

(19)



Евразийское
патентное
ведомство

(11) 039384

(13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2022.01.20

(21) Номер заявки
201590167

(22) Дата подачи заявки
2013.08.29

(51) Int. Cl. C12N 5/071 (2010.01)

(54) БЕЛОК "ЦИНКОВЫЕ ПАЛЬЦЫ" ДЛЯ МОДУЛИРОВАНИЯ ЭКСПРЕССИИ ГЕНА BCL11A И СПОСОБ МОДУЛИРОВАНИЯ ЭКСПРЕССИИ ГЛОБИНОВОГО ГЕНА

(31) 61/694,693

(32) 2012.08.29

(33) US

(43) 2015.11.30

(86) PCT/US2013/057214

(87) WO 2014/036219 2014.03.06

(71)(73) Заявитель и патентовладелец:
САНГАМО БАЙОСАЕНСЕС, ИНК.
(US)

(56) US-A1-20120142062
US-B1-8012946
WO-A2-200242459
US-A1-20100261271

(72) Изобретатель:
Кост Грегори Дж., Грегори Филип Д.,
Гущин Дмитрий, Холмс Майкл К.,
Миллер Джефффри К., Пашен Дэвид,
Ребар Эдвард Дж., Рейк Андреас,
Урнов Фёдор, Чжан Лей (US)

(74) Представитель:
Нилова М.И. (RU)

(57) Изобретение относится к области генной инженерии. В частности, предложен белок "цинковые пальцы" для модулирования экспрессии гена BCL11A, содержащий 5 или 6 доменов "цинковых пальцев", каждый из которых содержит спиральную область распознавания, которая распознает сайт-мишень, причем указанный белок связывается с определенным сайтом-мишенью. Также предложен полинуклеотид, кодирующий указанный белок, клетки, модифицированные указанным белком или полинуклеотидом (варианты), а также способ модулирования экспрессии глобинового гена и набор для модулирования экспрессии гена BCL11A.

B1

039384

039384

B1

Перекрестная ссылка на родственные заявки

Настоящая заявка испрашивает приоритет предварительной заявки США № 61/694693, поданной 29 августа 2012 г., раскрытие которой включено в настоящий документ посредством ссылки в полном объеме.

Область техники

Настоящее изобретение относится к области генной инженерии гемопоэтических стволовых клеток, особенно для лечения гемоглобинопатии.

Известный уровень техники

Генная терапия имеет огромный потенциал для новой эры в медицине. Эти методики позволят лечить заболевания, которые до сих пор были недоступны стандартной медицинской практике. Одной из областей особенно перспективной является способность генетически конструировать клетку таким образом, чтобы эта клетка экспрессировала продукт, ранее не вырабатываемый в этой клетке. Примеры применения этой технологии включают введение гена, кодирующего новый терапевтический белок, вставку кодирующей последовательности, которая кодирует белок, отсутствующий в клетке или у индивидуума, введение гена дикого типа в клетку, содержащую мутантную генную последовательность, и вставку последовательности, кодирующей структурную нуклеиновую кислоту, такую как микроРНК или миРНК.

