РОССИЙСКАЯ ФЕДЕРАЦИЯ

(19) RU (11) 2 688 462 (13) C2

(51) МПК
C12N 15/09 (2006.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
C12N 15/09 (2019.02)

(21)(22) Заявка: **2016106649, 25.07.2014**

(24) Дата начала отсчета срока действия патента:
25.07.2014

Дата регистрации:
21.05.2019

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

(43) Дата публикации заявки: **31.08.2017** Бюл. № **25**

(45) Опубликовано: **21.05.2019** Бюл. № **15**

(85) Дата начала рассмотрения заявки PCT на национальной фазе: **26.02.2016**

(86) Заявка PCT:
US 2014/048140 (25.07.2014)

(87) Публикация заявки PCT:
WO 2015/013583 (29.01.2015)

Адрес для переписки:
109012, Москва, ул. Ильинка, 5/2, ООО "Союзпатент"

(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: **HOLKERS M. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res. 2013 Mar 1;41(5):e63. doi: 10.1093/nar/gks1446. Epub 2012 Dec 28. THOMAS GAJ, CHARLES A. GERSBACH, and CARLOS F. BARBAS. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in** (см. прод.)

(54) ГЕНОМНАЯ ИНЖЕНЕРИЯ

(57) Реферат:

Изобретение относится к области биотехнологии, молекулярной биологии и биохимии. Предложен способ изменения ДНК-мишени в клетке, экспрессирующей фермент Cas, предусматривающий формирование комплекса колокализации с направляющей РНК, комплементарной ДНК-мишени, который расщепляет ДНК-мишень сайт-специфическим образом, с последующим встраиванием донорной нуклеиновой кислоты в ДНК-мишень. Изобретение позволяет осуществлять редактирование генома с высокой эффективностью и может быть использовано в медицине. 24 з.п. ф-лы, 20 ил., 6 табл., 20 пр.

(56) (продолжение):
Biotechnology 09 may 2013, Vol. 31, No. 7, p.399-402. US 0008450471 B2, 28.05.2013. RU 2267533 C2, 10.01.2006.

RU 2688462 C2

RU 2688462 C2

RUSSIAN FEDERATION

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(19) **RU** (11) **2 688 462** (13) **C2**

(51) Int. Cl.
C12N 15/09 (2006.01)

(12) ABSTRACT OF INVENTION

(52) CPC
C12N 15/09 (2019.02)

(21)(22) Application: **2016106649, 25.07.2014**

(24) Effective date for property rights:
25.07.2014

Registration date:
21.05.2019

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

(43) Application published: **31.08.2017** Bull. № **25**

(45) Date of publication: **21.05.2019** Bull. № **15**

(85) Commencement of national phase: **26.02.2016**

(86) PCT application:
US 2014/048140 (25.07.2014)

(87) PCT publication:
WO 2015/013583 (29.01.2015)

Mail address:
109012, Moskva, ul. Ilinka, 5/2, OOO "Soyuzpatent"

(72) Inventor(s):
CHERCH Dzhordzh M. (US),
YAN Lukhan (US),
KARGOL Mark Guel (US),
YAN Dzhojs Lichi (US)

(73) Proprietor(s):
PREZIDENT END FELLOUZ OF KHARVARD KOLLIDZH (US)

(54) GENOMIC ENGINEERING

(57) Abstract:

FIELD: biotechnology.

SUBSTANCE: invention refers to biotechnology, molecular biology and biochemistry. Disclosed is a method for changing a DNA target in a Cas-expressing cell expressing a complex of colocalization with an RNA guide complementary to a DNA target which cleaves the DNA target in a site-specific manner, followed by insertion of the donor nucleic acid into the target DNA.

EFFECT: invention makes it possible to perform genome editing with high efficiency and can be used in medicine.

25 cl, 20 dwg, 20 ex, 6 tbl

RU 2688462 C2

RU 2688462 C2

Данные связанных заявок

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

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.

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 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 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 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 траснкрипции) 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.

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, 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 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 последовательность нуклеиновой кислоты, кодирующую TALEN, утратившую повторяющиеся последовательности в 100 п. н. 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 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 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 множестве 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, 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 сайт-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.

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 сайт-специфическим образом.

Аспекты настоящего раскрытия относятся к клетке, включающую 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, 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 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-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 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 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 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, 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 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 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, 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 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 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 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 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 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.

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 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 одна или несколько последовательностей донорной нуклеиновой кислоты экспрессируются, где одна или несколько РНК и фермент колокализуются с ДНК-мишенью, фермент расщепляет

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

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

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 нуклеиновой кислоты, кодирующей ДНК-связывающий фермент, с использованием вектора, который может быть легко удален с помощью фермента. 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, нуклеиновая кислота, кодирующая ДНК-связывающую фермент, который был добавлен, например, с помощью кассеты или вектора, к примеру, может быть удалена полностью вместе с кассетой и вектором, не оставляя часть такой нуклеиновой 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 нуклеиновой кислоты, кассеты или вектора.

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.

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:

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

(a) 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 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 .

(b) Столбиковая диаграмма, демонстрирующая процент GFP^+ клеток, полученных с помощью только tGFP-донора отдельно, с помощью TALENs с tGFP-донором и re-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 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).

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 re-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 re-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 re-TALENs, которые находятся в положениях -26 п.о. и + 26 п.о. относительно центра двух участков связывания re-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 PAM. 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 котрансфецировали парой re-TALENs и их соответствующего донорного ssODN, несущего 2 п.о. несовпадение относительно геномной ДНК. Эффективность редактирования генома анализировали через 6 дней после трансфекции. Аналогичным образом, 15 Cas9-гидРНК трансфецировали 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 re-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 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 NIHi7 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.

(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 эффективности делеции и вставки описана в данном документе.

(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, смещенных 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.

(c) ssODNs, нацеленные на участки с варьирующими расстояниями (-620п.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 располагалось на расстоянии 40 п.о. от середины участка связывания пары re-TALEN.

(d) PGP1 hiPSCs котрансфецировали Cas9-гидРНК (AAVS1) и ssODNs различной ориентации (Oc: комлпементарная гидРНК; On: некомплементарная 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. 70-мерный Oc 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.

(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, PGPl-IPS-3-13.

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

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

Фиг. 5. Дизайн reTALE. (a) 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, (b) Тест повторяемости re-TALE с помощью ПЦР. Верхняя панель показывает структуру 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), (a) Схематическое представление библиотеки димерных блоков re-TALE для 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. Нуклеотидная и аминокислотная последовательности одного димера (Блок 6_AC) представлены справа.

(b) 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 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 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.

(c) 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.

(A) 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 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 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.

(b) 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 (|HDRl-HDR2|)/((HDR+HDR2)/2) =AHDR/HDR и V2 (|NHEJ1-NHEJ2|) /((NHEJl+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.

(a) Корреляция эффективностей HR и NHEJ при использовании reTALENs на одинаковых участках в iPSCs (i= 0,91, P <IX 10-5).

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

(c) Корреляция эффективности NHEJ, опосредованной Cas9-гидРНК и температуры Tm целевого участка для гидРНК в iPSCs (г = 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 1 и эффективностью NHEJ/HR.

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

(a) В эксперименте, описанном на рисунке 3b, общая HDR снижалось, 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.

(b) 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 шкала дистанций в десятки п.о. для одноцепочечных олигодезоксинуклеотидов по сравнению с сотнями пар оснований для дцДНК-доноров.

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

Фиг. 12. Эффективности редактирования генома Cas9-гидРНК нуклеазой и никазами. PGP1 iPSCs котрансфецировали 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.

Фиг. 13. Дизайн и оптимизация последовательности re-TALE.

Последовательность re-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 эволюционирующей ДНК. Окончательная последовательность имела cai = 0,59 ΔG = -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.

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 направляющей РНК- донорной ДНК.

Фиг. 18 является схемой, демонстрирующей генотип изогенных клеточных линий PGP 1, полученных с помощью системы, описанной в данном документе. PGPl-IPS-BTHH имеет фенотип с однонуклеотидными делециями, как у пациентов с BTHH. PGP 1-NHEJ имеет 4 п.о. делеции, которые образуют мутации со сдвигом рамки считывания 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 re-TALEN-основной и основной последовательности re-TALE-TF.

Подробное описание

Аспекты настоящего изобретения относятся к применению TALEN, которая утратила определенные повторяющиеся последовательности, для инженерии нуклеиновых 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 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).

Эфекторные 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 (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 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 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 доменом Fokl, 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 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 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; 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): el9509; 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): el9722, 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 аллелеями, 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 трансфецированные 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-рерио.

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 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 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 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.

В S. pyogenes, Cas9Cas9 образует двухцепочечной разрыв с тупыми концами на 3 п.о. выше мотива PAM (прилегающий к протоспейсеру мотив) с помощью процесса, опосредованного двумя каталитическими доменами в белке: доменом HNH, который расщепляет комплементарную нить ДНК и доменом подобного RuvC, который расщепляет некомплементарную цепь. См. Jinke et al., Science, 337, 816-821 (2012), 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 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 flavida DSM 17836 uid43465; Thermomonospora curvata DSM 43183; Bifidobacterium dentium Bdl; 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 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 gallolyticus UCN34 uid46061; Streptococcus gordonii Challis subst CHI; Streptococcus mutans NN2025 uid46353; Streptococcus mutans; Streptococcus pyogenes Ml 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; 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; Rhodospirillum rabram ATCC 11170; Diaphorobacter TPSY uid29975; Verminephrobacter eiseniae EF01-2; Neisseria meningitides 053442; Neisseria meningitides alphal4; Neisseria meningitides 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 T6c; Shewanella pealeana ATCC 700345; Legionella pneumophila Paris; Actinobacillus succinogenes 130Z; Pasteurella multocida; Francisella tularensis novicida U112; Francisella tularensis holarctica; 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 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.

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES
EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
GETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEA
KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE
KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIR
EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL
GGD-

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. pyogenes и белковых последовательностях, гомологичных, по меньшей мере, на 30%, 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 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 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 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 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 модификаций ДНК. Способы настоящего описания соответственно направлены на редактирование генов-мишеней 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-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.

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 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, в случае которой TALEN или направляемый РНК ДНК-связывающий белок, обладающий нуклеазной активностью, как описано в данном документе, могут быть полезны для внесения одноцепочечных 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 внесения одноцепочечных 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 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 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 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, 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 мишеней CRISPR-Cas в геноме дрожжей, и т. д. Дополнительные источники, включая информацию, материалы и методы, полезные для специалиста в данной области 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. 1126fscience. 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,l., 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 любого способа, известного специалистам в данной области техники для такого введения. Такие способы включают трансфекцию, трансдукцию, вирусную трансдукцию, микроинъекции, липофекцию, нуклеофекцию, бомбардировку наночастицами, трансформацию, конъюгацию 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 нуклеиновой кислоты, как описано в данном 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.

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

(TTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGN19GTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCC).

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 AflII и очищали вектор. Реакцию сборки (10 мкл) гидРНК 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

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 последовательность подвергали эволюции в течение нескольких циклов для устранения 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 эволюционирующей ДНК. Последовательность одного из re-TALE, обладающего 16,5 мономерами представлена ниже

CTAACCCCTGAACAGGTAGTCGCTATAGCTTCAAATATCGGGGGCAAGCAAG
CACTTGAGACCGTTCAACGACTCCTGCCAGTGCTCTGCCAAGCCCATGGATTGACTC
CGGAGCAAGTCGTCGCGATCGCGAGCAACGGCGGGGGGAAGCAGGCGCTGGAAAC
TGTTCAGAGACTGCTGCCTGTACTTTGTCAGGCGCATGGTCTCACCCCCGAACAGGT
TGTCGCAATAGCAAGTAATATAGGCGGTAAGCAAGCCCTAGAGACTGTGCAACGCC
TGCTCCCCGTGCTGTGTCAGGCTCACGGTCTGACACCTGAACAAGTTGTCGCGATAG
CCAGTCACGACGGGGGAAAACAAGCTCTAGAAACGGTTCAAAGGTTGTTGCCCGTT
CTGTGCCAAGCACATGGGTTAACACCCGAACAAGTAGTAGCGATAGCGTCAAATAA
CGGGGGTAAACAGGCTTTGGAGACGGTACAGCGGTTATTGCCGGTCCTCTGCCAGG
CCCACGGACTTACGCCAGAACAGGTGGTTGCAATTGCCTCCAACATCGGCGGGAAA
CAAGCGTTGGAAACTGTGCAGAGACTCCTTCCTGTTTTGTGTCAAGCCCACGGCTTG
ACGCCTGAGCAGGTTGTGGCCATCGCTAGCCACGACGGAGGGAAGCAGGCTCTTGA
AACCGTACAGCGACTTCTCCCAGTTTTGTGCCAAGCTCACGGGCTAACCCCCGAGCA
AGTAGTTGCCATAGCAAGCAACGGAGGAGGAAAACAGGCATTAGAAACAGTTCAG
CGCTTGCTCCCGGTACTCTGTCAGGCACACGGTCTAACTCCGGAACAGGTCGTAGCC
ATTGCTTCCCATGATGGCGGCAAACAGGCGCTAGAGACAGTCCAGAGGCTCTTGCCT
GTGTTATGCCAGGCACATGGCCTCACCCCGGAGCAGGTCGTTGCCATCGCCAGTAAT
ATCGGCGGAAAGCAAGCTCTCGAAACAGTACAACGGCTGTTGCCAGTCCTATGTCA
AGCTCATGGACTGACGCCCGAGCAGGTAGTGGCAATCGCATCTCACGATGGAGGTA

AACAAGCACTCGAGACTGTCCAAAGATTGTTACCCGTACTATGCCAAGCGCATGGTT
TAACCCCAGAGCAAGTTGTGGCTATTGCATCTAACGGCGGTGGCAAACAAGCCTTG
GAGACAGTGCAACGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAG
CAAGTCGTAGCTATCGCCAGCAACATAGGTGGGAAACAGGCCCTGGAAACCGTACA
ACGTCTCCTCCCAGTACTTTGTCAAGCACACGGGTTGACACCGGAACAAGTGGTGGC
GATTGCGTCCAACGGCGGAGGCAAGCAGGCACTGGAGACCGTCCAACGGCTTCTTC
CGGTTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAGGTGGTAGCAATAGCGTCG
AACATCGGTGGTAAGCAAGCGCTTGAAACGGTCCAGCGTCTTCTGCCGGTGTTGTGC
CAGGCGCACGGACTCACACCAGAACAAGTGGTTGCTATTGCTAGTAACAACGGTGG
AAAGCAGGCCCTCGAGACGGTGCAGAGGTTACTTCCCGTCCTCTGTCAAGCGCACG
GCCTCACTCCAGAGCAAGTGGTTGCGATCGCTTCAAACAATGGTGGAAGACCTGCC
CTGGAA

В соответствии с некоторыми аспектами, могут быть использованы TALEs, имеющие, 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. Специалист легко поймет, где указанная последовательность может варьировать при сохранении ДНК-связывающей активности TALE.

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

Блок 0	CGCAATGCGCTCACGGGAGCACCCCTCAACCTAACCCCTGAACAGGTAGTCGCTATAGCTTCANNNNNN GGGGGCAAGCAAGCACTTGAGACCGTTCAACGACTCCTGCCAGTGCTCTGCCAAGCCCATGGATTGACTC CGGAGCAAGTCGTCGCGATCGCGAGCNNNNNNGGGGGGAAGCAGGCGCTGGAAACTGTTCAGAGACTG CTGCCTGTACTTTGTCAGGCGCATGGTCTC
Блок 1	AGACTGCTGCCTGTACTTTGTCAGGCGCATGGTCTCACCCCCGAACAGGTTGTCGCAATAGCAAGmNNNN NGGCGGTAAGCAAGCCCTAGAGACTGTGCAACGCCTGCTCCCCGTGCTGTGTCAGGCTCACGGTCTGACA CCTGAACAAGTTGTCGCGATAGCCAGTNNNNNNGGGGGAAAACAAGCTCTAGAAACGGTTCAAAGGTTG TTGCCCGTTCTGTGCCAAGCACATGGGTTA
Блок 1'	TGCGCTCACGGGAGCACCCCTCAACCTCACCCCCGAACAGGTTGTCGCAATAGCAAGTNNNNNNGGCGG TAAGCAAGCCCTAGAGACTGTGCAACGCCTGCTCCCCGTGCTGTGTCAGGCTCACGGTCTGACACCTGAA CAAGTTGTCGCGATAGCCAGTNNNNNNGGGGGAAAACAAGCTCTAGAAACGGTTCAAAGGTTGTTGCCC GTTCTGTGCCAAGCACATGGGTTA
Блок 2	AGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGTTAACACCCGAACAAGTAGTAGCGATAGCGTCANNNN NNGGGGGTAAACAGGCTTTGGAGACGGTACAGCGGTTATTGCCGGTCCTCTGCCAGGCCCACGGACTTAC GCCAGAACAGGTGGTTGCAATTGCCTCCNNNNNNGGCGGGAAACAAGCGTTGGAAACTGTGCAGAGACT CCTTCCTGTTTTGTGTCAAGCCCACGGCTTGACGCCT
Блок 3	AGACTCCTTCCTGTTTTGTGTCAAGCCCACGGCTTGACGCCTGAGCAGGTTGTGGCCATCGCTAGCNNNN NNGGAGGGAAGCAGGCTCTTGAAACCGTACAGCGACTTCTCCCAGTTTTGTGCCAAGCTCACGGGCTAAC CCCCGAGCAAGTAGTTGCCATAGCAAGCNNNNNNGGAGGAAAACAGGCATTAGAAACAGTTCAGCGCTT GCTCCCGGTACTCTGTCAGGCACACGGTCTA
Блок 4	CGCTTGCTCCCGGTACTCTGTCAGGCACACGGTCTAACTCCGGAACAGGTCGTAGCCATTGCTTCCNNNN NNGGCGGCAAACAGGCGCTAGAGACCGTCCAGAGGCTCTTGCCTGTGTTATGCCAGGCACATGGCCTCA CCCCGGAGCAGGTCGTTGCCATCGCCAGTNNNNNNGGCGGAAAGCAAGCTCTCGAAACAGTACAACGGC TGTTGCCAGTCCTATGTCAAGCTCATGGACTG

Блок 5	CGGCTGTTGCCAGTCCTATGTCAAGCTCATGGACTGACGCCCGAGCAGGTAGTGGCAATCGCATCTNNNN NNGGAGGTAAACAAGCACTCGAGACTGTCCAAAGATTGTTACCCGTACTATGCCAAGCGCATGGTTTAAC CCCAGAGCAAGTTGTGGCTATTGCATCTTSnSINNNNGGTGGCAAACAAGCCTTGGAGACCGTGCAACGAT TACTGCCTGTCTTATGTCAGGCCCATGGCCTT
Блок 6	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAGGTGGTCGCTATCGCCAGCNNNN NNGGGGGCAAGCAAGCACTGGAAACAGTCCAGCGTTTGCTTCCAGTACTTTGTCAGGCGCATGGATTGAC ACCGGAACAAGTGGTGGCTATAGCCTCANNNNNNGGAGGAAAGCAGGCGCTGGAAACCGTCCAACGTCT TTTACCGGTGCTTTGCCAGGCGCACGGGCTC
Блок 6'	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAAGTCGTAGCTATCGCCAGCNNNN NNGGTGGGAAACAGGCCCTGGAAACCGTACAACGTCTCCTCCCAGTACTTTGTCAAGCACACGGGTTGAC ACCGGAACAAGTGGTGGCGATTGCGTCCNNNNNNGGAGGCAAGCAGGCACTGGAGACCGTCCAACGGC TTCTTCCGGTTCTTTGCCAGGCTCATGGGCTC
Блок 7	CGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAGGTGGTAGCAATAGCGTCGNNNN NNGGTGGTAAGCAAGCGCTTGAAACGGTCCAGCGTCTTCTGCCGGTGTTGTGCCAGGCGCACGGACTCAC ACCAGAACAAGTGGTTGCTATTGCTAGTTSnsnsnsnSINGGTGGAAAGCAGGCCCTCGAGACGGTGCAGAG GTTACTTCCCGTCCTCTGTCAAGCGCACGGCCTC

Таблица 2. Последовательности праймеров блоков re-TALE

Блок O-F	CGCAATGCGCTCACGGGAGCACCCCTCAACctAACCCCTGAACAGGT*A*G
Блок O-R	GAGACCATGCGCCTGACAAAGTACAGGCAGCAGTCTCTGAACAG*T*T
Блок 1'-F	TGGCGCAATGCGCTCACGGGAGCACCCCTCA*A*C
Блок 1-F	AGACTGCTGCCTGTACTTTGTCAGGCGCATGGTCTCACCCCCGAACA*G*G
Блок -R/Блок l'-R	TAACCCATGTGCTTGGCACAGAACGGGCAACAACCTTTGAACCG*T*T
Блок 2-F	AGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGTTAACACCCgaac*a*a
Блок 2-R	AGGCGTCAAGCCGTGGGCTTGACACAAAACAGGAAGGAGTCTCTGCACAG*T*T
Блок 3-F	AGACTCCTTCCTGTTTTGTGTCAAGCCCACGGCTTGACGCCTG*A*G
Блок 3-R	TAGACCGTGTGCCTGACAGAGTACCGGGAGCAAGCGCT*G*A
Блок 4-F	CGCTTGCTCCCGGTACTCTGTCAGGCACACGGTCTAA*C*T
Блок 4-R	CAGTCCATGAGCTTGACATAGGACTGGCAACAGCCGTT*G*T
Блок 5-F	CGGCTGTTGCCAGTCCTATGTCAAGCTCATGGACTGA*C*G
Блок 5-R	AAGGCCATGGGCCTGACATAAGACAGGCAGTAATCGTT*G*C
Блок 6-F	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTA*C*T
Блок 6-R	GAGCCCGTGCGCCTGGCAAAGCACCGGTAAAAGACGTTGGA*C*G
Блок 6'-F	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAA
Блок 6'-R	GAGCCCATGAGCCTGGCAAAGAACCGGAAGAAGCCGTT*G*G
Блок 7-F	CGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAGG*T*G
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 MgC12, 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 IIs, 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 pfu, 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.

