGCGAC
TCGTTCTAGACCTGCTTAGGCAGTTCGACCCCTCACTGTTAACACATCGTTCTCG
ACTCCCTTCCCTCGTTGGGGCCACCATACGGAGGGCCACCCGGAGTGGGAT
GAGGTGCACTCGGGATTCGAGCCTGGGAATGCCACACCCCCAACCATGGGCTGGC
CGTCACCGCTGCCGACGCCGATGCCACGGCCAAAGGGCAAGCCCAAGGGCAGGGCAGGC
AACCCTGCCAGCGCAAGGCCCGAACGCCGAAGGCAAGTAGATTGAGAACTTGGGATATCA
CAGCAGCAGGAGAAAGATAAGCTAACGCCAAAGTGAGGTGACAGTCGGCAGCAGC
ACGAAGGGCTGGTGGTCATGGGTTAACATGCCAACATCGTAGGCTTGTGCGAGC
ACCCCTGAGGCCCTTGGCACGGTCGGCTCAACTACGGACATGATTGGGGCTGGTGC
CGGAAGGCCACACATGAGGGATCGTGGTGGGGAAACAGTGGAGGGCCG
AGGGCTTGAAGGGCCAGTTGCTGAAGATCGCAAGGGGGAGGTCAAGGGGTGAG
TGGCACACGGCCATGGCTGGGAGGGGGAGGTCAAGGGAGGAAACAGTGGAGGGCAG
GGGTGCAAGCGTGGCAATGGCTCACGGAGCACCCCAAACAGTTCACTG
ACAGAGACCCGGCCATTAGGACACCCAGGGTTAACCTTACACTTATGGCTTCCGGCTCG
TATAATGTTGGGATTTGAGTTGGATCCGTGAGATTTCAGGGCTAACGGAGCT
AAAATGGAGAAAAAAATCACTGGATAACCAACCGTTGATATAATCCAATGGCATCG
TAAAGAAACATTGGGCAATTGGGCAATTGGCTAACATGTAACCTAACAGAACCGT
TCAGCTGGAAATTACGGCCTTAAAGACGGTAAGAAAATAAGCACAAAGTTTA
TCCGGCCATTTCACATTCTGGCCCTGATGAATGCTCATCCGGAAATTCCGTATG
GCAATGAAAGACGGTGGCTGAGCTGGGATATGGGATAGTGTCAACCTGTACACCGTT
  
```

Фиг. 20В

TICCGATGGCAAACGTGAACTTTCATCGGCTCTGGAGTGAAATACCGACGATTCCGGCAGTTCTACACATAATTCCGAAGATGTGGGTACGGTGAAGAACCTGGCC
TATTTCCTTAAGGGTTATTGAAATATGTTTCGCTCTAGCCATTCGCAACTCTTCGCCATCCCTGGGTGA
GTTTACCAAGTTGATTTAAGCTGAACTTACGAAAGGCACAAAGTGGCCAAATATGGACA
CACCAGGGCAAAATAATACGAAAGGGTCCATGTGGCCAAATGGCTTAATGAATTAC
ACGTTCATCATGCCGTTGTGATGGCTTCCATGTGGCCAAATGGCTTAATGAATTAC
AACAGTACTGCGATGAGTGGCAGGGCAGGGCGGGCGTAAGAACCTGGATCGGCCCTA
AAAGCCAGATAACAGTATGCGTATTGCGCGCTGATTTGCGGTATAAGAAATTAT
ACTGATAATGTTACCCGAAGTGTCAAAAGAGGTATGGCTATGAAAGCAGCTATT
ACAGTGACAGTTGACAGGGACAGCTATCAGTTGCTCAAGGCAATTATGATGTCATA
TCTCCGGCTGGTAAGGACAACCATGCGAAATGCGAGAACTGAAAGCCGTCGCTGCGTGGCGAAC
GCTGGAAAGCGGAAATCAGGAAGGGATGGCTGAGGTGCCCGTTATGAAATIG
AACGGCTCTTTTGCAGCAGGAAACAGGGCTGGTGAATGCGATTAGGTTCAC
CTATAAAAGAGAGAGCCGTTATCGCTGTTGGATACAGGIGATAATTATGAA
CACGCCCGGGGACGGATGGTGAATCCCCTGGCCAGTGACGTCGCTGTCAGATA