Трансгены могут быть доставлены в клетку с помощью различных способов, таким образом, что трансген интегрируется в собственный геном клетки и поддерживается там. В последние годы была разработана стратегия интеграции трансгена, в которой используется расщепление с помощью сайт-специфичных нуклеаз для целевого введения в выбранный геномный локус (см., например, патент США 7888121 того же заявителя). Нуклеазы, специфичные для генов-мишеней, могут быть использованы таким образом, что трансгенный конструкт вставляется либо путем гомологично направленной репарации (HDR), или с помощью процессов, приводимых в движение концевым захватом во время нехомологичного концевого присоединения (NHEJ). Локусы-мишени включают локусы "безопасной гавани", например ген CCR5, ген CXCR4, ген PPP1R12C (также известный как AAVS1), ген альбумина или ген Rosa. См., например, патентные публикации США 20080299580; 20080159996; 201000218264; 20110301073; 20130177983 и 20130177960 и предварительную заявку на патент США № 61/823689. Опосредованная нуклеазой интеграция открывает перспективу улучшения экспрессии трансгенов, повышения безопасности и устойчивости экспрессии по сравнению с классическими подходами к интеграции, которые полагаются на случайную интеграцию трансгена, так как она делает возможным точное позиционирование трансгена с минимальным риском для сайленсинга гена или активации соседних онкогенов. Красные клетки крови (RBC), или эритроциты, являются основными клеточными компонентами крови. В самом деле, эритроциты составляют четверть клеток у человека. У человека зрелые эритроциты не содержат ядра и многих других органелл и полны гемоглобином - металлопротеином, находящегося в эритроцитах, который переносит кислород к тканям, а также переносит диоксид углерода из тканей обратно в легкие для удаления. Белок составляет примерно 97% от сухой массы эритроцитов и повышает способность крови переносить кислород приблизительно в семьдесят раз. Гемоглобин является гетеротетрамером, состоящим из двух α -подобных глобиновых цепей и двух β -подобных глобиновых цепей и 4 темных групп. У взрослых $\alpha_2\beta_2$ тетрамер называется гемоглобином А (HbA) или взрослым гемоглобином. Как правило, α - и β -глобиновые цепи синтезируются в приблизительном соотношении 1:1, и это отношение, по-видимому, имеет решающее значение в плане стабилизации гемоглобина и RBC. В самом деле, в некоторых случаях, когда один тип гена глобина неадекватно экспрессируется (см. ниже), снижая экспрессию (например, с использованием конкретной миРНК) другого типа глобина, восстановление указанного соотношения 1:1 снижает некоторые аспекты мутантного клеточного фенотипа (см. Voon et al. (2008) Haematologica 93(8): 1288). У развивающегося плода вырабатывается другая форма гемоглобина, фетальный гемоглобин (HbF), который обладает более высокой аффинностью связывания кислорода, чем гемоглобин А, таким образом, что кислород может доставляться в систему ребенка через кровоток матери. Фетальный гемоглобин также содержит две α -цепи глобина, но вместо взрослых β -глобиновых цепей он содержит две фетальные γ -цепи глобина (т.е. фетальный гемоглобин $\alpha_2\gamma_2$). Приблизительно на 30 неделе беременности синтез γ -глобина у плода начинает падать, а выработка β -глобина увеличивается. Примерно к 10-месячному возрасту гемоглобин новорожденного почти полностью представляет собой $\alpha_2\beta_2$, хотя некоторое количество HbF сохраняется и в зрелом возрасте (примерно 1-3% от общего гемоглобина). Регулирование перехода от выработки γ к β является довольно сложным, и в первую очередь включает отрицательную регуляцию экспрессии γ -глобина с одновременной положительной регуляцией экспрессии β -глобина. Генетические дефекты в последовательностях, кодирующих гемоглобиновые цепи, могут быть ответственными за целый ряд заболеваний, известных как гемоглобинопатии, в том числе серповидно-клеточную анемию и талассемию. У большинства пациентов с гемоглобинопатиями гены, кодирующие γ -глобин, остаются в наличии, но экспрессия является относительно низкой из-за нормальной генной репрессии, происходящей накануне родов, как раскрыто выше.

Считается, что у 1 из 5000 человек в США присутствует серповидно-клеточная анемия (SCD), в основном у людей, происходящих из Африки к югу от Сахары. Похоже, что там серповидно-клеточная ге-

терозиготность полезна для защиты от малярии, следовательно, произошел отбор этой аномалии с течением времени, таким образом, что, по оценкам, в Африке к югу от Сахары у одной трети населения присутствует серповидно-клеточная аномалия эритроцитов. Серповидно-клеточная анемия вызвана мутацией в гене β -глобина, в котором валин замещается глутаминовой кислотой в аминокислотном положении 6 (GAG в GTG на уровне ДНК), причем образующийся гемоглобин называют "гемоглобином S" или "HbS". В условиях сниженного содержания кислорода конформационный сдвиг в дезоксиформу HbS воздействует на гидрофобный участок на белке между спиралями E и F. Гидрофобные остатки валина в положении 6 β -цепи гемоглобина способны ассоциировать с гидрофобным участком, вызывая агрегацию HbS и образование волокнистых преципитатов. Эти агрегаты, в свою очередь, вызывают аномалии или образование серповидных форм эритроцитов, что приводит к потере гибкости клеток. Образованные серповидные формы RBC больше не в состоянии протиснуться в капиллярное ложе и могут привести к вазоокклюзивному кризу у пациентов с серповидными клетками. Кроме того, серповидные эритроциты являются более хрупкими, чем нормальные эритроциты, и, как правило, проявляют тенденцию к гемолизу, что, в конечном счете, приводит к анемии у пациента.

Лечение и ведение больных с серповидно-клеточной анемией является пожизненной проблемой, включающей лечение антибиотиками, лечение боли и переливания во время острых эпизодов. Один из подходов состоит в использовании гидроксимочевина, которая оказывает свое действие в части увеличения производства γ -глобина. Однако долгосрочные побочные эффекты хронической терапии гидроксимочевинной до сих пор неизвестны, и лечение дает нежелательные побочные эффекты и может иметь переменную эффективность в зависимости от пациента. Несмотря на увеличение эффективности лечения серповидных клеток, продолжительность жизни больных по-прежнему остается на уровне середины-конца 50-х годов, и сопутствующие заболевания оказывают глубокое влияние на качество жизни пациента.