Пример III

Клеточная линия и Клеточная культура

Клетки PGP1 iPS поддерживали в покрытых матригелем (BD Biosciences) планшетах в mTeSRl (StemCell Technologies). Культуры пересевали каждые 5-7 дней с TrypLE Express (Invitrogen). Клетки 293T и 293FT выращивали и поддерживали в модифицированной 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 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 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-луночный планшет и трансфецировали их 1 мкг каждой плазмиды re-TALENs и 2 мкг плазмиды 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, 25 000 событий анализировали для каждого трансфецированного образца. Для

экспериментов по направленному воздействию на эндогенный локус AAVS1 в клетках 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

Лентивирусные векторы были созданы с помощью стандартных методов ПЦР и клонирования. Лентивирусные плазмиды трансфецировали 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 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, лентивирус-трансдуцированные клетки 293T трансфецировали 30 нг плазмиды, несущей репортер mCherry и 500 нг плазмиды pUC19 с помощью Lipofectamine 2000 (Invitrogen). Изображения клеток анализировали с помощью Axio Observer Zl (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 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) 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 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 использовано 2 мкг донорной двухцепочечной ДНК, а среда mTeSRl 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 (начало) /chr3:	re-TALENs целевой участок пары (конец) /chr3:	re-TALEN-L целевая последовательность	re-TALEN-R целевая последовательность	гидРНК целевая последовательность	гидРНК начальная позиция целевой последовательности
1	46409942	46409993	TCCCCACTTTCTTGTGAA	TAACCACTCAGGACAGGG	CACTTTCTTGTGAATCCTT	46409946
2	46410227	46410278	TCACACAGCAAGTCAGCA	TAGCGGAGCAGGCTCGGA	TGGGCTAGCGGAGCAGGCT	46410264
3	46411260	46411311	TACCCAGACGAGAAAGCT	TCAGACTGCCAAGCTTGA	ACCCAGACGAGAAAGCTGA	46411261
4	46411464	46411515	TCTTGTGGCTCGGGAGTA	TATTGTCAGCAGAGCTGA	AGAGGGCATCTTGTGGCTC	46411456
5	46411517	46411568	TTGAGATTTTCAGATGTC	TATACAGTCATATCAAGC	ATCAAGCTCTCTTGGCGGT	46411538
6	46411634	46411685	TTCAGATAGATTATATCT	TGCCAGATACATAGGTGG	GCTTCAGATAGATTATATC	46411632
7	46412396	46412447	TTATACTGTCTATATGAT	TCAGCTCTTCTGGCCAGA	ACGGATGTCTCAGCTCTTC	46412437
8	46412432	46412483	TGGCCAGAAGAGCTGAGA	TTACCGGGGAGAGTTTCT	CCGGGGAGAGTTTCTTGTA	46412461
9	46412750	46412801	TTTGCAGAGAGATGAGTC	TTAGCAGAAGATAAGATT	GAAATCTTATCTTCTGCTA	46412782
10	46413152	46413203	TATAAGACTAAACTACCC	TCGTCTGCCACCACAGAT	AATGCATGACAT TCATCTG	46413172
11	46414305	46414356	TAAAACAGTTTGCATTCA	TATAAAGTCCTAGAATGT	AACAGTTTGCATTCATGGA	46414308
12	46414608	46414659	TGGCCATCTCTGACCTGT	TAGTGAGCCCAGAAGGGG	CCAGAAGGGGACAGTAAGA	46414632
13	46414768	46414820	TAGGTACCTGGCTGTCGT	TGACCGTCCTGGCTTTTA	CTGACAATCGATAGGTACC	46414757
14	46415017	46415068	TGTCATGGTCATCTGCTA	TCGACACCGAAGCAGAGT	ACACCGAAGCA	46415046
15	46420034	46420084	TGCCCCCGCGAGGCCACA	TCTGGAAGTTGAACACCC	GGAAGTTGAACACCCTTGC	46420062

Таблица 4. Дизайн ssODN для изучения ssODN-опосредованного редактирования генома

Фигура 3b	Дистанция между второй мутацией и DSB	90-*1	CTACTGTCATTCAGGGCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90-*2	CTACTGTCATTCAGCCCAATACCCTAACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90-*3	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAAGTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90M-0	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90-*4	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTGTAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90-*5	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCTCTCTGACTACAGAGGCCACTGGCTT
		90-*6	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTAGTGAGGCCACTGGCTT
Фигура 3c	Дистанция между ssODN и DSB	L670 bp_9 0M	CACTTTATATTTCCCTGCTTAAACAGTCCCCCGAGGGTGGGTGCGGAAAAGGCTCTACACTTGTTATCATTCCCTCTCCACCACAGGCAT
		L570 bp_9 OM	TAAGGCACAGAGCTTCAATAATTTGGTCAGAGCCAAGTAGCAG
		L480 bp_9 OM	GGAGGTTAAACCCAGCAGCATGACTGCAGTTCTTAATCAATGCCCCTTGAATTGCACATATGGGATGAACTAGAACATTTTCTCGATGAT
		L394 bp_9 OM	CTCGATGATTCGCTGTCCTTGTTATGATTATGTTACTGAGCTCTACTGTAGCACAGACATATGTCCCTATATGGGGCGGGGGTGGGGGTG
		L290 bp_9 OM	GGTGTCTTGATCGCTGGGCTATTTCTATACTGTTCTGGCTTTTCGGAAGCAGTCATTTCTTTCTATTCTCCAAGCACCAGCAATTAGCTT

		L200 bp_9 OM	GCTTCTAGTTTGCTGAAACTAATCTGCTATAGACAGAGACTCCGACGAACCAATTTTATTAGGATTTGATCAAATAAACTCTCTCTGACA
		L114 bp_9 OM	GAAAGAGTAACTAAGAGTTTGATGTTTACTGAGTGCATAGTATGCACTAGATGCTGGCCGTGGATGCCTCATAGAATCCTCCCAACAACT
		L45b p_90 M	GCTAGATGCTGGCCGTGGATGCCTCATAGAATCCTCCCAACAACCGATGAAATGACTACTGTCATTCAGCCCAATACCCAGACGAGAAAG
		R40b p_90 M	ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTTTACCCCTGGGTTAGTCTGCCTCTGTAGGATTGGGGGCACGTAATTT
		R100 bp_90 M	TTAGTCTGCCTCTGTAGGATTGGGGGCACGTAATTTTGCTGTTTAAG
		R200 bp_90 M	GGAAGCCCAGAGGGCATCTTGTGGCTCGGGAGTAGCTCTCTGCTACCTTCTCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACC
		R261 bp_90 M	TCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACCAACCAGCAAGAGAGCTTGATATGACTGTATATAGTATAGTCATAAAGAAC
		R322 bp_90 M	CATAAAGAACCTGAACTTGACCATATACTTATGTCATGTGGAAATCTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAAGAATCCTG
		R375 M_90 M	GTGGAAAATTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAGCAATCCTGCCACCTATGTATCTGGCATAGTGTGAGTCCTCATAAA
		R448 bp_90 M	GGTTTGAAGGGCAACAAAATAGTGAACAGAGTGAAAATCCCCACCTAGATCCTGGGTCCAGAAAAAGATGGGAAACCTGTTTAGCTCACC
		Комплементарный-30-мер	GGCCACTAGGGACAAAATTGGTGAcagaaa

Фигура 3d	Длина и ориентация ssODN для направленного воздействия Cas9-гидРНК	Комплементарный-50-мер	CCCACAGTGGGGCCACTAGGGACAAAATTGGTGAcagaaaagccccatcc
		Комплементарный-70-мер	TCCCCTCCACCCCACAGTGGGGCCACTAGGGACAAAATTGGTGAcag aaaagccccatccttaggcctcc
		Комплементарный-90-мер	cttTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACAAAAT TGGTGAcagaaaagccccatccttaggcctcctccttcctag
		Комплементарный-110-мер	gttctgggtacttTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTAGGGA CAAAATTGGTGAcagaaaagccccatccttaggcctcctccttcctagtctcctgata
		Некомплементарный - 30-Мер	TTTCTGTCACCAATGGTGTCCCTAGTGGCC
		Некомплементарный -50-мер	GGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGG
		Некомплементарный -70-Мер	GGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGA
		Некомплементарный-90-мер	CTAGGAAGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGACAGATAAAAG
		Некомплементарный-110-мер	TATCAGGAGACTAGGAAGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGACAGATAAAAGTACCCAGAAC
Фигура 2c	Донорный ssO DN для направленного воздействия Cas9-гидРНК на CC R5	Cas9 гидРНК-CCR 5-1	TTCTAGTAACCACTCAGGACAGGGGGGTTCAGCCCAAAAATTCACA AGAAAGTGGGGACCCATGGGAAAT

Cas9 гидРНК -CCR 5-2	CAGCAAGTCAGCAGCACAGCGTGTGTGACTCCGAGGGTGCTCC GCT AGCCCACATTGCCCTCTGGGGGTG
Cas9 гидРНК -CCR 5-3	GTCAGACTGCCAAGCTTGAAACCTGTCTTACCCTCTACTTTCTC GTC TGGGTATTGGGCTGAATGACAGT
Cas9 гидРНК -CCR 5-4	CAGAGCTGAGAAGACAGCAGAGAGCTACTCCCGAAGCACAAG ATG CCCTCTGGGCTTCCGTGACCTTGGC
Cas9 гидРНК -CCR 5-5	CTGACAATACTTGAGATTTTCAGATGTCACCAACGACCAAGAG AGC TTGATATGACTGTATATAGTATAG
Cas9 гидРНК -CCR 5-6	CAGATACATAGGTGGCAGGATTCTTCACTCCAGACTTAATCTAT CT GAAGCTATGAGAAATTTTCCACAT
Cas9 гидРНК -CCR 5-7	TATATGATTGATTTGCACAGCTCATCTGGCCAGATAAGCTGAGA CA TCCGTTCCCCTACAAGAAACTCTC
Cas9 гидРНК -CCR 5-8	ATCTGGCCAGAAGAGCTGAGACATCCGTTCCCCTTGAAGAAAC TCTCCCCGGTAAGTAACCTCTCAGCTG
Cas9 гидРНК -CCR 5-9	AGGCATCTCACTGGAGAGGGTTTAGTTCTCCTTAAGAGAAGATA AGATTTCAAGAGGGAAGCTAAGACTC
Cas9 гидРНК -CCR 5-10	ATAATATAATAAAAAATGTTTCGTCTGCCACCACTAATGAATGT CATGCATTCTGGGTAGTTTAGTCTTA
Cas9 гидРНК -CCR 5-11	TTTATAAAGTCCTAGAATGTATTTAGTTGCCCTCGTTGAATGCA AACTGTTTTATACATCAATAGGTTTT
Cas9 гидРНК -CCR 5-12	GCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCCCACTGTCCC CTTCTGGGCTCACTATGCTGCCGCC
Cas9 гидРНК -CCR 5-13	TTTTAAAGCAAACACAGCATGGACGACAGCCAGGCTCCTATCG ATTGTCAGGAGGATGATGAAGAAGATT

	Cas9 гидРНК -CCR 5-14	GCTTGTCATGGTCATCTGCTACTCGGGAATCCTAATTACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGG
	Cas9 гидРНК -CCR 5-15	ATACTGCCCCCGCGAGGCCACATTGGCAAACCAGCTTGGGTGTTCAACTTCCAGACTTGGCCATGGAGAA
Доно-рный ssO DN для направ лен-ного воздейс твия RETAL ENs на CC R5	reTALE N-CCR 5-1	CTGAAGAATTTCCCATGGGTCCCCACTTTCTTGTGAATCCTTGGAGTGAACCCCCCTGTCCTGAGTGGTTACTAGAACACACCTCTGGAC
	reTALE N-CCR 5-2	TGGAAGTATCTTGCCGAGGTCACACAGCAAGTCAGCAGCACAGCCAGTGTGACTCCGAGCCTGCTCCGCTAGCCCACATTGCCCTCTGGG
	reTALE N-CCR 5-3	CTACTGTCATTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATAACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
	reTALE N-CCR 5-4	GGAAGCCCAGAGGGCATCTTGTGGCTCGGGAGTAGCTCTCTGCTACCTTCTCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACC
	reTALE N-CCR 5-5	TCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACCAACGCCCAAGAGAGCTTGATATGACTGTATATAGTATAGTCATAAAGAAC
	reTALE N-CCR 5-6	GTGGAAAATTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAGCAATCCTGCCACCTATGTATCTGGCATAGTGTGAGTCCTCATAAA
	reTALE N-CCR 5-7	GAAACAGCATTTCCTACTTTTATACTGTCTATATGATTGATTTGGTCAGCTCATCTGGCCAGAAGAGCTGAGACATCCGTTCCCCTACAA
	reTALE N-CCR 5-8	TTGATTTGCACAGCTCATCTGGCCAGAAGAGCTGAGACATCCGTATCCCTACAAGAAACTCTCCCCGGTAAGTAACCTCTCAGCTGCTTG
	reTALE N-CCR 5-9	GGAGAGGGTTTAGTTCTCCTTAGCAGAAGATAAGATTTCAAGATGAGAGCTAAGACTCATCTCTCTGCAAATCTTTCTTTTGAGAGGTAA
	reTALE N-CCR 5-10	TAATATAATAAAAAATGTTTCGTCTGCCACCACAGATGAATGTCGAGCATTCTGGGTAGTTTAGTCTTATAACCAGCTGTCTTGCCTAGT

reTALEN-CCR5-11	TTAAAAACCTATTGATGTATAAAACAGTTTGCATTCATGGAGGGTGACTAAATACATTCTAGGACTTTATAAAAGATCACTTTTTATTTA
reTALEN-CCR5-12	GACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTATTTACTGTCCCCTTCTGGGCTCACTATGCTGCCGCCCAGTGGGAC
reTALEN-CCR5-13	TCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTACGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAA
reTALEN-CCR5-14	GGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAGACCTAAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACA
reTALEN-CCR5-15	GGCAAGCCTTGGGTCATACTGCCCCCGCGAGGCCACATTGGCAAGTCAGCAAGGGTGTTCAACTTCCAGACTTGGCCATGGAGAAGACAT

Пример VII

Получение библиотеки ампликонов целевых участков

Собирали клетки через 6 дней после нуклеофекции 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 направленного воздействия в CCR5	название	последовательность праймера
	Site1-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATTTT GCAGTGTGCGTTACTCC
	Site1-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGTTT GCAGTGTGCGTTACTCC
1	Site1-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAATTT GCAGTGTGCGTTACTCC
	Site1-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCATTT GCAGTGTGCGTTACTCC CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCAAGCAACTAAGTCACAGCA
	Site1-R	
	Site2-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATATG AGGAAATGGAAGCTTG
	Site2-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAT GAGGAAATGGAAGCTTG

2	Site2-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAT GAGGAAATGGAAGCTTG
	Site2-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAATG AGGAAATGGAAGCTTG
	Site2-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCATTAGGG TATTGGAGGA
	site3-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATAAT CCTCCCAACAACTCAT
	site3-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAA TCCTCCCAACAACTCAT
3	site3-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAA TCCTCCCAACAACTCAT
	site3-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAAAT CCTCCCAACAACTCAT
	site3_R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCCCAATCCT ACAGAGGCAG
	site4-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATAA
	site4-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAA
4	site4-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAA
	site4-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAAA
	site4_R	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGCCAAA GCTTTTTATTCT
5	Site5-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATATC TTGTGGCTCGGGAGTAG
	Site5-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGATC TTGTGGCTCGGGAGTAG
	Site5-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTTGGCAGGA TTCTTCACTCCA
	site6-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCTA TTTTGTTGCCCTTCAAA
6	site6-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCTA TTTTGTTGCCCTTCAAA
	site6-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAACCTGAA CTTGACCATATACT
	site7-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCA GCTGAGAGGTTACTTACC
7	site7-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCA GCTGAGAGGTTACTTACC
	site7-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAATGATTTA ACTCCACCCTC
	site8-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATACT CCACCCTCCTTCAAAAGA
8	site8-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGACT CCACCCTCCTTCAAAAGA
	site8-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTTGGTGTTTG CCAAATGTCT
	site9_F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATGG GCACATATTCAGAAGGCA
9	site9_F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGGG GCACATATTCAGAAGGCA
	site9_R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAGTGAAAG ACTTTAAAGGGAGCA
	site10-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCAC AATTAAGAGTTGTCATA
10	site10-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCA CAATTAAGAGTTGTCATA
	site10-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTCTCAGCTA GAGCAGCTGAAC
11	site11-Fl	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTGACACTTG ATAATCCATC
	site11-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGTCA
		ATGTAGACATCTATGTAG
	site11-R	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATTCA ATGTAGACATCTATGTAG
	site12-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATACT GCAAAAGGCTGAAGAGC
	site12-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGACT GCAAAAGGCTGAAGAGC
12	site12-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAACT GCAAAAGGCTGAAGAGC
	site12-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAACT GCAAAAGGCTGAAGAGC
	site12-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTGCCTATAA AATAGAGCCCTGTCAA
	site13-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCTC TATTTTATAGGCTTCTTC
13	site13-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCTC TATTTTATAGGCTTCTTC
	site13-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAGCCACCA CCCAAGTGATC
	site14-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGTTC CAGACATTAAAGATAGTC
14	site14-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATTTC CAGACATTAAAGATAGTC
	site14-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAATCATGA TGGTGAAGATAAG
	site15-Fl	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCCG GCAGAGACAAACATTAAA
15	site15-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCGGCAGA GACAAACATTAAA
	site15-R	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCTAGCTAGGA AGCCATGGCAAG
адаптер Illumina	PE-PCR-F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACA cgac*g*c
	PE-PCR-R	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCT GCTGAACc*g*c
Мультиплексные сиквенсные праймеры для ПЦР		
3	site3-MF	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTGCATA GTATGTGCTAGATGCTG
	site3-MR	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGATCTC TAAGAAGGCAAATGAGAC
Illumina адаптер	Index-PCR	CAAGCAGAAGACGGCATACGAGATN1N2N3N4N5N6GTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT

	universal-PCR	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACA CGACGCTCTTCCGATCT

* 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

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

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

Пример IX

Скрининг Генотипа колонизированных hiPSCs

Человеческие клетки iPS на безфидерных культурах предварительно обрабатывали средой mTesr-1, обогащенной SMC4 (5 мкМ тиазовивин, 1 мкМ CHIR99021, 0,4 мкМ 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 mTesr-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 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). 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.