AAGTCCTCGTGAACATTACCCGGTGGCATATGGGATGAAAGCTGGCGATGA
TGACCCACCGGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGAAAGAAAGTGGCTGATC
TCAGCCACCGCGAAATGACATCAAAACGCCATTAACTGTGATGTTCTGGGAAATA
TAAATGTCAGGCTCCCTTATACAGCCAGTCTGCAGGTCAGGGCTCACTCCAGAGCAAGTGGTTGC
AAGGTTACTTCCGGTCTGTCAAGCAGGGCCCTCACTGGAAAGACCTGCCCTGGAAATCAATCGTGGCCAGCTTC
GATCGCTTCAAACACGGTGGAAAGACCTGCCCTGGAAATCAATCGTGGCCAGCTTC

ФИГ. 20C

GAGGCCGGACCCCGGCTGGCCACTCACTAATGATACTCTTGTAGCGCTGGCCTG
 CCTCGGGGACGCCGCCTTGATGCGGTGAAGAAGGGGCTCCGGCACGGCCTG
 CATTGATTAAAGGGACCAACAGAAGGATTCCCGAGAGACATCACAATCGAGTGGCA
 GGTTCCTCAAACSTCGTAAAGAGTGAACCTTGAGGAGAAAAGTGGAGCTGGGACAA
ATTTGAAATACGTAACCGCATGAAATACATCGAACTTATCGAAATTGCTAGGAAC
TCAAGACAGAAATCCSTTGAGATGAGGTAAATGGAGTTATGAGGTTTATGGATA
CCGAGGGAGGACATCTCGGTTGATGATACGGGAGTCACTGGAGGAGCAATTCTATA
GGAGCCCCGATTTGATTAATGGAGGAGTCACTGGAGGAGCTAACGAAAGGAC
AATCTTCCCATCGGGAGGAGATGAGTCACTGGAGGAGTCACTGGAGGAGCTAAC
CAGGAACAAACATCAATCCAATGAGTGGTGGAGTGGTGGAGTGGTGGTGGTGG
CCGAGTTTAAGTTTGTGTTGCTCTGGGAATGAACTGCAATTGAGTGGTGGGTTG
CACACGGTTGAATGAACTTACGAACATGAACTTACGAACATATAAGGCCAGCT
GCTCATTTGGAGAAATGATAAAAGGGAAACTGACACTGGAAAGTCAGAC
GCAAGTTAACAAATGGCGAGAATCAATTCCG

каркасная последовательность ге-TALEN-TF

(фиолетовый: ге-TALE-N; красный: участок SapI; зеленый: 0,5 мономера;

синий: ге-TALE-N-C, оранжевый:-NLS-M364;
 \ 2A-GFP выделен зеленым)

ATGTCGCGGACCCGGCTCCCTTCCCAACCOGCCACCCAGGGTTCGGGCGAC
 TCGTTCTAGACCTGCTTAGGCAGTTGACCCCTCACTGTTAACACATCGTTGTTG

ФИГ. 20D

```

ACTCCCTCCTCCGTGTTGGGGCCACCATACGGAGGGGGCACCGGGGAGTGGAT
GAGGTGCAAGTCGGATTGAGGGTGGGATGCACCACCCCAACATGGGGTGGC
CGTCACCGCTGCCGACCGCGAAGGGGAAGGCCGACCAAGGGGAGGGCAGGCG
AACCGTCCGACGCCAAGGCCCGAAGGGCAAGTAGATTGAGAACTTGGGATATTC
CAGCAGCAGGCCAAGGGCAAAGATCAAGGCCAACAGTGAGGTGACAGTCGGCAGCATC
ACGGAGGGCTGGTGGGTCAATGGGTTAACACATGCCAACATCGTAGGCTTGTGCA
ACCGCTGCAAGCCCTTGGCACCGGTTGGCGTCAAGTACCGAGGACATGATTGGGG
CGGAAGGCCACACATGAGGGATCGTCGGTGGGTGGGGAAACAGTGGAGGGAGCCG
AGGGCTTGAGGCCCTGTTGACGGTGGGCCAGTGTGTAAGATGCCAACAGTTCA