Талассемии также являются заболеваниями, связанными с гемоглобином, и обычно включают пониженную экспрессию цепей глобина. Это может происходить за счет мутаций в регуляторных областях генов или из-за мутации в кодирующей последовательности глобина, что приводит к снижению экспрессии. α -талассемии ассоциируются с выходцами из Западной Африки и Южной Азии и могут обеспечивать резистентность к малярии. β -талассемия связана с людьми средиземноморского происхождения, как правило, из Греции и прибрежных районов Турции и Италии. Лечение талассемии обычно включает переливание крови и терапию хелаторами железа. Пересадка костного мозга также используется для лечения людей с тяжелыми талассемиями, если можно выявить подходящего донора, но эта процедура может содержать в себе значительный риск. Одним из подходов для лечения SCD и β -талассемий, который был предложен, является повышение экспрессии γ -глобина, чтобы HbF функционально заменил аберрантный взрослый гемоглобин. Как упоминалось выше, лечение пациентов с SCD гидроксимочевинной считается успешным, отчасти из-за ее влияния на повышение экспрессии γ -глобина. Первой обнаруженной группой соединений, влияющих на реактивацию активности HbF, были цитотоксические средства. Способность вызывать *de novo* синтез γ -глобина путем фармакологических манипуляций впервые была показана с использованием 5-азациитидина у экспериментальных животных (DeSimone (1982) Proc. Natl. AcadSci USA 79(14):4428-31). Последующие исследования подтвердили способность 5-азациитидина увеличивать содержание HbF у пациентов с β -талассемией и серповидно-клеточной анемией (Ley, et al., (1982) N. Engl. J. Medicine, 307: 1469-1475, и Ley, et al., (1983) Blood 62: 370-380). Кроме того, короткоцепочечные жирные кислоты (например, бутират и его производные), как было показано в экспериментальных системах, увеличивают содержание HbF (Constantoulakis et al., (1988) Blood 72(6): 1961-1967). Также существует сегмент человеческой популяции с состоянием, известным как "наследственная персистенция фетального гемоглобина" (HPFH), при котором повышенные количества HbF сохраняются в зрелом возрасте (10-40% у гетерозигот HPFH (см. Thein et al. (2009) Hum. Mol. Genet. 18 (R2): R216-R223)). Это редкое состояние, но в отсутствие каких-либо связанных β -глобиновых аномалий, не связано с какими-либо значительными клиническими проявлениями, даже если 100% гемоглобина человека является HbF. У людей с β -талассемией, у которых одновременно присутствует HPFH, экспрессия HbF может уменьшить тяжесть заболевания. Кроме того, тяжесть естественного течения серповидно-клеточной анемии может значительно варьировать от пациента к пациенту, и эту изменчивость, в частности, можно отнести к тому, что у некоторых людей с более легкой формой болезни экспрессируются более высокие уровни HbF.

Один из подходов к увеличению экспрессии HbF включает идентификацию генов, продукты которых играют важную роль в регуляции экспрессии γ -глобина. Одним из таких генов является BCL11A, впервые идентифицированный благодаря его роли в развитии лимфоцитов. BCL11A кодирует белок "цинковые пальцы", который, как считается, участвует в стадии специфического регулирования экспрессии γ -глобина. BCL11A экспрессируется у взрослых в эритроидных клетках-предшественниках и отрицательная регуляция его экспрессии приводит к увеличению экспрессии γ -глобина. Кроме того, по-видимому, сплайсинг мРНК BCL11A регулируется стадией развития.

По-видимому, в эмбриональных клетках изначально экспрессируются более короткие варианты мРНК BCL11A, известные как BCL11A-S и BCL11A-XS, в то время как во взрослых клетках преимущественно экспрессируются более длинные варианты мРНК - BCL11A-L и BCL11A-XL. См. Sankaran et al.