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

Иммуннноокрашивание и анализы тератом, полученных из hiPSCs

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 на Axio Observer Z.1 (ZIESS).

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

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 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 в полном объеме. Соответственно, один аспект настоящего изобретения относится к TALEs, лишенных повторов, например, полностью лишенных повторов. Такие перекодированные TALE является эффективными, поскольку позволяют быстрее и проще синтезировать расширенные массивы TALE RVD.

Для устранения повторов, нуклеотидные последовательности массивов TALE RVD эволюционировали вычислительным образом с тем, чтобы минимизировать количество последовательностей-повторов при сохранении аминокислотной композиции. Перекодированные TALE (Re-TALEs), кодирующие 16 тандемных мономеров RVD, распознающих ДНК, плюс конечный полуповтор RVD, лишены любых 12 п.о. повторов (см. Фигуру 5a). Примечательно, что этот уровень перекодирования достаточен для того, чтобы провести амплификацию ПЦР какого-либо конкретного мономера или подраздела из полноразмерного конструкции с re-TALE (см. Фиг. 5b). 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 с обогащенными повтрами 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 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 (PI и P2). С помощью нулевой гипотезы о том, что изменения 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 наблюдаемой эффективности HDR. Для анализа чувствительности HDR, донорные ssODN содержали 2 п.о. несовпадение относительно генома-мишени, что делает вероятной одновременную замену оснований в двух п. о., являющихся мишенями 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 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 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 время (см. Фиг. 8A). 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	тип клеток	HDR (ReTALEN) (%)	NHEJ (reTALE) (%)	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/hgl9/bbi/wgEncodeOpenChromDnaseIpsnihi7Sig. 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 перекодироваными 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. 1a. 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 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. 1b). Активность reTALENs была протестирована в локусе AAVS1 в PGP1 hiPSCs (см. Фиг. 1c), при этом успешно восстановленные клеточные клоны, содержащие специфические вставки (см. Фиг. 1d, e), подтвердили, что reTALENs активны как в соматических, так и в плюрипотентных клетках человека.

Устранение повторов позволяет получить функциональный лентивирирус, 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, инфецированных lenti-reTALE-2A-GFP. Клетки 293T, трансдуцированные lenti-re-TALE-TF показали 36x активацию экспрессии репортера по сравнению с только отрицательным репортером (см. Фиг. 7a,b,c). Проверяли целостность последовательности re-TALE-TF в лентивирус-инфицированных клетках и обнаружили полноразмерные reTALEs во всех 10 протестированных клонах (см. Фиг. 1d).

Пример XII

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

Для сравнения эффективности редактирования re-TALENs относительно Cas9-гидРНК 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, 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. 2a). Нуклеазные конструкции и донорный ssODN трансфецировали в hiPSCs. Для количественной оценки эффективности генного 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, в то время как доставка re-TALENs и ssODN давала эффективности 1,7% по HDR и 1,2% по NHEJ (см. Фиг. 2b). 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 Associated Motif (PAM) участка направленного воздействия Cas9-гидРНК (см. Фиг. 2b), 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 является одновременно высоко воспроизводимым (коэффициент вариации между повторами = ± 15% * измеренная эффективность) и в 400X более чувствительным, чем наиболее часто используемые тесты с чувствительной к несовпадениям эндонуклеазой (см. Фиг. 8).

Пары re-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 трансфецировали вместе с соответствующими донорными ssODNs (см. таблицу 3) в PGP1 hiPSCs. Через шесть дней после трансфекции профилировали эффективности редактирования генома в этих участках (таблица 6). Для 13 из 15 пар re-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 (R2 = 0,81) обнаружили между эффективностями HR и NHEJ в одних и тех же целевых локусах (P <1 x 10-4) (см. Фиг. 9a), что предполагает, что образование DSB, общей вышестоящей стадии как для HDR, так и для NHEJ, является скорость-лимитирующей стадией для reTALEN-опосредованного редактирования генома.

В противоположность этому, все 15 пар Cas9-гидРНК показали значительные уровни 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) (R2 = 0,52, p = 0,003), что согласуется с наблюдениями для reTALENs. Эффективность NHEJ, достигаемая Cas9-гидРНК была значительно выше, чем та, что достигается с помощью reTALENs (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) (R2 = 0,28, p = 0,04), что позволяет 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 NHEJ в среднем в 7 раз выше, чем соответствующие reTALEN, Cas9-гидРНК достигает уровней HDR (среднее = 1,0%), схожих с уровнями соответствующих reTALENs (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 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

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 спейсерной области участков направленного воздействия пары re-TALEN #3 для CCR5. Наблюдаемая 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. 3a). Поскольку более длинные участки гомологии улучшают показатели HDR, когда донорные дцДНК 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 дцДНК-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. 3a), 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 re-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 спейсерной области пары re-TALEN # 3. Каждый ssODN также содержала второе 2 п.о. несовпадение (B) на различных расстояниях от центра (см. Фиг. 3b). ssODN, обладающий лишь одним 2 п.о. несовпадением использовали в качестве контроля. Каждый из этих ssODNs вводили в индивидуальном порядке с парой re-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. 3b, см. Фиг. 11a). 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. 3b (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. 11b). Оцененное распределение подразумевает, что тракты генной конверсии постепенно становятся менее частыми, по мере увеличения их длин, результат, очень похожий на распределения трактов генной конверсии, наблюдаемых с донорными дцДНК, но на сильносжатой 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 2 п.о. несовпадения «A», позволил получить паттерн, в котором A отдельно был включен в 85% случаев, с множественными несовпадениями B, включенными в других случаях (см. Фиг. 11c). Хотя количество событий с включением только B 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). Наконец, во всех экспериментах с одиночными несовпадениями B, наблюдалась небольшая часть событий включения только B (0,04% ~ 0,12%), что приблизительно постоянно для всех расстояний от A до B.

Кроме того, был проведен анализ того, как далеко может быть расположен донор ssODN от re-TALEN-индуцированного разрыва двухцепочечной ДНК, при сохранении 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 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 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 воздействующий на локус AAVS 1, и ssODN-доноры различных ориентаций (Oc: комплементарная к гидРНК и On: некомплементарная к гидРНК) и длин (30, 50, 70, 90, 110 нуклеотидов). Oc были более эффективны, чем On, а 70-мерный Oc достиг оптимального показателя HDR 1,5% (см. Фиг. 3d.) Такую же тенденциозность цепи ssODN детектировали с помощью никазы из Cas9 (CC: Cas9_D10A), несмотря на тот факт, что эффективности HDR, опосредованные Cc с ssODN были значительно меньше, 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 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 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. 4a), PGP1 hiPSCs трансфецировали парой reTALENs и ssODN, предназначенных для направленного воздействия на CCR5 в участке #3 (таблица 3). GEAS проводили на части трансфецированных клеток, при этом обнаружили частоту HDR 1,7% (см. Фиг. 4b). 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 $\mu = 0{,}017 * 0{,}25 * 96 * 5 * 2$). 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. 4c). Эффективность направленного воздействия 1% (1% = 2/2 * 100, 2 моноаллельных скорректированных клона из 100 клеток, подвергнутых скринингу) была сопоставима с данными, полученными при анализе секвенированием следующего поколения (1,7%) (см. Фиг. 4b). Плюрипотентность полученных hiPSCs подтвердили иммуногистохимическое окрашивание SSEA4 и TRA-1-60 (см. Фиг. 4d). Подвергнутые успешному направленному воздействию клоны hiPSCs были способны образовывать зрелые тератомы с признаками всех трех зародышевых листков (см. Фиг. 4E).

Пример XV

Способ непрерывного редактирования генома клетки

Согласно некоторым аспектам, предлагается способ редактирования генома в клетках, включающий человеческую клетку, например, человеческую стволовую клетку, где клетка генетически модифицирована для включения нуклеиновой кислоты, кодирующей фермент, который образует комплекс колокализации с РНК, комплементарной ДНК-мишени, и который расщепляет ДНК-мишень сайт-специфическим образом. Такой фермент включает направляемый РНК ДНК-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 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 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 нуклеиновой кислоты, кодирующей 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 реактива, 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 редактировать геном HiPSCs без «шрамов» с эффективностями вплоть до 50% при однодневных обработках, которые в 2-10 раз более эффективны, чем наилучшие эффективности, известные к настоящему времени. Кроме того, способ прост в использовании и имеет значительно меньшую клеточную токсичность. Воплощения настоящего изобретения включают одиночное 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 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 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 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 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 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, обработанных фосфатазой, для мультиплексного редактирования генома. Аспекты настоящего изобретения включают получение библиотеки клеточных линий PGP, 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-индуцируемым 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 плазмиды 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, кодирующей 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 удаления вектора известны специалистам в данной области техники. Проводили скрининг на колонизированных iPS-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 перманетно вставленной кассеты с Cas9.

Пример XIX

Редактирование генома в клетках iPGP1

Исследование патогенеза кардиомиопатии исторически было затруднено из-за отсутствия подходящих модельных систем. Дифференцировка кардиомиоцитов из индуцированных плюрипотентных стволовых клеток, полученных из пациентов (iPSCs) 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, 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. Однораундная трансфекция РНК дает ~30% эффективность HDR. Модифицированные клетки Cas9-iPGP1 с искомыми мутациями выращивали до уровня 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, наблюдаемые в iPSCs, которые были получены из пациентов, и в неонатальной крысиной модели с нокаутом TAZ (см. Фиг. 19). Соответственно, предлагаются способы коррекции мутаций, вызывающих заболевания, в плюрипотентных клетках с последующей дифференцировкой клеток в искомые типы клеток.

Пример XX

Материалы и методы

<u>1. Получение стабильных клеточных линий iPS/ES с PiggyBacCas9, индуцируемым DOX</u>

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 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 mTeSRl; осторожно промывали клетки 2 мл DPBS на лунку шестилуночного планшета.

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 mTeSRl. Добавляли 1 мл mTeSRl и сбивали клетки с помощью 1000 мкл микропипетки медленным потоком mTeSRl.

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.

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.

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

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

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 mTeSRl и ингибитором ROCK. Инкубировали клетки при 37°C в течение ночи.

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.

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

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

# 1 Смесь нативной РНК	[Конечная] (мМ)
GTP	7,5
АТФ	7,5
CTP	7,5
UTP	7,5
Общий объем	
# 2 Смесь Кэпированной нативной РНК	[Конечная] (мМ)
структурный аналог 3'-O-Me-m7G Cap (NEB)	6
GTP	1,5
ATP	7,5
CTP	7,5
UTP	7,5
Общий объем	
#3 Модифицированная смесь РНК	[Конечная] (мМ)
GTP	7,5
ATP	7,5
5-Me-CTP (Tri-Link)	7,5

Pseudo-UTP (Tri-Link)	7,5
Общий объем	
#4 смесь кэпированной/модифицированной РНК	[Конечная] (мМ)
структурный аналог 3'-O-Me-m7G Cap (NEB)	6
GTP	1,5
ATP	7,5
5-Me-CTP (Tri-Link)	7,5
Псевдо-UTP (Tri-Link)	7,5
Общий объем	

4. Готовили смесь для транскипции 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. Инкубировали в течение 4 часов (3-6 часов OK) при 37 °C (амплификатор).

6. Добавляли 2мкл Turbo ДНКазы (набор MEGAscript из Ambion) к каждому образцу. Тщательно перемешивали и инкубировали при 37 °C в течение 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).

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

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

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.

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

1. 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.

2. Аккуратно аспирировали версен и добавляли mTeSRl. Добавляли 1 мл mTeSRl и отсоединяли клетки с помощью 1000 мкл микропипетки медленным потоком mTeSRl.

3. Собирали отсоединенные клетки, мягко разбивали их в одноклеточную суспензию,

количественно оценивали с помощью гематоцитометра, и корректировали плотность клеток до 100 тыс. клеток на мл.

4. Рассевали клетки на покрытые матригелем 10 см чашки с mTeSRl плюс ингибитор ROCK при плотности клеток 50К, 100К и 400К на 10 см чашку.

5. Скрининг клонов, образованных из одиночных клеток

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 10 мкл) с наконечниками с фильтрами аспирировали 10 мкл среды из каждой лунки 24 луночных планшетов. 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 достаточно большими для разделения.

4. Аспририровали 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.

7. Добавляли 250 мкл mTeSRl. 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 Эппендорф для экстракции геномной ДНК.

<u>6. Скрининг клонов</u>

1. Центрифугировали пробирку из стадии 7.7

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

3. NaCl + 10% ДСН, 40 мкг/мл протеиназы К + (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 70% этанолом.

7. Аккуратно удаляли этанол. 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.

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

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

<u>7. Удаление вектора PiggyBac</u>

1. Повторяли стадии 2.1-2.9

2. Переносили клетки в кювету Nucleofector с помощью 1 мл наконечника для пипетки. Добавляли 2 мкг плазмиды с транспозазой в суспензию клеток в кювете. Смешивали клетки и ДНК осторожным покачиванием.

3. Повторяли стадию 2.10-2.11

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.

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. 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/gksl446.

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., 0'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, 0., 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. andKim, 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.

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.

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.

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.

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, el52.

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.

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.

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.

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.

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.

48. Mimitou, E. P. and Symington, L. S. (2008) Sae2, Exol and Sgsl 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.

(57) Формула изобретения

1. Способ изменения ДНК-мишени в клетке, экспрессирующей фермент Cas, который формирует комплекс колокализации с направляющей РНК, комплементарной ДНК-мишени, и который расщепляет ДНК-мишень сайт-специфическим образом, включающий

(a) введение в клетку последовательности донорной нуклеиновой кислоты,

(b) введение в клетку направляющей РНК, комплементарной ДНК-мишени, где направляющая РНК и фермент Cas являются членами комплекса колокализации на ДНК-мишени,

где направляющая РНК и фермент Cas колокализуются на ДНК-мишени, фермент Cas расщепляет ДНК-мишень и донорная нуклеиновая кислота встраивается в ДНК-мишень с получением измененной ДНК в клетке, и

повтор стадий (a) и (b), результатом чего являются множественные инсерции экзогенных нуклеиновых кислот в клетке.

2. Способ по п. 1, в котором направляющая РНК включает структуру 5' Cap.

3. Способ по п. 1, где направляющая РНК утратила фосфатные группы.

4. Способ по п. 1, где фермент является Cas9.

5. Способ по п. 1, где длина направляющей РНК составляет от около 10 до около 500 нуклеотидов.

6. Способ по п. 1, где длина направляющей РНК составляет от около 20 до около 100 нуклеотидов.

7. Способ по п. 1, где направляющая РНК представляет собой химеру tracrRNA-crRNA.

8. Способ по п. 1, где ДНК представляет собой геномную ДНК, митохондриальную ДНК, вирусную ДНК или экзогенную ДНК.

9. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается в результате рекомбинации.

10. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается путем гомологичной рекомбинации.

11. Способ по п. 1, где донорная последовательность нуклеиновой кислоты встраивается негомологичным соединением концов.

12. Способ по п. 1, дополнительно включающий введение двух или более донорных нуклеиновых кислот и двух или более направляющих РНК, чтобы получить два или более изменения в ДНК клетки.

13. Способ по п. 1, где после получения измененной ДНК в клетке нуклеиновая кислота, кодирующая фермент, удаляется из генома клетки.

14. Способ по п. 1, где РНК и донорные нуклеотидные последовательности экспрессируются в виде связанной последовательности нуклеиновой кислоты.

15. Способ по п.1, в котором клетку генетически модифицируют путем введения ферментативно удаляемого вектора или кассеты, включающего нуклеиновую кислоту,

кодирующую этот фермент, в геномную ДНК клетки.

16. Способ по п.15, в котором:

i) вектор представляет собой вектор PiggyBac;

ii) ферментативно удаляемый вектор или кассету, включающий нуклеиновую кислоту, кодирующую фермент Cas, удаляют ферментом или

iii) ферментативно удаляемый вектор или кассету, включающий нуклеиновую кислоту, кодирующую фермент, удаляют с помощью транспозазы.

17. Способ по п.15, в котором:

i) нуклеиновая кислота, кодирующая фермент, является индуцибельной;

ii) нуклеиновая кислота, кодирующая фермент, находится под управлением промотора, индуцируемого доксициклином, или

iii) нуклеиновую кислоту, кодирующую фермент, индуцируют с помощью доксициклина.

18. Способ по п.1, в котором направляющая РНК вводится в клетку из среды, окружающей клетку.

19. Способ по п.1, в котором направляющую РНК обрабатывают фосфатазой, чтобы избежать иммунной реакции Toll2 в клетке.

20. Способ по п.1, в котором клетка является эукариотической клеткой.

21. Способ по п.1, в котором клетка представляет собой дрожжевую клетку, клетку растения или животную клетку.

22. Способ по п. 1, в котором клетка представляет собой соматическую клетку.

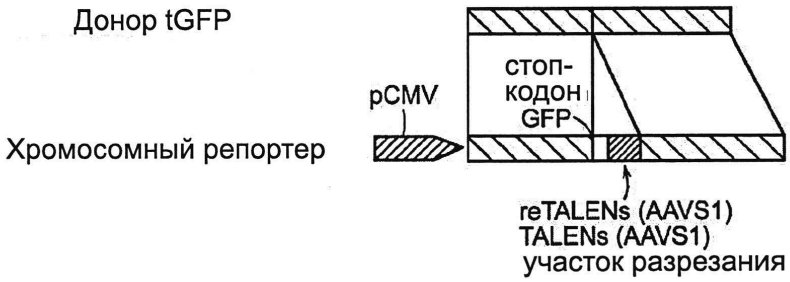
23. Способ по п. 1, в котором клетка представляет собой стволовую клетку.

24. Способ по п. 1, в котором клетка представляет собой человеческую стволовую клетку.

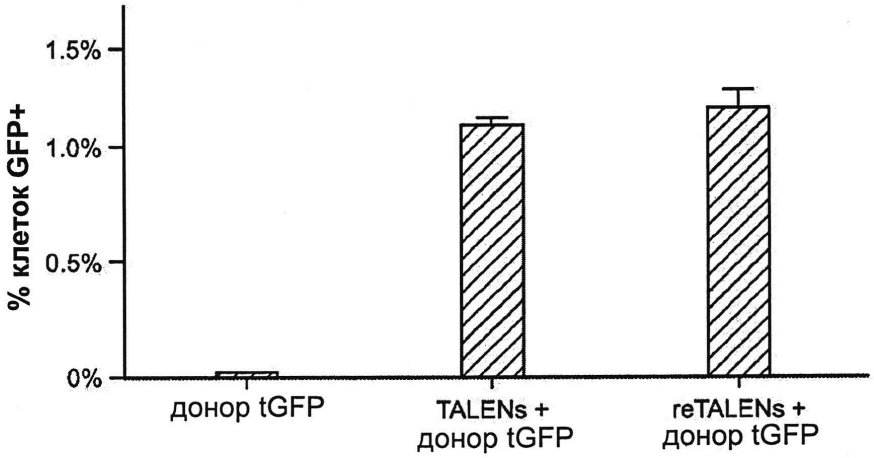
25. Способ по п.18, в котором направляющая РНК находится в комплексе с липосомой в окружающей клетку среде.