TGGGACACGGGCCAGTGTGTAAGATGCCAACGGGAGCTGAGGGCTGAGGGCT
GGGTGCAACGGCTGGCAATGGCTCACGGGAGCACCCCTAACAGTTCA
ACAGAGACCGGGGCCGCAATTAGGGACCCCAGGCTTAACATTAGCTTCCGGCTG
TATAATGGGAAAAAAATCACTGGATATAACCACCGTTGATATAATCCAATGGCATCG
AAAATGGGAAAAAAATCACTGGATATAACCACCGTTGATATAATCCAATGGCATCG
TAAGGAACATTGTAGGCATTTCAGTCAGTGTGCTOAATGTACCTTAACCAGAACCGT
TCAGCTGGATATTAGGCCCTTAAAGACGTAAGAAAAATAAGCACAAGTTTA
TCCGGCCCTTATTACCATTTCTGGCCCTGATGAAATGGCTCATCCGAATTCCGGAT
GCAATGAAAGACGGTGAGCTGGTGAATGGGATAGTGGTACCCUTGTTACACGTT
TICCAATGAGCAAACTGAAACGTTCAUCGCTCTGGAGTGAATACACAGACGATTIC
GGCAGTCTACACATATTCGGCAAGATGTCGGTGTGAGGGGAAACCTGGG
TATTICCCCTAAAGGGTTIATTGAGAATATGTTTTCGCTCAGCCAAATCCCTGGGTGA
GTTTACACAGTTTGATTAAACGTGGCCAAATATGGACAACTTCTCGCCCCGGTTT

```

ФИГ. 20Е

CACCATGGCAATTATTAAGCGAAGGGCAAGGTGCTATGCCGC^TGGCGATTIC
 AGGTTCACTATGCCGTTTGTGATGGCTTCATGTCGGAGAATGC^TAA^TGAATTAC
 AACAGTACTGCGATGAGTGGCAGGGGGCTAAGZTCGGATCCGGCTTACTA
 AAAGCCAGATAAACGATAGCGIATTGGCGCTGATTTCGGGTATAAGAAATAT
 ACTGATAATGTTACCCGAAGTAAAGTCAAGGCAAGGTAAAGGTTAAGCAGCGIATT
 ACAGTGACAGTTGACAAGCGACAGCTATCAGTTGCTCAAGGGATATGA^TGTCATA
 TCTCCGGCTGGTAAGCACAACCATGCGAATGAAGGCCCGTCTGC^TGCGGAAC
 GCTGGAAAGCGGAAATTCAGGAAGGGATGGCTGAGGTGCGCCGGTTATGAAATG
 AACGGCTCTTTTGCTGACGAGAACGGGCTTATCGTTGTTGGTGAATTCAGTAAAGTTACAC
 CTATAAAGAGAGAGGCGTTATCGTTGTTGGATGTTACAGGATGATTTATTGA
 CACGCCCGGGCAGCGATGGTGAATCCCCCTGGCCAGTGCACGTCTGC^TG^CCAGATA
 AAGTCTCCGTGAACCTTACCCGGTGGTGCATACTGGGGATGAAGC^TGGCGATGA
 TGACCCACCGATAGGCCAGTGGCCAGTGGCCGGTCTCGGTATCGGGAAAGAAGTGGCTGATIC
 TCAGGCCACCGCGAAATGACATCAAACGCCATTAAACCTGATGTTCTGGGAATA
 TAAATGTCAGGCTCCCTTATACACAGCCAGCTGCAAGGGTCAACGGCTC^C**GCTCTTCG**
 AAGTTACTTCCGGTCTCTGTAAGGCAGGGCTCACTCCAGGCAAGTGGTTGC
 GATCGCTCAAAACACGGTGGAAAGACCTGCCC^TGGAAATCAATCGTGGCC^CAGCTTC
GAGGCCGGACCCCGGGCTGGCCGACTCACTTAATGATCATCTTGAGGCTGGCTG
CTCTCGGGGACGACCGCCCTTGATGGGTGAAGAACGGGCTCCGGCACGGCTG
CATTGATTAAGGGMCCAACAGAAGGATTCCGGAGAGAACATAGCCCAAGANGAA
CAGAAACCTGGAGCCACCCGCTCCGGACGGCTTGAGGCAATGGACATTTGATC

40/40

ФИГ. 20F