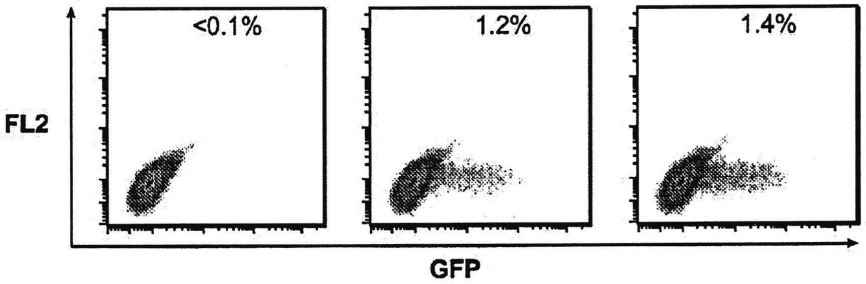
1

Фиг. 1A



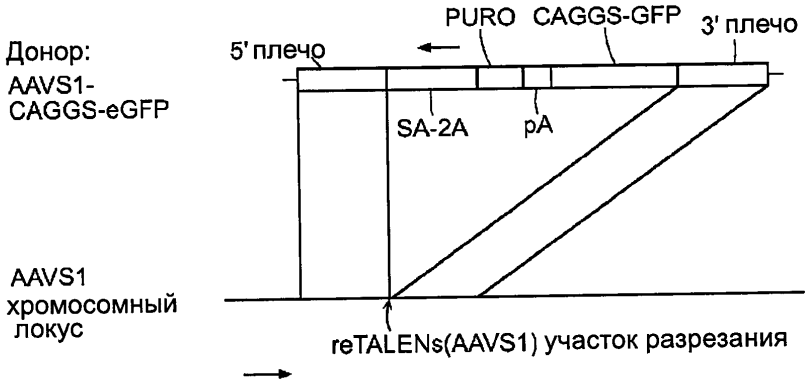
Фиг. 1B





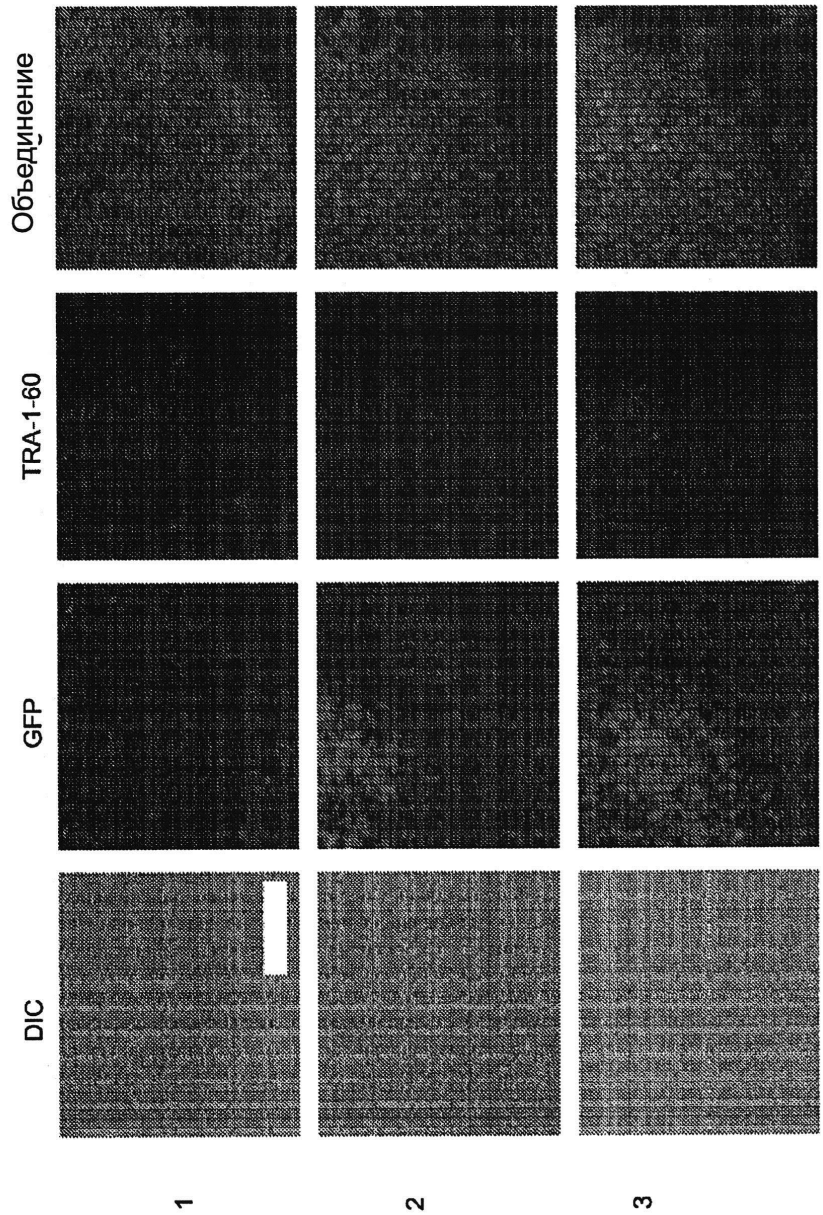
2

Фиг. 1C


Донор:
AAVS1-
CAGGS-eGFP
5' плечо
PURO
CAGGS-GFP
3' плечо
SA-2A
pA
AAVS1
хромосомный
локус
reTALENs(AAVS1) участок разрезания

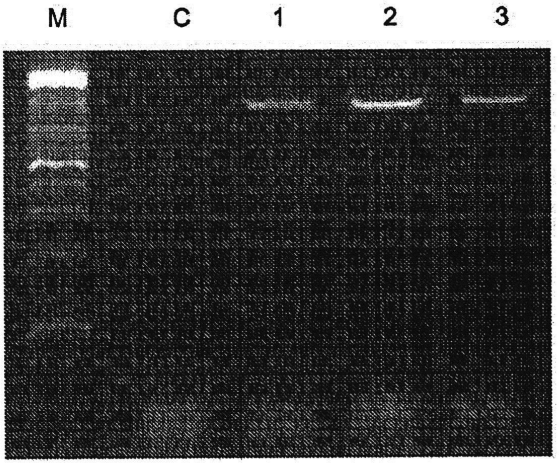
3

Фиг. 1D

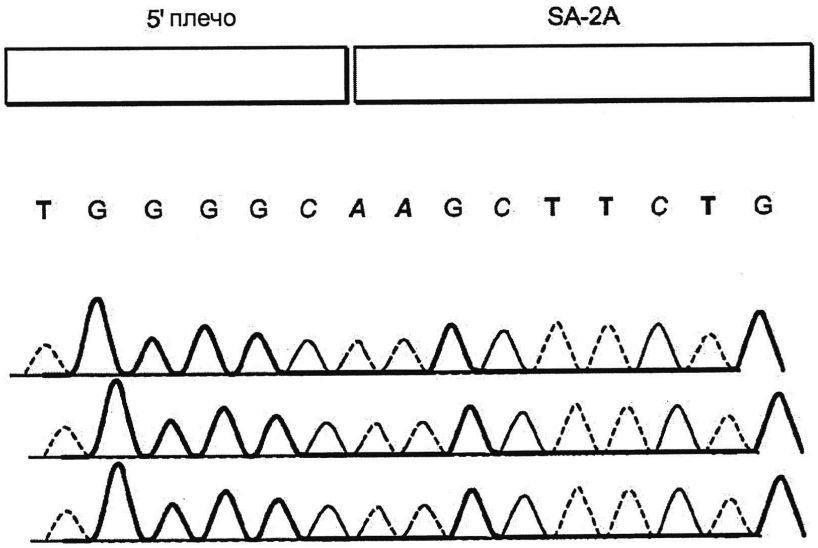


4

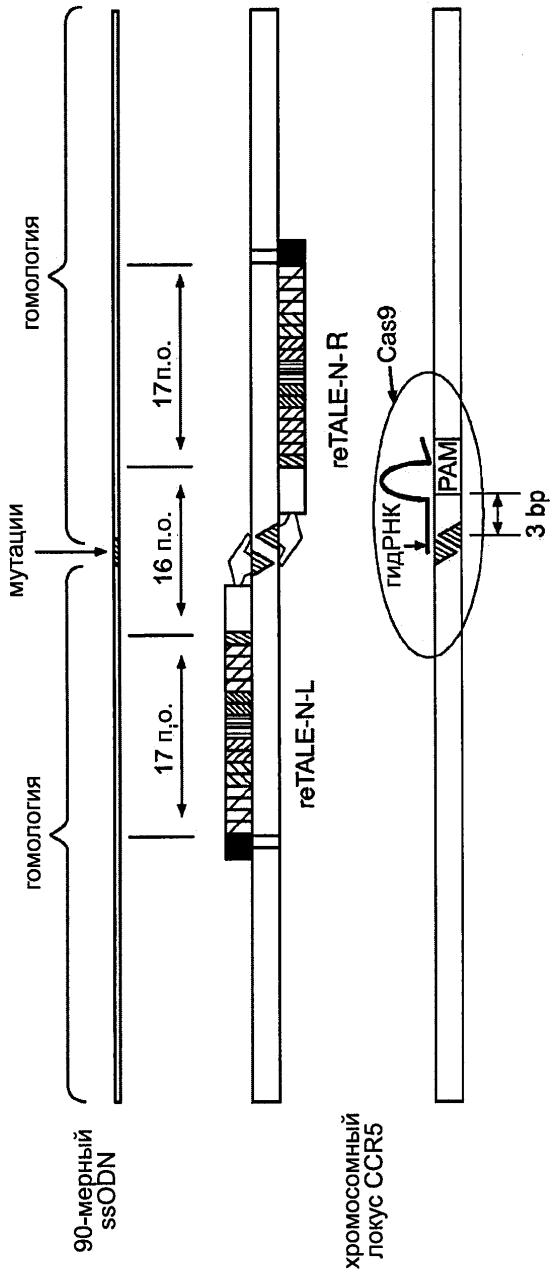
Фиг. 1E



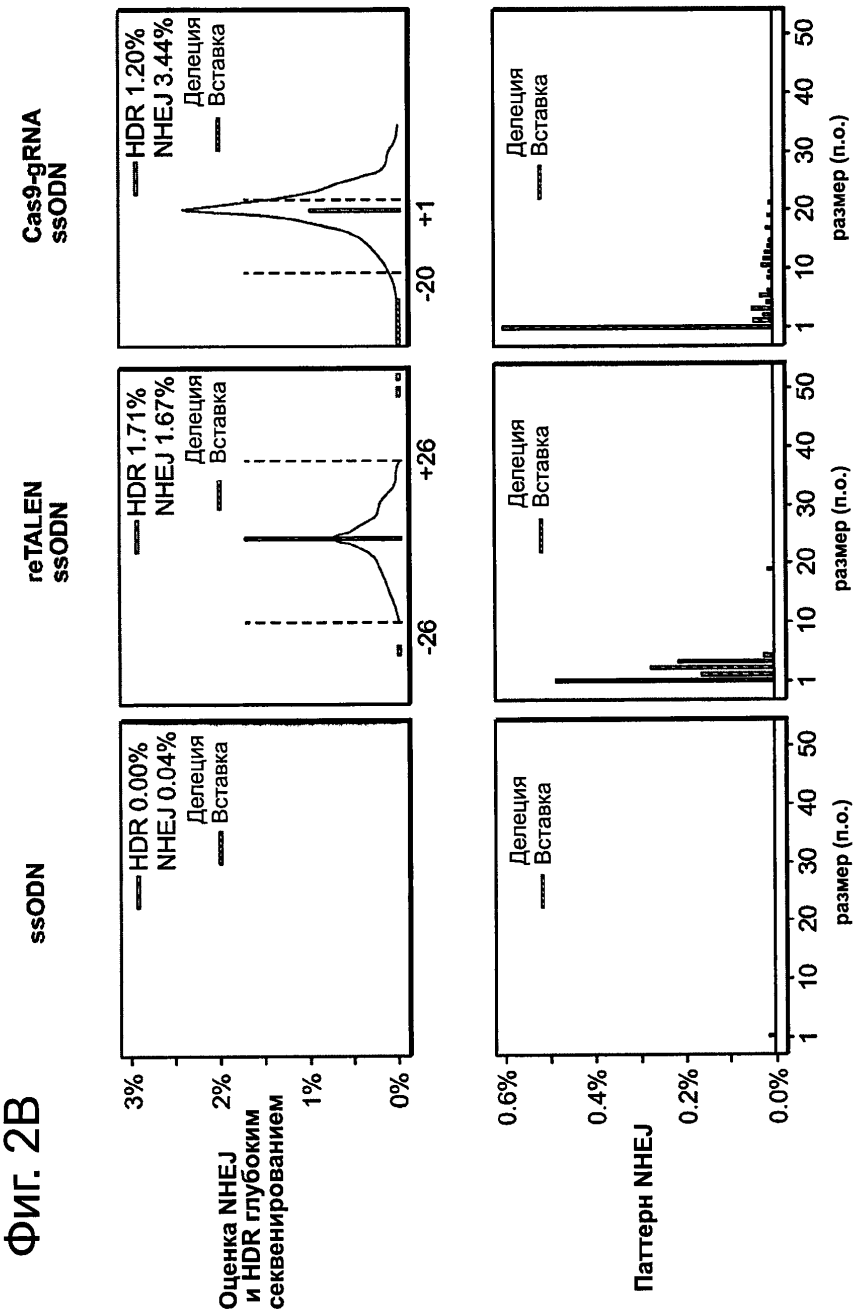
Фиг. 1F



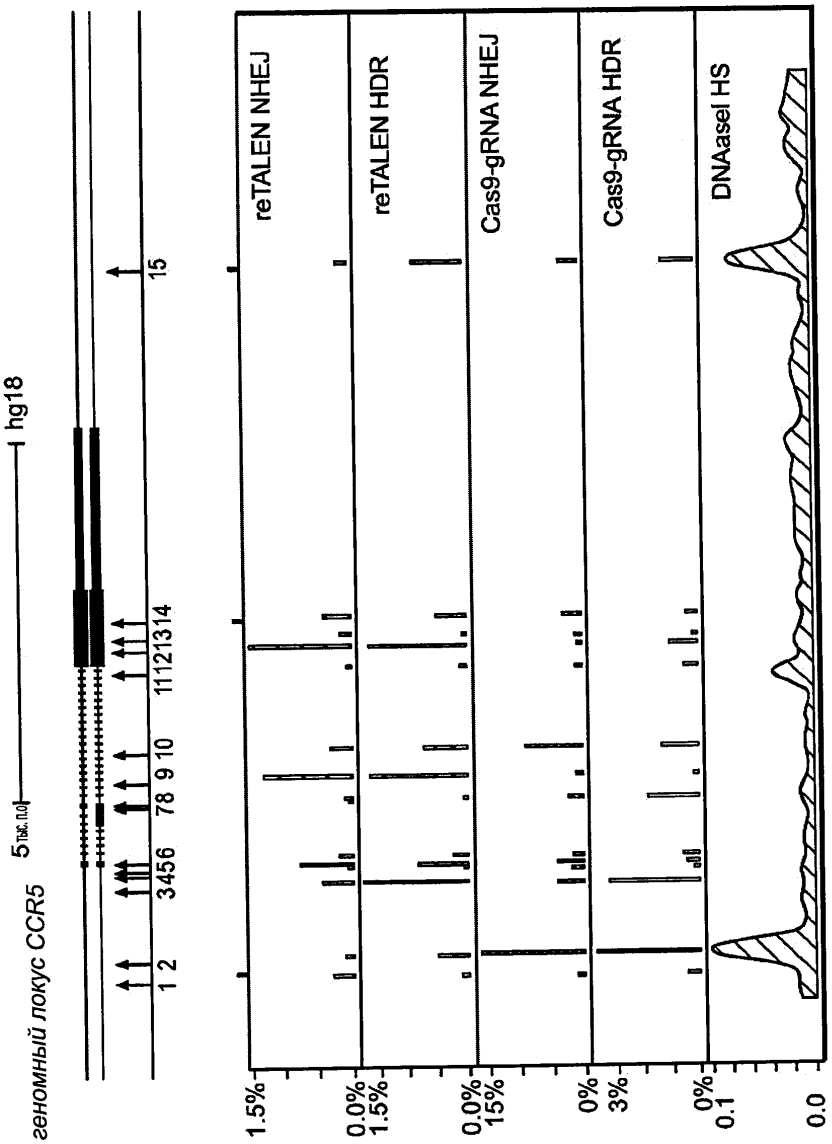
Фиг. 2A


90-мерный ssODN
гомология
мутации
гомология
17 п.о.
16 п.о.
17п.о.
хромосомный локус CCR5
reTALE-N-L
reTALE-N-R
Cas9
гидРНК
PAM
3 bp

Фиг. 2B

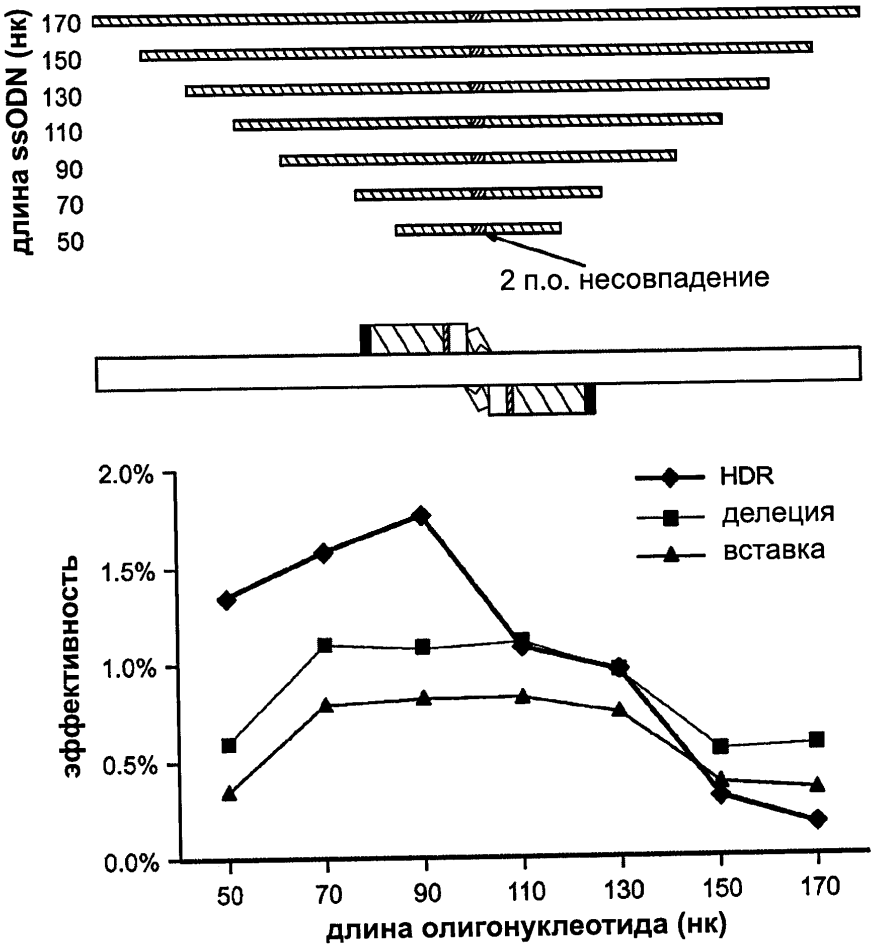

ssODN
reTALEN ssODN
Cas9-gRNA ssODN
HDR 0.00%
NHEJ 0.04%
HDR 1.71%
NHEJ 1.67%
HDR 1.20%
NHEJ 3.44%
Делеция
Вставка
-26
+26
-20
+1
Оценка NHEJ и HDR глубоким секвенированием
3%
2%
1%
0%
Паттерн NHEJ
0.6%
0.4%
0.2%
0.0%
1
10
20
30
40
50
размер (п.о.)

Фиг. 2C

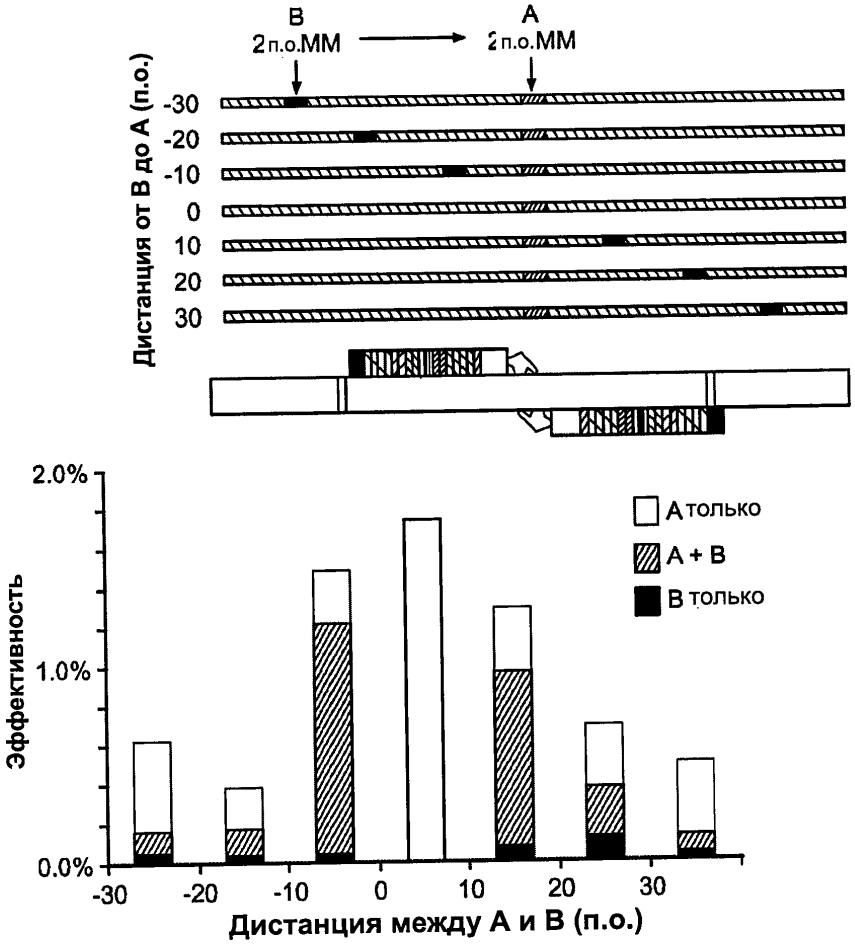

геномный локус CCR5
5тыс. п.о.
hg18
1 2
3456
78 9 10
1112 13 14
15
reTALEN NHEJ
reTALEN HDR
Cas9-gRNA NHEJ
Cas9-gRNA HDR
DNAasel HS
1.5%
0.0%
1.5%
0.0%
15%
0%
3%
0%
0.1
0.0

8

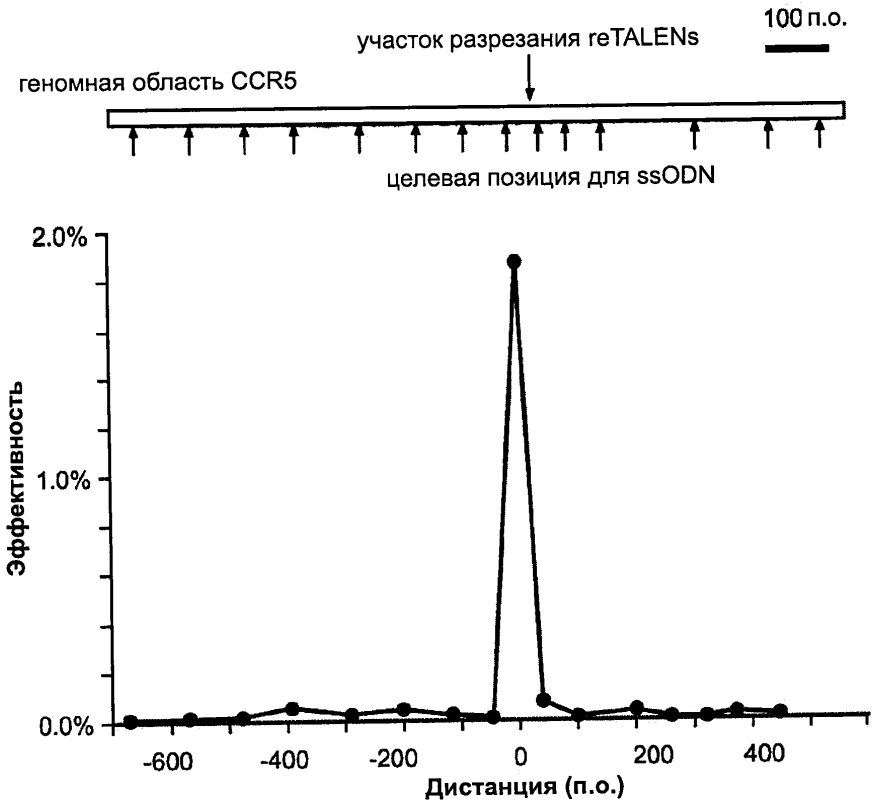
Фиг. 3А


длина ssODN (нк)
170
150
130
110
90
70
50
2 п.о. несовпадение
2.0%
1.5%
1.0%
0.5%
0.0%
эффективность
HDR
делеция
вставка
50
70
90
110
130
150
170
длина олигонуклеотида (нк)

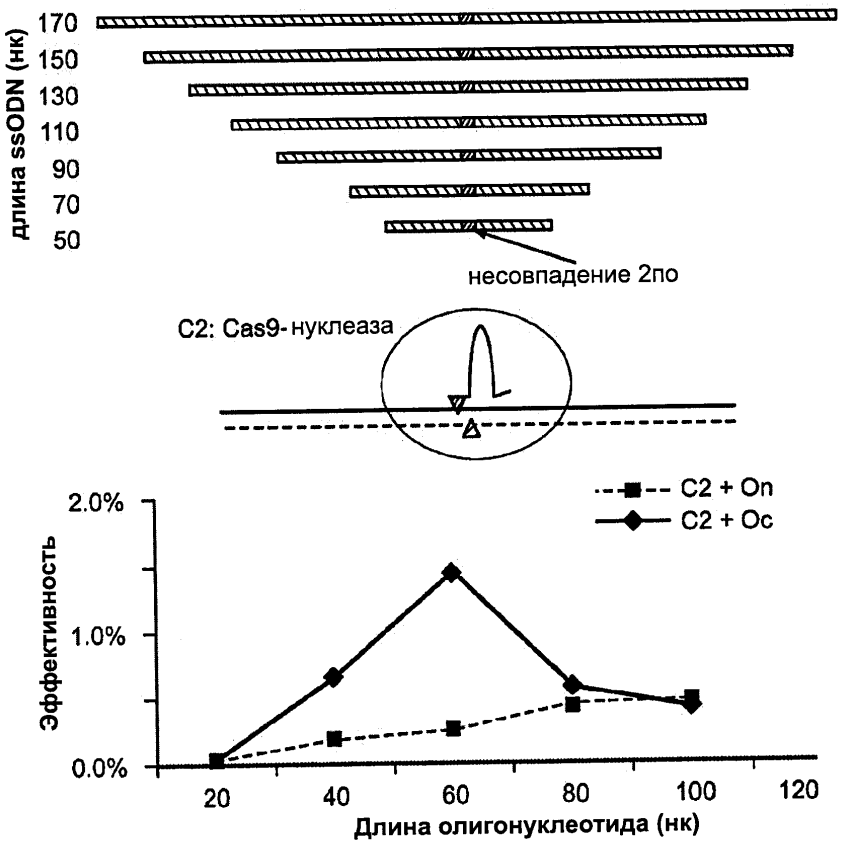
Фиг. 3B

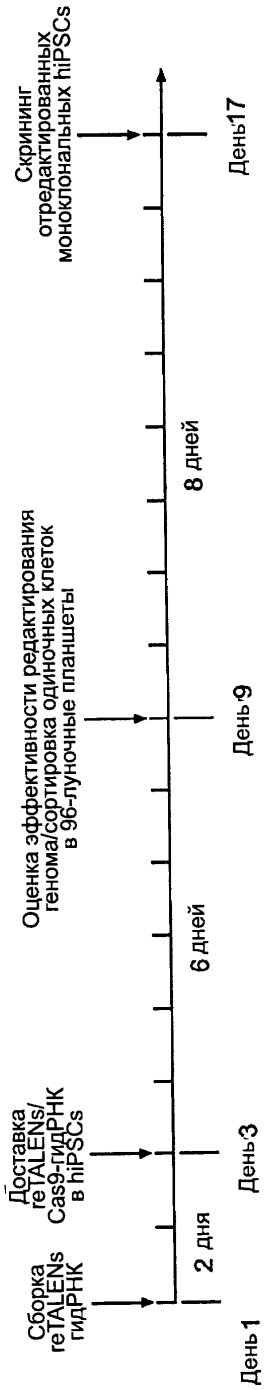

B
2 п.о.MM
A
2 п.о.MM
Дистанция от B до A (п.о.)
-30
-20
-10
0
10
20
30
2.0%
1.0%
0.0%
Эффективность
A только
A + B
B только
-30
-20
-10
0
10
20
30
Дистанция между A и B (п.о.)

Фиг. 3C

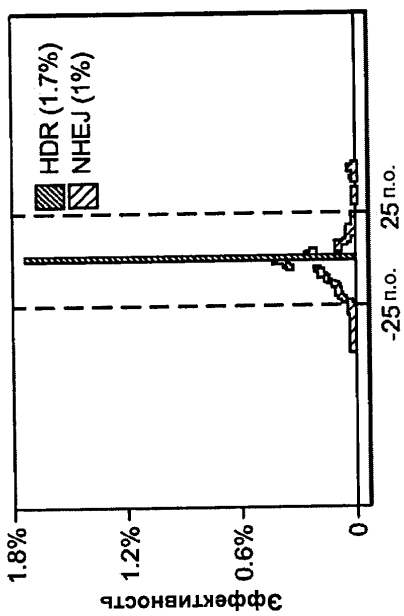

100 п.о.
участок разрезания reTALENs
геномная область CCR5
целевая позиция для ssODN
2.0%
1.0%
0.0%
Эффективность
-600
-400
-200
0
200
400
Дистанция (п.о.)

Фиг. 3D

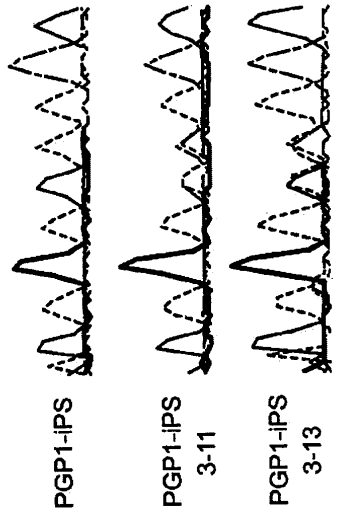

длина ssODN (нк)
170
150
130
110
90
70
50
несовпадение 2по
C2: Cas9-нуклеаза
C2 + On
C2 + Oc
Эффективность
2.0%
1.0%
0.0%
20
40
60
80
100
120
Длина олигонуклеотида (нк)



Фиг. 4A

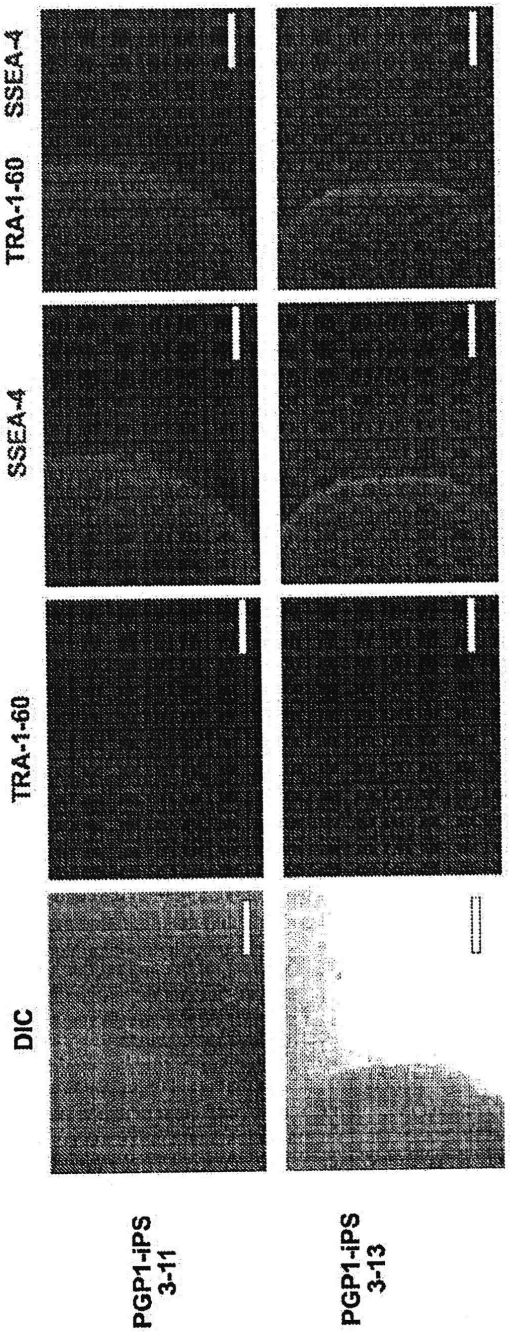


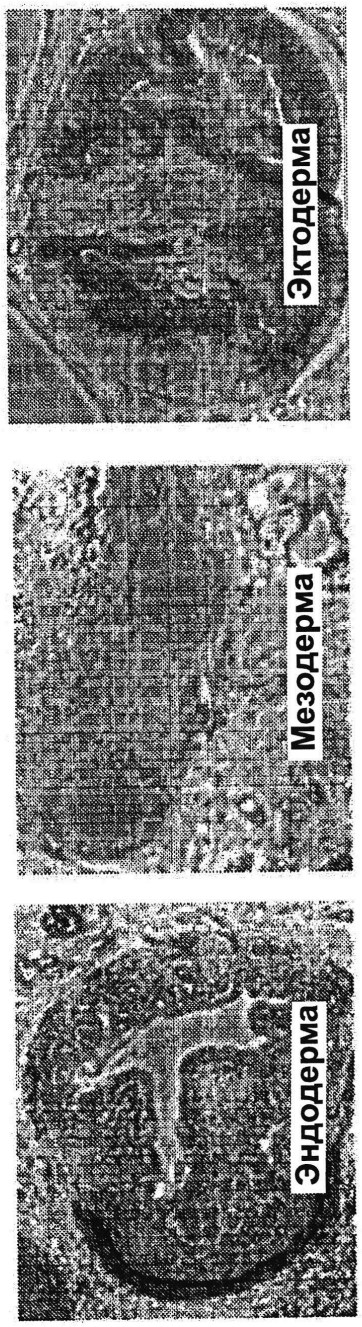
Фиг. 4B



Фиг. 4C

Фиг. 4D



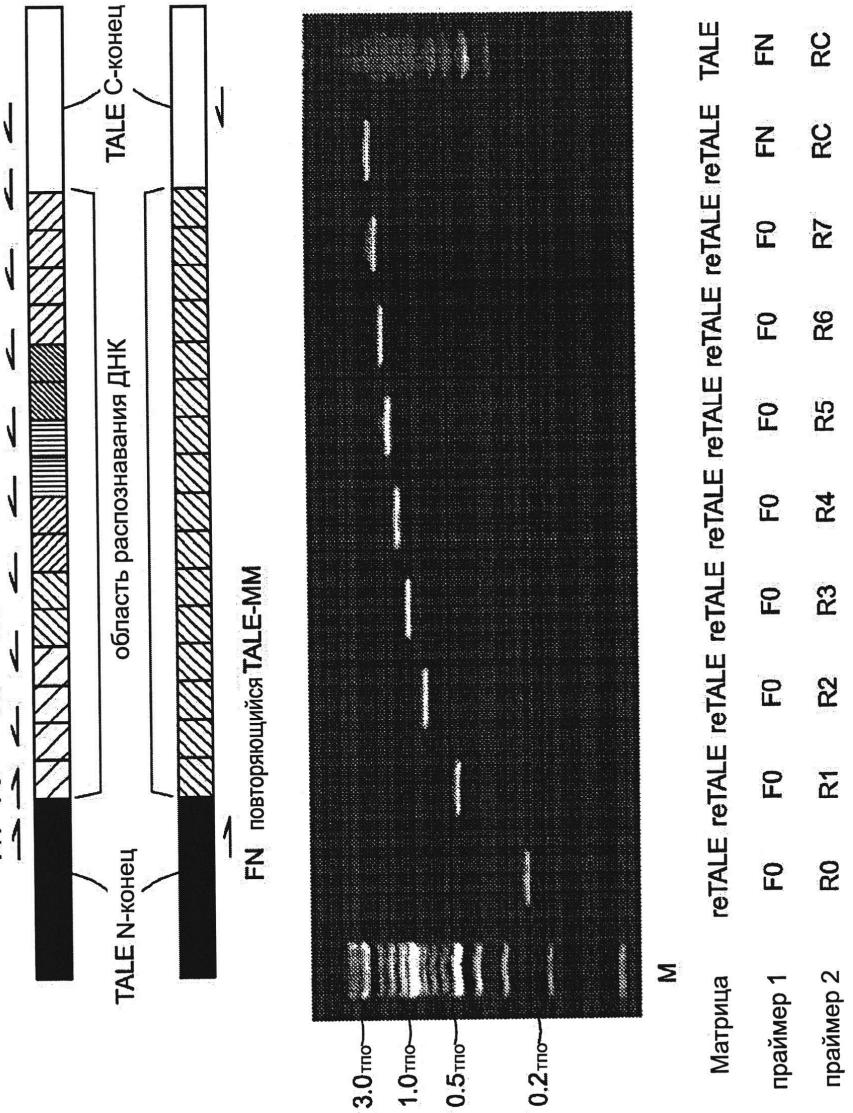


Фиг. 4E

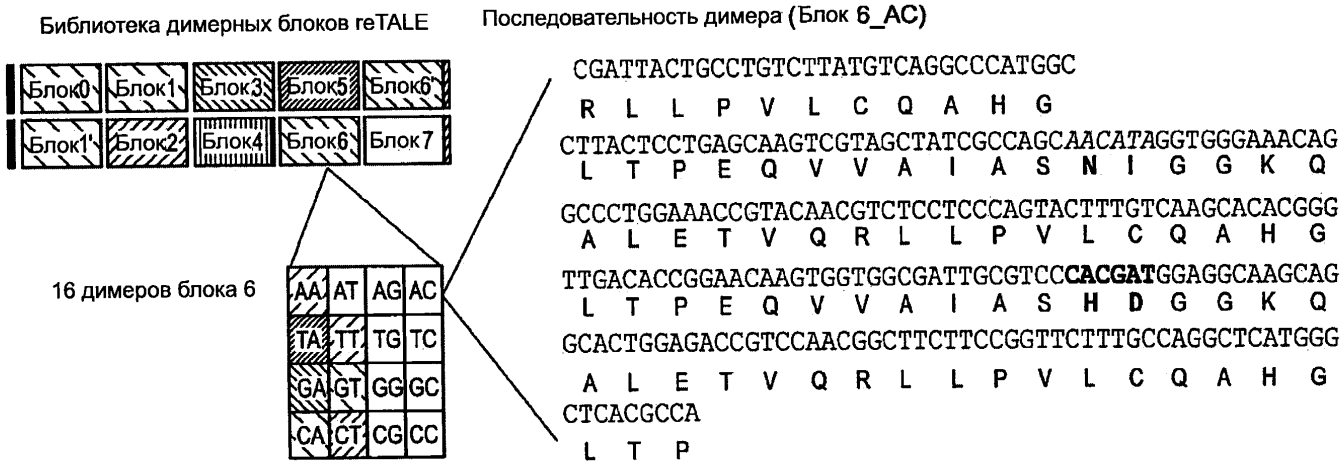
Фиг. 5A

```
                 10        20        30        40        50        60        70        80        90       100
                  L  T  P  E  Q  V  V  A  I  A  S  X  X  G  G  K  Q  A  L  E  T  V  Q  R  L  L  P  V  L  C  Q  A  H  G
TALE-MM          CTAACCCCAGAGCAGGTCGTGGCAATCGCCTCCNNNNNNGGCGGAAAACAGGCATTGGAAACAGTACAGCGGCTGCTGCCGGTGCTGTGCCAAGCGCACGGA
reTALE-M1        CTAACCCCTGAACAGGTAGTCGCTATAGCTTCANNNNNNGGGGGCAAGCAAGCACTTGAGACCGTTCAACGACTCCTGCCAGTGCTCTGCCAAGCCCATGGA
reTALE-M2        TTGACTCCGGAGCAAGTCGTCGCGATCGCGAGCNNNNNNGGGGGGAAGCAGGCGCTGGAAACTGTTCAGAGACTGCTGCCTGTACTTTGTCAGGCGCATGGT
reTALE-M3        CTCACCCCCGAACAGGTTGTCGCAATAGCAAGTNNNNNNGGCGGTAAGCAAGCCCTAGAGACTGTGCAACGCCTGCTCCCCGTGCTGTGTCAGGCTCACGGT
reTALE-M4        CTGACACCTGAACAAGTTGTCGCGATAGCCAGTNNNNNNGGGGGAAAACAAGCTCTAGAAACGGTTCAAAGGTTGTTGCCCGTTCTGTGCCAAGCACATGGG
reTALE-M5        TTAACACCCGAACAAGTAGTAGCGATAGCGTGANNNNNNGGGGGTAAACAGGCTTTGGAGACGGTACAGCGGTTATTGCCGGTCCTCTGCCAGGCCCACGGA
reTALE-M6        CTTACGCCAGAACAGGTGGTTGCAATTGCCTCCNNNNNNGGCGGGAAACAAGCGTTGGAAACTGTGCAGAGACTCCTTCCTGTTTTGTGTCAAGCCCACGGC
reTALE-M7        TTGACGCCTGAGCAGGTTGTGGCCATCGCTAGCNNNNNNGGAGGGAAGCAGGCTCTTGAAACCGTACAGCGACTTCTCCCAGTTTTGTGCCAAGCTCACGGG
reTALE-M8        CTAACCCCCGAGCAAGTAGTTGCCATAGCAAGCNNNNNNGGAGGAAAACAGGCATTAGAAACAGTTCAGCGCTTGCTCCCGGTACTCTGTCAGGCACACGGT
reTALE-M9        CTAACTCCGGAACAGGTCGTAGCCATTGCTTCCNNNNNNGGCGGCAAACAGGCGCTAGAGACAGTCCAGAGGCTCTTGCCTGTGTTATGCCAGGCACATGGC
reTALE-M10       CTCACCCCGGAGCAGGTCGTTGCCATCGCCAGTNNNNNNGGCGGAAAGCAAGCTCTCGAAACAGTACAACGGCTGTTGCCAGTCCTATGTCAAGCTCATGGA
reTALE-M11       CTGACGCCCGAGCAGGTAGTGGCAATCGCATCTNNNNNNGGAGGTAAACAAGCACTCGAGACTGTCCAAAGATTGTTACCCGTACTATGCCAAGCGCATGGT
reTALE-M12       TTAACCCCAGAGCAAGTTGTGGCTATTGCATCTNNNNNNGGTGGCAAACAAGCCTTGGAGACAGTGCAACGATTACTGCCTGTCTTATGTCAGGCCCATGGC
reTALE-M13       CTTACTCCTGAGCAAGTCGTAGCTATCGCCAGCNNNNNNGGTGGGAAACAGGCCCTGGAAACCGTACAACGTCTCCTCCCAGTACTTTGTCAAGCACACGGG
reTALE-M14       TTGACACCGGAACAAGTGGTGGCGATTGCGTCCNNNNNNGGAGGCAAGCAGGCACTGGAGACCGTCCAACGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGG
reTALE-M15       CTCACGCCAGAGCAGGTGGTAGCAATAGCGTCGNNNNNNGGTGGTAAGCAAGCGCTTGAAACGGTCCAGCGTCTTCTGCCGGTGTTGTGCCAGGCGCACGGA
reTALE-M16       CTCACACCAGAACAAGTGGTTGCTATTGCTAGTNNNNNNGGTGGAAAGCAGGCCCTCGAGACGGTGCAGAGGTTACTTCCCGTCCTCTGTCAAGCGCACGGC
reTALE-M17       CTCACTCCAGAGCAAGTGGTTGCGATCGCTTCANNNNNNGGTGGAAGACCTGCCCTGGAA
```

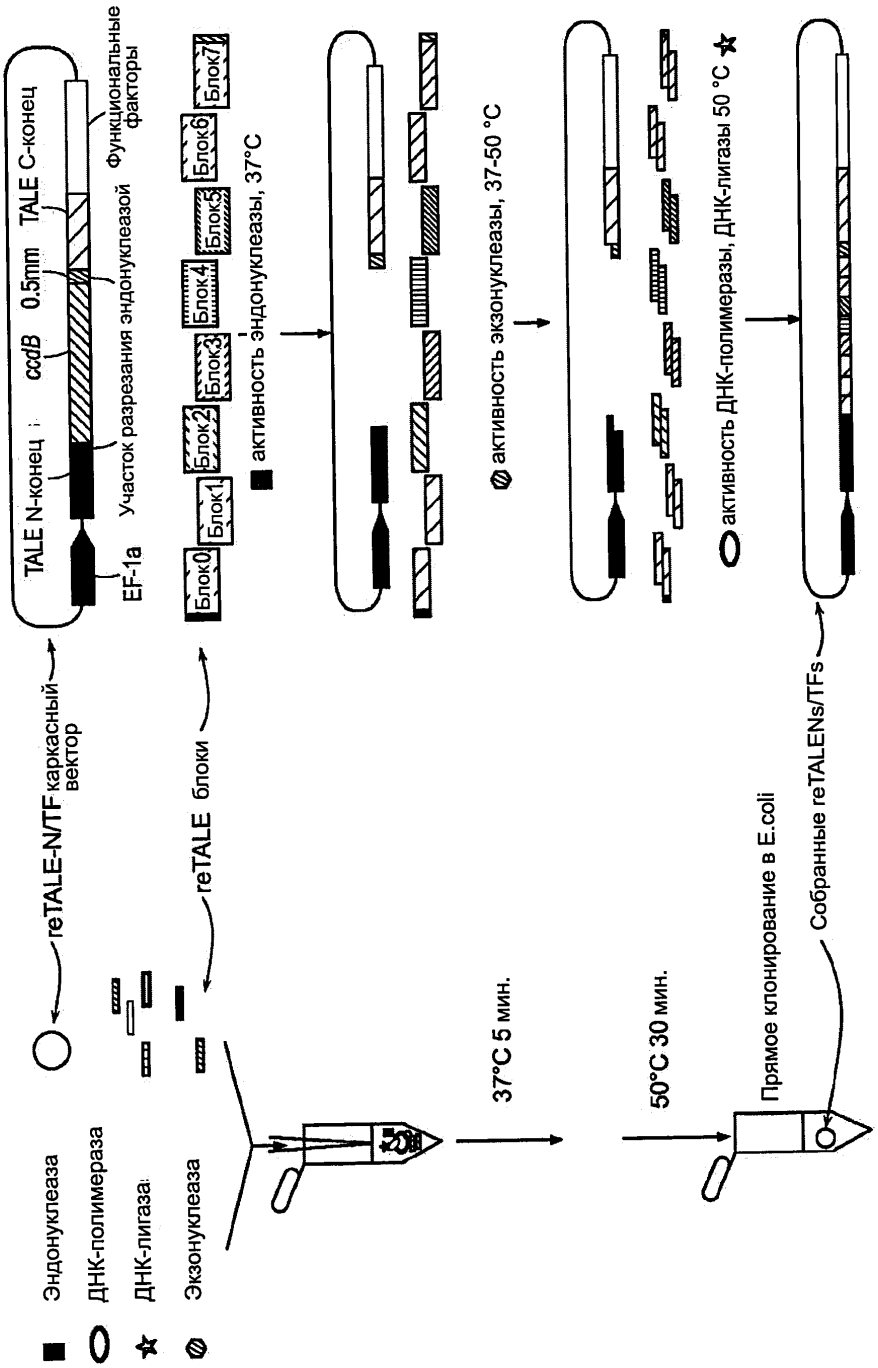
Фиг. 5B

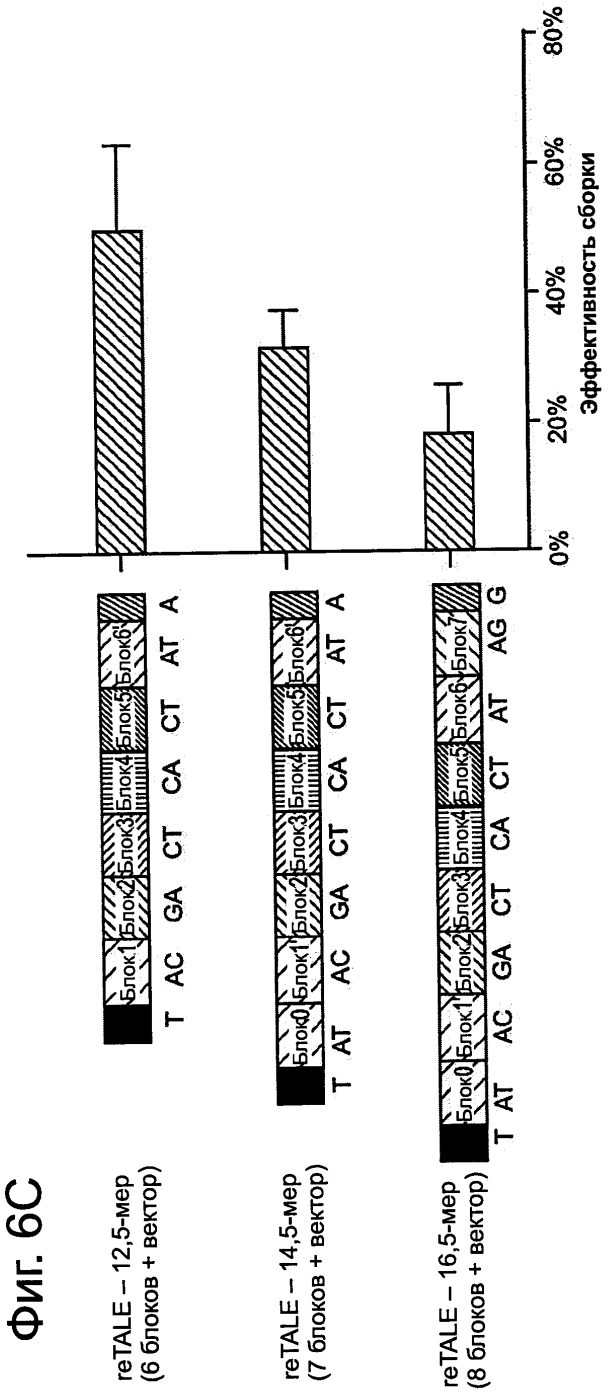

reTALE
TALE
FN F0 R0 R1 R2 R3 R4 R5 R6 R7 RC
TALE N-конец
TALE C-конец
область распознавания ДНК
FN повторяющийся TALE-MM
3.0тпо
1.0тпо
0.5тпо
0.2тпо
M
Матрица reTALE reTALE reTALE reTALE reTALE reTALE reTALE reTALE reTALE TALE
праймер 1 F0 F0 F0 F0 F0 F0 F0 F0 FN FN
праймер 2 R0 R1 R2 R3 R4 R5 R6 R7 RC RC

Фиг. 6A


Библиотека димерных блоков reTALE
Последовательность димера (Блок 6_AC)
Блок0
Блок1
Блок3
Блок5
Блок6'
Блок1'
Блок2
Блок4
Блок6
Блок7
16 димеров блока 6
AA AT AG AC
TA TT TG TC
GA GT GG GC
CA CT CG CC
CGATTACTGCCTGTCTTATGTCAGGCCCATGGC
R L L P V L C Q A H G
CTTACTCCTGAGCAAGTCGTAGCTATCGCCAGCAACATAGGTGGGAAACAG
L T P E Q V V A I A S N I G G K Q
GCCCTGGAAACCGTACAACGTCTCCTCCCAGTACTTTGTCAAGCACACGGG
A L E T V Q R L L P V L C Q A H G
TTGACACCGGAACAAGTGGTGGCGATTGCGTCCCACGATGGAGGCAAGCAG
L T P E Q V V A I A S H D G G K Q
GCACTGGAGACCGTCCAACGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGG
A L E T V Q R L L P V L C Q A H G
CTCACGCCA
L T P

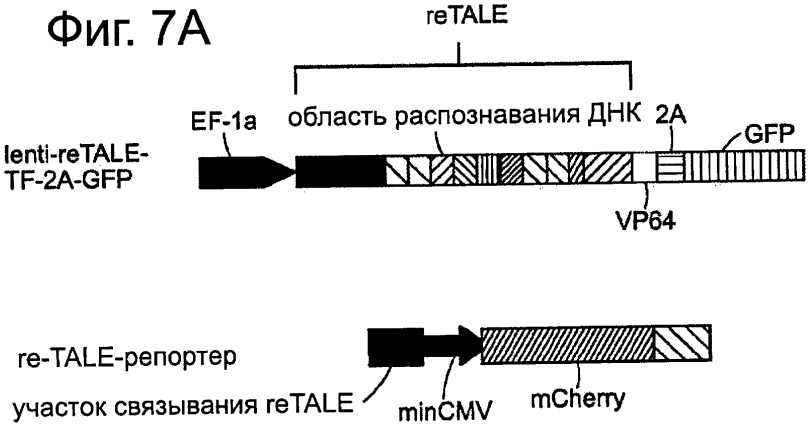
Фиг. 6B


Эндонуклеаза
ДНК-полимераза
ДНК-лигаза
Экзонуклеаза
reTALE-N/TF каркасный вектор
reTALE блоки
TALE N-конец
ccdB
0.5mm
TALE C-конец
EF-1a
Участок разрезания эндонуклеазой
Функциональные факторы
Блок0
Блок1
Блок2
Блок3
Блок3
Блок4
Блок6
Блок7
активность эндонуклеазы, 37°C
активность экзонуклеазы, 37-50 °C
активность ДНК-полимеразы, ДНК-лигазы 50 °C
37°C 5 мин.
50°C 30 мин.
Прямое клонирование в E.coli
Собранные reTALENs/TFs

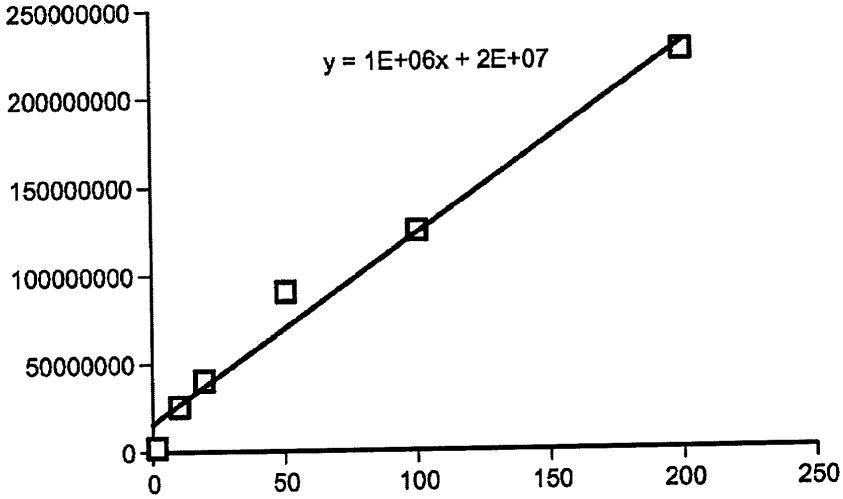

reTALE – 12,5-мер
(6 блоков + вектор)
reTALE – 14,5-мер
(7 блоков + вектор)
reTALE – 16,5-мер
(8 блоков + вектор)
Блок0
Блок1
Блок2
Блок3
Блок4
Блок5
Блок6
Блок7
T AC GA CT CA CT AT A
T AT AC GA CT CA CT AT A
T AT AC GA CT CA CT AT AG G
0%
20%
40%
60%
80%
Эффективность сборки

Фиг. 6C

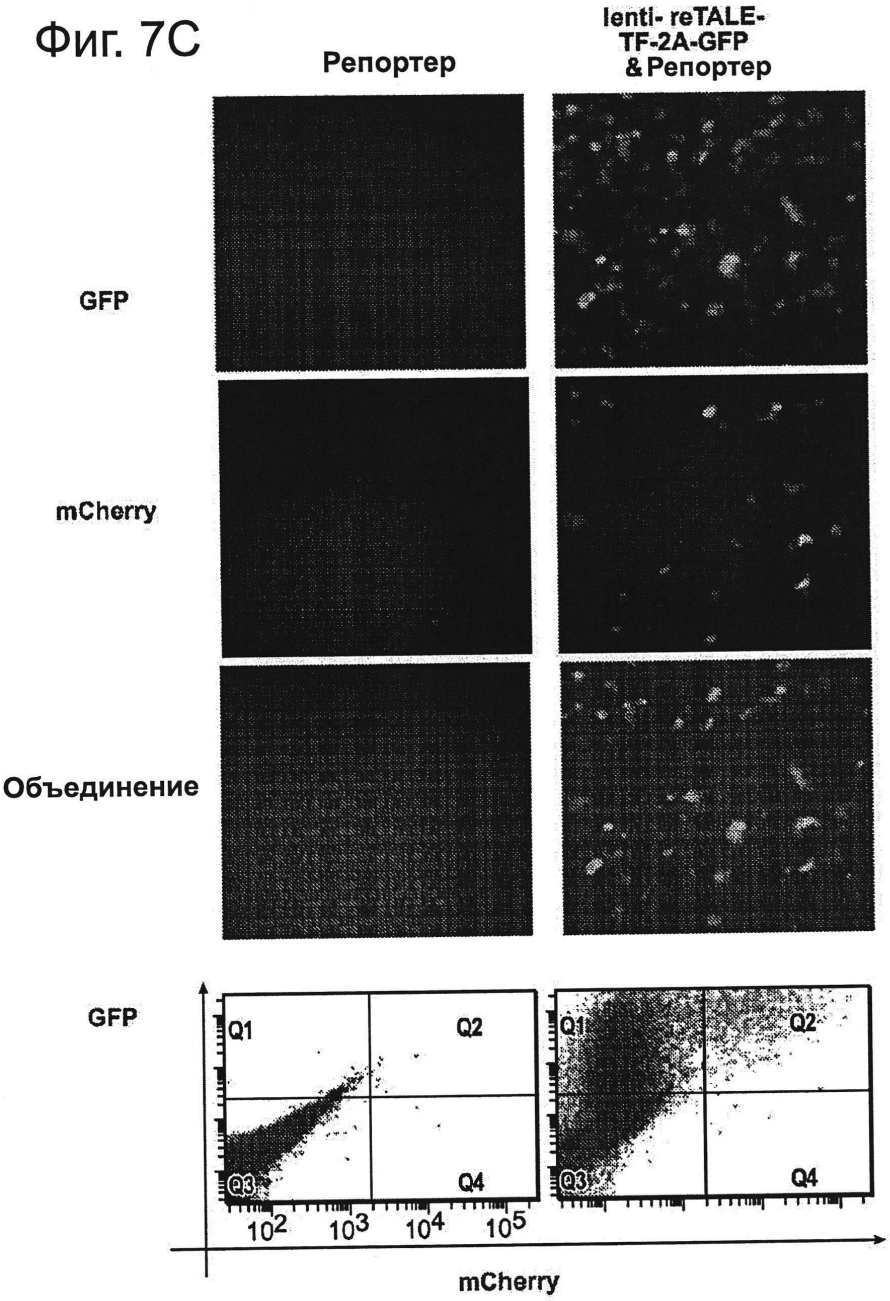
Фиг. 7A


reTALE
EF-1a
область распознавания ДНК
2A
GFP
lenti-reTALE-
TF-2A-GFP
VP64
re-TALE-репортер
участок связывания reTALE
minCMV
mCherry

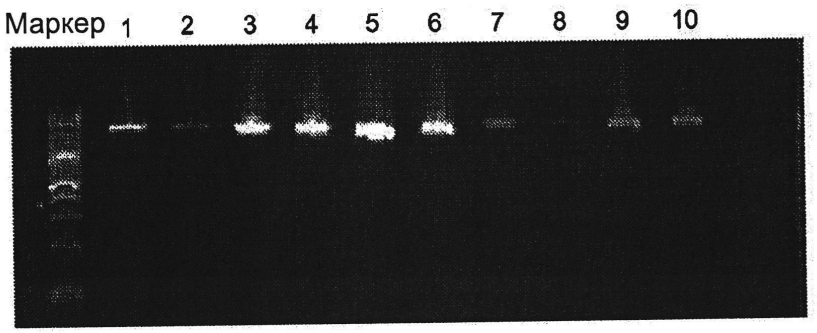
Фиг. 7B


y = 1E+06x + 2E+07
250000000
200000000
150000000
100000000
50000000
0
0
50
100
150
200
250

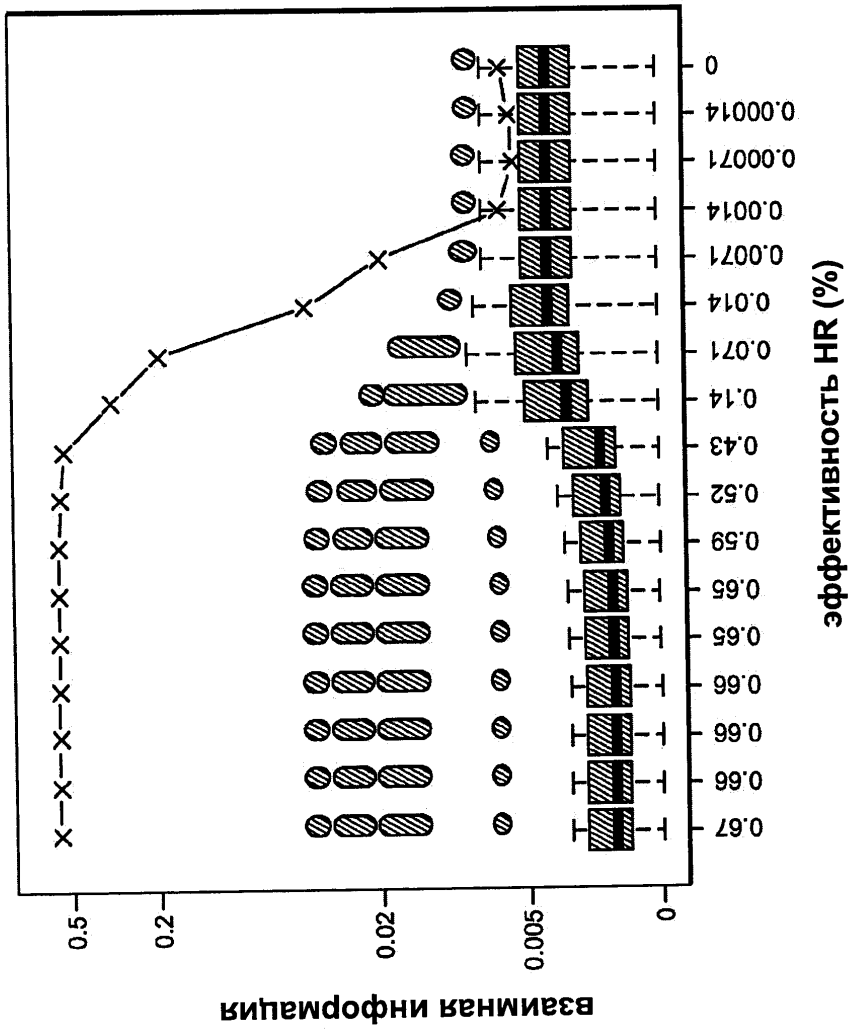
Фиг. 7C


Репортер
lenti- reTALE-TF-2A-GFP & Репортер
GFP
mCherry
Объединение
GFP
Q1
Q2
Q3
Q4
Q1
Q2
Q3
Q4
10^2
10^3
10^4
10^5
mCherry

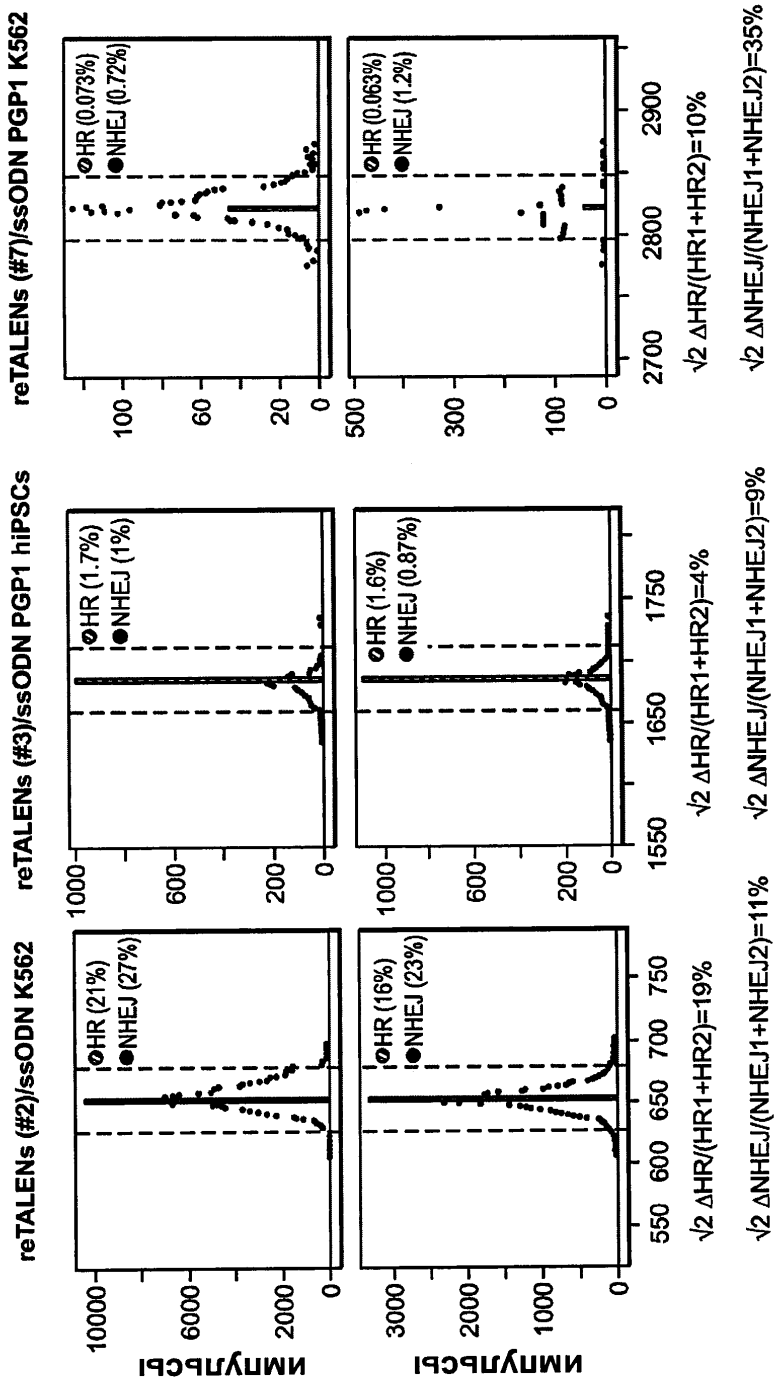
Фиг. 7D



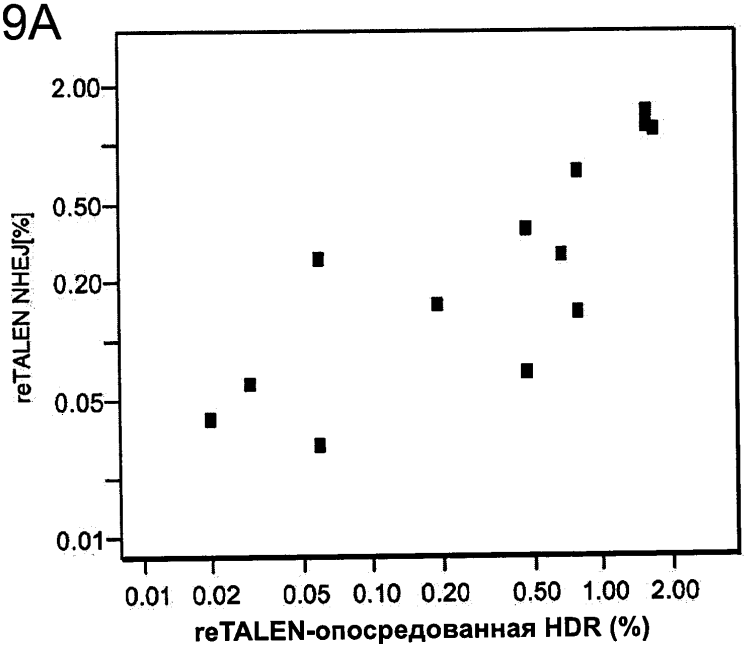
Фиг. 8A


взаимная информация
0.5
0.2
0.02
0.005
0
эффективность HR (%)
0.67
0.66
0.66
0.66
0.66
0.65
0.65
0.59
0.52
0.43
0.14
0.071
0.014
0.0071
0.0014
0.00071
0.00014
0

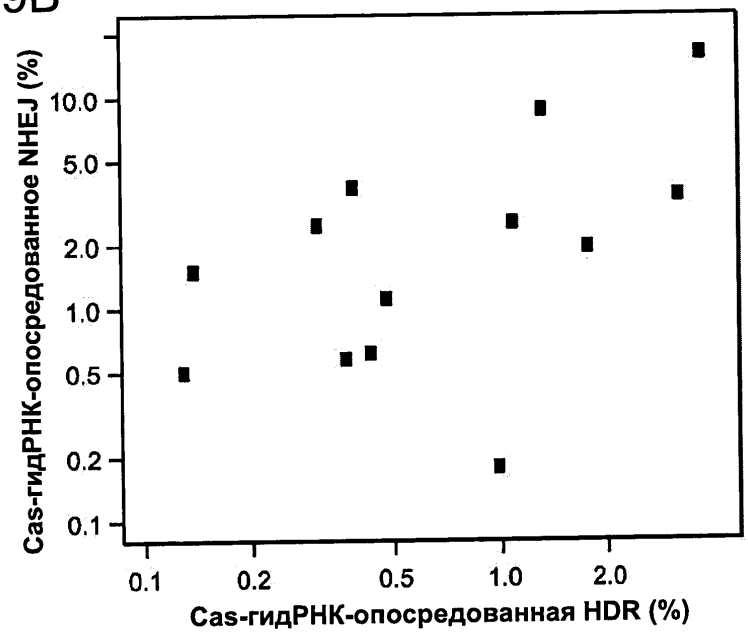
Фиг. 8B


reTALENs (#2)/ssODN K562
reTALENs (#3)/ssODN PGP1 hiPSCs
reTALENs (#7)/ssODN PGP1 K562
HR (21%)
NHEJ (27%)
HR (1.7%)
NHEJ (1%)
HR (0.073%)
NHEJ (0.72%)
HR (16%)
NHEJ (23%)
HR (1.6%)
NHEJ (0.87%)
HR (0.063%)
NHEJ (1.2%)
нмулрсы
√2 ΔHR/(HR1+HR2)=19%
√2 ΔHR/(HR1+HR2)=4%
√2 ΔHR/(HR1+HR2)=10%
√2 ΔNHEJ/(NHEJ1+NHEJ2)=11%
√2 ΔNHEJ/(NHEJ1+NHEJ2)=9%
√2 ΔNHEJ/(NHEJ1+NHEJ2)=35%

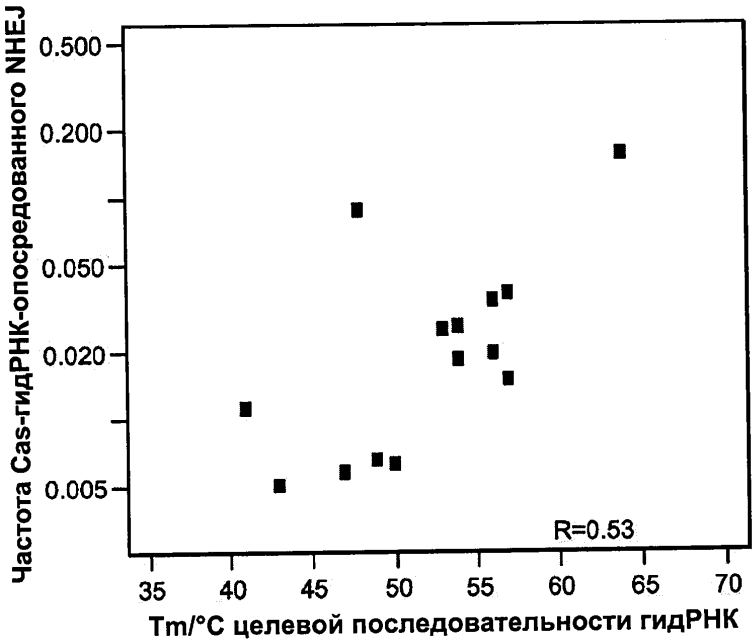
Фиг. 9A


2.00
0.50
0.20
0.05
0.01
reTALEN NHEJ[%]
0.01 0.02 0.05 0.10 0.20 0.50 1.00 2.00
reTALEN-опосредованная HDR (%)

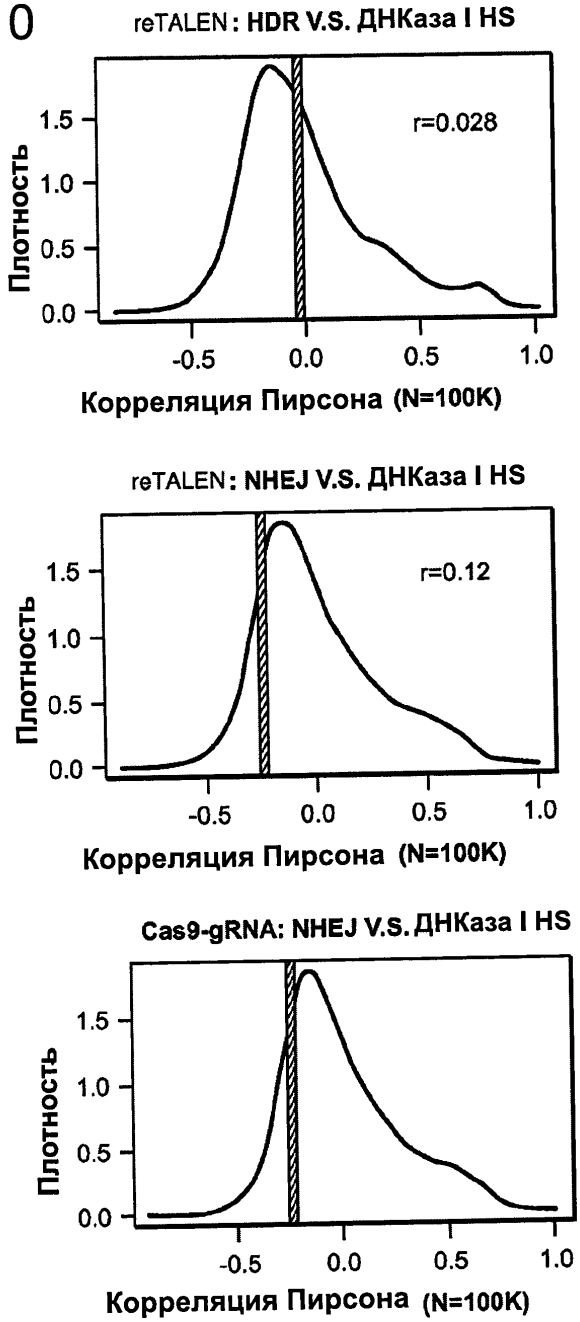
Фиг. 9B


10.0
5.0
2.0
1.0
0.5
0.2
0.1
Cas-гидРНК-опосредованное NHEJ (%)
0.1 0.2 0.5 1.0 2.0
Cas-гидРНК-опосредованная HDR (%)

Фиг. 9C

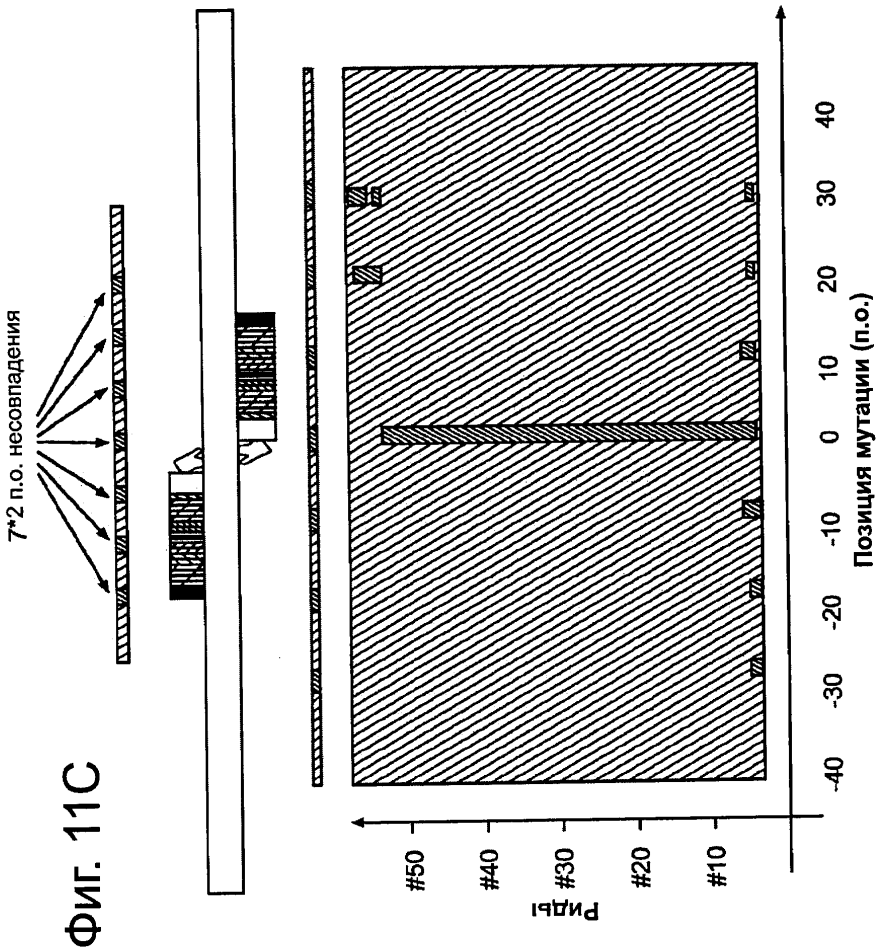

Частота Cas-гидРНК-опосредованного NHEJ
0.500
0.200
0.050
0.020
0.005
R=0.53
35
40
45
50
55
60
65
70
Tm/°C целевой последовательности гидРНК

Фиг. 10


reTALEN: HDR V.S. ДНКаза I HS
r=0.028
Плотность
1.5
1.0
0.5
0.0
-0.5
0.0
0.5
1.0
Корреляция Пирсона (N=100K)
reTALEN: NHEJ V.S. ДНКаза I HS
r=0.12
Плотность
1.5
1.0
0.5
0.0
-0.5
0.0
0.5
1.0
Корреляция Пирсона (N=100K)
Cas9-gRNA: NHEJ V.S. ДНКаза I HS
Плотность
1.5
1.0
0.5
0.0
-0.5
0.0
0.5
1.0
Корреляция Пирсона (N=100K)

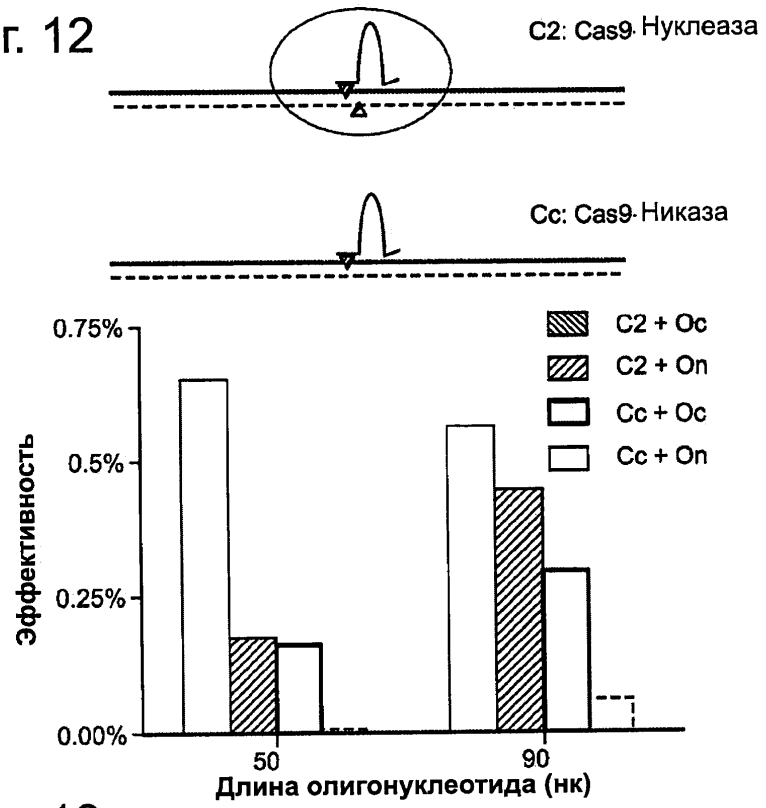
Фиг. 11A

Фиг. 11B


7*2 п.о. несовпадения
#50
#40
#30
#20
#10
Риды
-40
-30
-20
-10
0
10
20
30
40
Позиция мутации (п.о.)

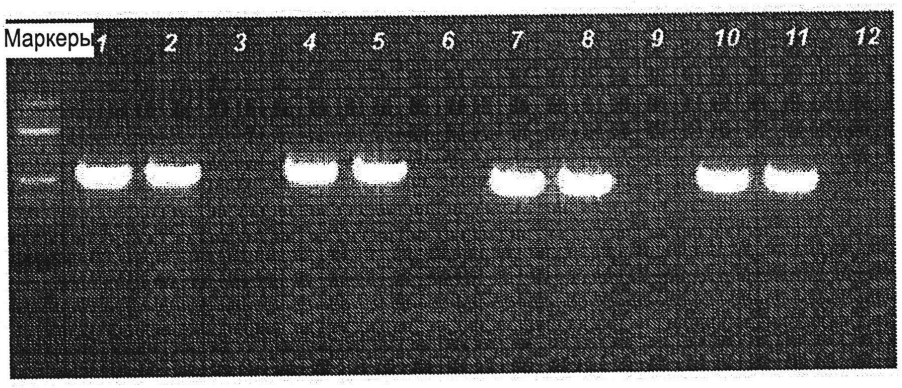
Фиг. 11C

Фиг. 12

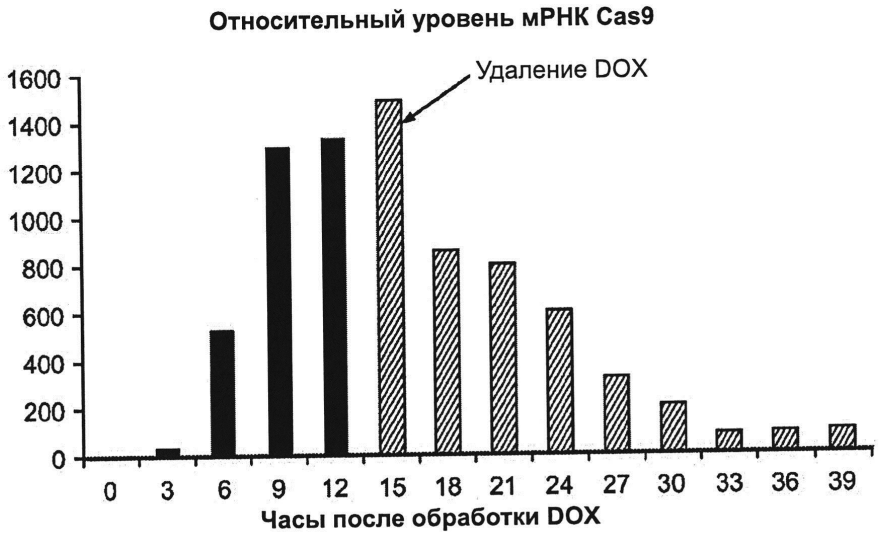

C2: Cas9-Нуклеаза
Cc: Cas9-Никаза
C2 + Oc
C2 + On
Cc + Oc
Cc + On
0.75%
0.5%
0.25%
0.00%
Эффективность
50
90
Длина олигонуклеотида (нк)

Фиг. 13


350
250
150
50
0
количество повторов
2
4
6
8
10
12
14
количество циклов

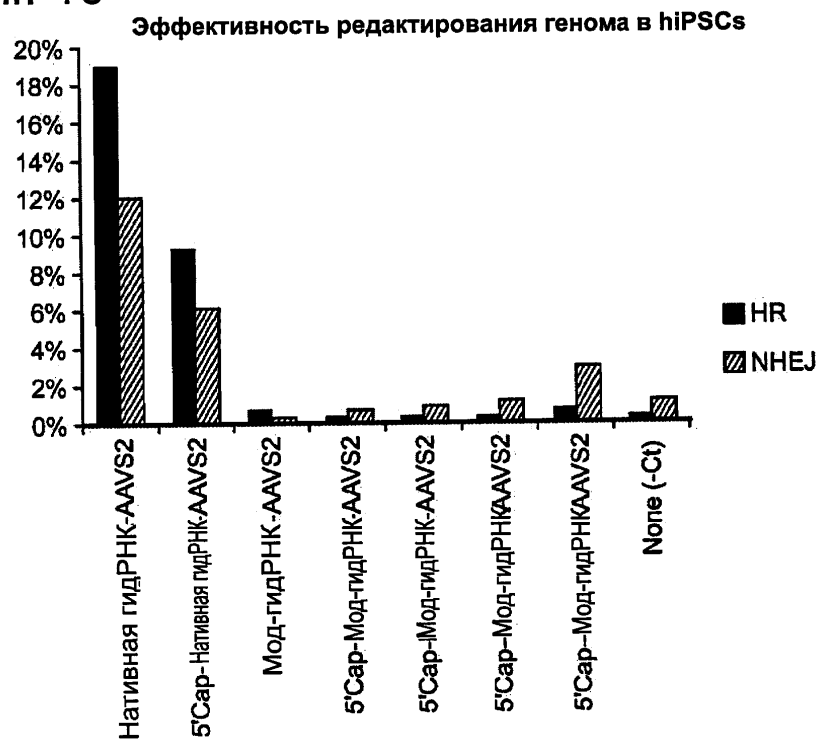

Маркеры
1
2
3
4
5
6
7
8
9
10
11
12

Фиг. 14

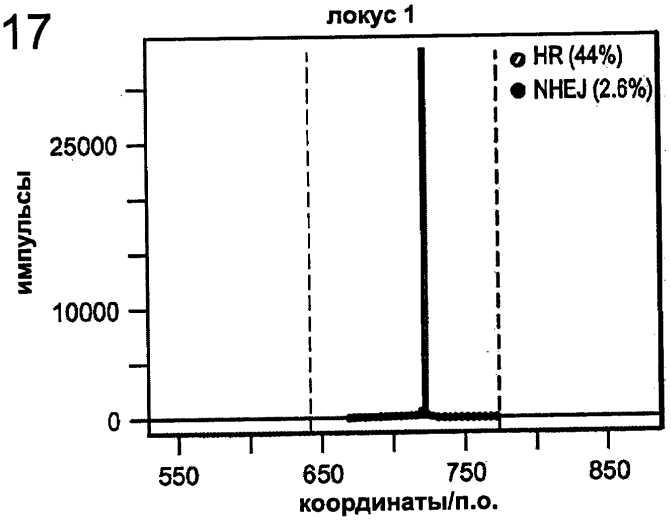

Относительный уровень мРНК Cas9
Удаление DOX
1600
1400
1200
1000
800
600
400
200
0
0
3
6
9
12
15
18
21
24
27
30
33
36
39
Часы после обработки DOX

Фиг. 15

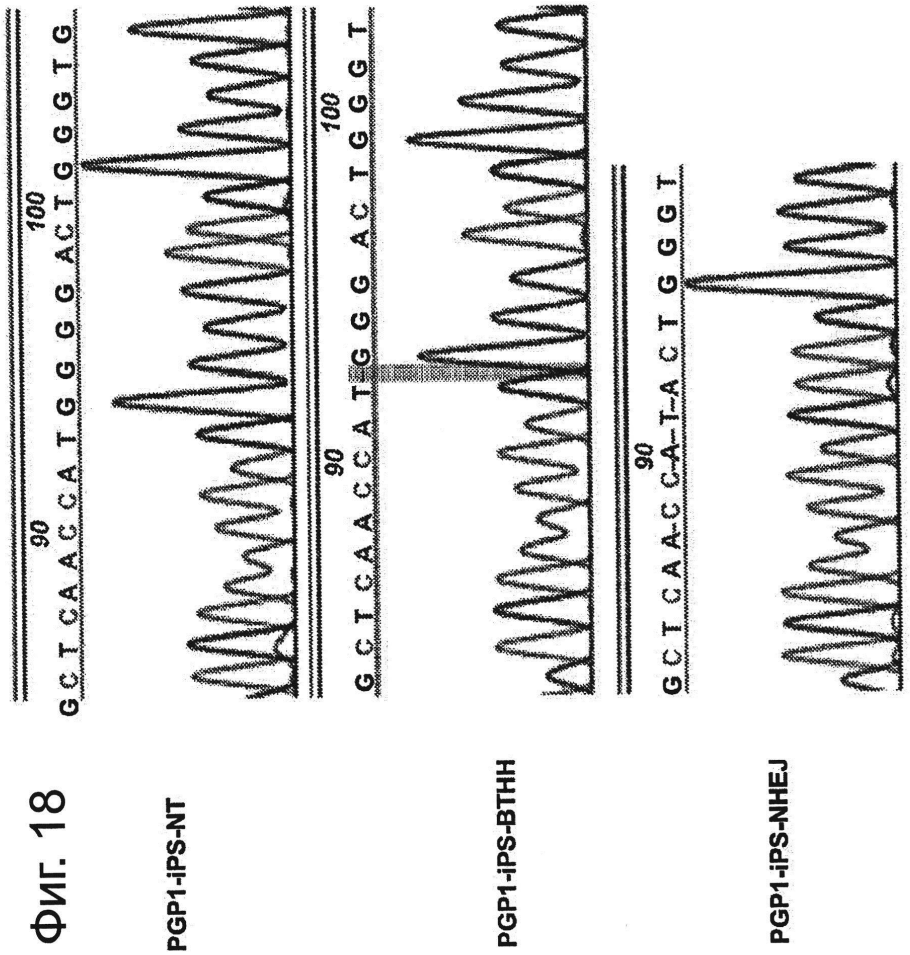
Фиг. 16


Эффективность редактирования генома в hiPSCs
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
HR
NHEJ
Нативная гидРНК-AAVS2
5'Cap-Нативная гидРНК-AAVS2
Мод-гидРНК-AAVS2
5'Cap-Мод-гидРНК-AAVS2
5'Cap-Мод-гидРНК-AAVS2
5'Cap-Мод-гидРНК-AAVS2
5'Cap-Мод-гидРНК-AAVS2
None (-Ct)

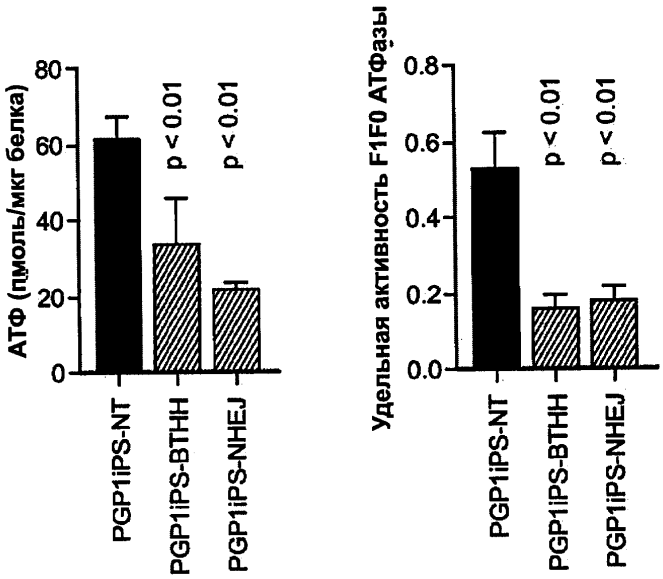
Фиг. 17


локус 1
HR (44%)
NHEJ (2.6%)
импульсы
25000
10000
0
550
650
750
850
координаты/п.о.

Фиг. 18


PGP1-iPS-NT
PGP1-iPS-BTHH
PGP1-iPS-NHEJ
90
100
G C T C A A C C A T G G G G A C T G G G T G
G C T C A A C C A T G G G A C T G G G T
G C T C A A C C-A- T-A C T G G G T

Фиг. 19


АТФ (пмоль/мкг белка)
80
60
40
20
0
p < 0.01
p < 0.01
PGP1iPS-NT
PGP1iPS-BTHH
PGP1iPS-NHEJ
Удельная активность F1F0 АТФазы
0.8
0.6
0.4
0.2
0.0
p < 0.01
p < 0.01
PGP1iPS-NT
PGP1iPS-BTHH
PGP1iPS-NHEJ

Фиг. 20A

re-TALEN-каркасная последовательность

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

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGAC
TCGTTCTCAGACCTGCTTAGGCAGTTCGACCCCTCACTGTTTAACACATCGTTGTTCG
ACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGTGGGAT
GAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGC
CGTCACCGCTGCCCGACCGCCGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGC
AACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTTTGGGATATTCA
CAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATC
ACGAAGCGCTGGTGGGTCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGC
ACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGCGTTGC
CGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCG
AGCGCTTGAGGCCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGC
TGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGGGAGGAGTCACGGCGGTCGAG
GCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAAC*AGTTCACG*TG
ACAGAGACCGCGGCCGCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCG
TATAATGTGTGGATTTTGAGTTAGGATCCGTCGAGATTTTCAGGAGCTAAGGAAGCT
AAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCG
TAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGT
TCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTA
TCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATG
GCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTT

Фиг. 20B

TTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTC
CGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCC
TATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGA
GTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTT
CACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTC
AGGTTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTAC
AACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAAGATCTGGATCCGGCTTACTA
AAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATAT
ACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTATGCTATGAAGCAGCGTATT
ACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATA
TCTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCGTCTGCGTGCCGAAC
GCTGGAAAGCGGAAAATCAGGAAGGGATGGCTGAGGTCGCCCGGTTTATTGAAATG
AACGGCTCTTTTGCTGACGAGAACAGGGGCTGGTGAAATGCAGTTTAAGGTTTACAC
CTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGA
CACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATA
AAGTCTCCCGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGA
TGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATC
TCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATA
TAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTCGACGGTCTC***GCTCTTC***G
AAGGTTACTTCCCGTCCTCTGTCAAGCGCACGGCCTCACTCCAGAGCAAGTGGTTGC
GATCGCTTCAAACAACGGTGGAAGACCTGCCCTGGAA<u>TCAATCGTGGCCCAGCTTTC</u>

Фиг. 20C

GAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTG
CCTCGGCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTG
*CATTGATTAAGCGGACCAACAGAAGGATTCCCGAGAGGACAT*CACATCGAGTGGCA
GGTTCCCAACTCGTGAAGAGTGAACTTGAGGAGAAAAAGTCGGAGCTGCGGCACAA
ATTGAAATACGTACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGAC
TCAAGACAGAATCCTTGAGATGAAGGTAATGGAGTTCTTTATGAAGGTTTATGGATA
CCGAGGGAAGCATCTCGGTGGATCACGAAAACCCGACGGAGCAATCTATACGGTGG
GGAGCCCGATTGATTACGGAGTGATCGTCGACACGAAAGCCTACAGCGGTGGGTAC
AATCTTCCCATCGGGCAGGCAGATGAGATGCAACGTTATGTCGAAGAAAATCAGAC
CAGGAACAAACACATCAATCCAAATGAGTGGTGGAAAGTGTATCCTTCATCAGTGA
CCGAGTTTAAGTTTTTGTTTGTCTCTGGGCATTTCAAAGGCAACTATAAGGCCCAGCT
CACACGGTTGAATCACATTACGAACTGCAATGGTGCGGTTTTGTCCGTAGAGGAACT
GCTCATTGGTGGAGAAATGATCAAAGCGGGAACTCTGACACTGGAAGAAGTCAGAC
GCAAGTTTAACAATGGCGAGATCAATTTCCGC

каркасная последовательность re-TALEN-TF

(**фиолетовый: re-TALE-N; красный: участок SapI**; *зеленый: 0,5 мономера;*
синий: re-TALEN-C; оранжевый: NLS-M364;
2A-GFP выделен зеленым)

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGAC
TCGTTCTCAGACCTGCTTAGGCAGTTCGACCCCTCACTGTTTAACACATCGTTGTTCG

Фиг. 20D

ACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGTGGGAT
GAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGC
CGTCACCGCTGCCCGACCGCCGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGC
AACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTTTGGGATATTCA
CAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATC
ACGAAGCGCTGGTGGGTCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGC
ACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACATGATTGCGGCGTTGC
CGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCG
AGCGCTTGAGGCCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGC
TGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGGGAGGAGTCACGGCGGTCGAG
GCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAACA*AGTTCACG*CTG
ACAGAGACCGCGGCCGCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCG
TATAATGTGTGGATTTTGAGTTAGGATCCGTCGAGATTTTCAGGAGCTAAGGAAGCT
AAAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCG
TAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGT
TCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTA
TCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATG
GCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTT
TTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTC
CGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCC
TATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGA
GTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTT

Фиг. 20E

CACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTC
AGGTTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTAC
AACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAAGATCTGGATCCGGCTTACTA
AAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAAGAATATAT
ACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTATGCTATGAAGCAGCGTATT
ACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATA
TCTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCGTCTGCGTGCCGAAC
GCTGGAAAGCGGAAAATCAGGAAGGGATGGCTGAGGTCGCCCGGTTTATTGAAATG
AACGGCTCTTTTGCTGACGAGAACAGGGGCTGGTGAAATGCAGTTTAAGGTTTACAC
CTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGA
CACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATA
AAGTCTCCCGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGA
TGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATC
TCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATA
TAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTCGACGGTCTC***GCTCTTC***G
AAGGTTACTTCCCGTCCTCTGTCAAGCGCACGGCCTCACTCCAGAGCAAGTGGTTGC
GATCGCTTCAAACAACGGTGGAAGACCTGCCCTGGAATCAATCGTGGCCCAGCTTTC
GAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTG
CCTCGGCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTG
CATTGATTAAGCGGACCAACAGAAGGATTCCCGAGAGGACATAGCCCCAAGAAGAA
GAGAAAGGTGGAGGCCAGCGGTTCCGGACGGGCTGACGCATTGGACGATTTTGATC

Фиг. 20F

TGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGG
ATGCCCTTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCG
ACCTGGACATGCTGATTAACTCTAGAGGCAGTGGAGAGGGCAGAGGAAGTCTGCTA
ACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGTGAGCAAGGGCGAGGAGCTGTT
CACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG
TTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTG
ACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTC
TTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCCACACCCTGGTGAA
CCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAG
AACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGC
CCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAG
CGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATG
GACGAGCTGTACAAG