(2008) Science 322 p. 1839. По-видимому, белок BCL11A взаимодействует с β -глобиновым локусом, чтобы изменить его конформацию и тем самым его экспрессию на разных стадиях развития. Кроме того, другой регуляторный белок KLF1, по-видимому, участвует в регуляции экспрессии γ -глобина. Было обнаружено, что уровни KLF1 прямо пропорциональны уровню BCL11A, и оба обратно пропорциональны уровню γ -глобина. Например, в мальтийском семействе с постоянной экспрессией HbF, семейство несет гетерозиготную мутацию гена KLF1 (Borg et al. (2010) Nat. Genet., 42(9):801-805). По-видимому, генный продукт KLF1 связывается непосредственно с геном BCL11A *in vivo* и, таким образом, может быть ответственным за его положительную регуляцию (см. Borg et al. там же; Bieker (2010) Nat. Genet. 42(9): 733-734; Zhou et al. (2010) Nat. Genet. 42(9):742-744). Таким образом, если KLF1 стимулирует экспрессию BCL11A, действие этого индуцированного BCL11A приведет к подавлению выработки γ -глобина и HbF. Было предложено использование ингибирующей РНК, нацеленной на ген BCL11A (см., например, патентную публикацию 20110182867), но этой технологии присущи несколько потенциальных недостатков, а именно, нельзя достигнуть полного нокдауна, доставка таких РНК может быть проблематичной и РНК должны присутствовать непрерывно, что требует многократного лечения на протяжении жизни. α -Талассемии также широко распространены у населения, особенно в Азии, а некоторые типы абберрации α -глобина, как считается, являются широко распространенным генетическим расстройством у человека. В тропических и субтропических районах мира α -глобиновое нарушение встречается у 80-90% населения (см. Harteveld and Higgs (2010) Orphanet Journal of Rare Diseases 5:13).

Люди несут две копии гена α -глобина в тандеме ($\alpha 1$ и $\alpha 2$) на хромосоме 16, следовательно, в нормальной диплоидной клетке присутствуют все 4 копии. Ген $\alpha 2$ обычно ответственен за количество мРНК α -глобина, в 2-3 раза большее, чем ген $\alpha 1$. Тандемная организация этих двух генов может быть связана с высокой распространенностью крупных делеций на α -глобиновых генах у больных α -талассемией, где, как правило, количество α -глобиновых генов, которые не являются функциональными, связано непосредственно с тяжестью любой α -талассемии (см. Chui et al. (2003) Blood 101(3):791). Делеция одной копии, по-видимому, в некоторой степени является обычной (30% афроамериканцев и 60-80% людей, проживающих в Саудовской Аравии, Индии и Таиланде), и, как правило, не обнаруживается у человека без проведения генетического тестирования. Делеция двух копий, на одной и той же хромосоме (*cis*) или по одной из каждой хромосомы (*trans*), может привести к легкой форме анемии у пациента. Если удалены три α -глобиновых гена, т.е. у индивидуума присутствует только один действующий α -глобиновый ген, то обнаруживается умеренная анемия, но что еще более важно, нарушается решающее соотношение α -глобина и β -глобина. $\beta 4$ тетрамеры, содержащие четыре β -глобиновые цепи, часто наблюдаются у пациентов только с одним функциональным α -глобиновым геном, состояние, известное как HbH. $\beta 4$ тетрамеры способны связывать кислород, но не высвобождать его в периферию, вызывая то, что известно как болезнь HbH. У людей с болезнью HbH присутствуют RBC с сокращенным периодом полувыведения, которые легко подвергаются гемолизу, что приводит к увеличению анемии. Потеря всех четырех α -глобиновых генов, как правило, приводит к смертельному исходу во время внутриутробного развития. Таким образом, существует потребность в дополнительных способах и композициях, которые можно использовать для редактирования генома с целью корректировки абберрантных генов или модификации экспрессии других, например, для лечения гемоглобинопатии, таких как серповидно-клеточная анемия и талассемия.

Сущность изобретения

В настоящем документе раскрыты способы и наборы для модификации экспрессии или корректировки одного или более генов, кодирующих белки, участвующих в генетическом заболевании (например, выработка белков, которые отсутствуют, являются дефицитными или абберрантными при болезни, и/или белков, которые регулируют эти белки), таком как серповидно-клеточная анемия или талассемия. Изменение этих белков может привести к лечению этих генетических заболеваний. В частности, геномное редактирование используют для корректировки абберрантного гена, вставки гена дикого типа или модификации экспрессии эндогенного гена. В качестве неограничивающего примера ген дикого типа, кодирующий β -глобин, может быть вставлен в клетку для выработки недостающего белка и/или лечения гемоглобинопатии, вызванной поврежденным β -глобином. В некоторых случаях ген дикого типа может быть вставлен в локус "безопасной гавани" или в локус, который, как известно, экспрессируется на высоком уровне в целевой ткани, такой как β -глобиновый локус в эритроидных клетках. Геномное редактирование подобным образом может быть использовано для выработки недостающего белка (и тем самым лечения) при α -талассемии, путем вставки α -глобинового гена дикого типа в локус "безопасной гавани". Другой подход предполагает использование коррекции гена, при которой воздействуют на поврежденный эндогенный α - или β -глобиновый ген, а мутантная последовательность заменяется. В качестве альтернативы регуляторный ген, участвующий в подавлении γ -глобина, может быть модифицирован или нокаутирован (например, для увеличения экспрессии γ -глобина путем инактивации и/или снижения количества подавляющего белка), и/или регуляторный сайт связывания, расположенный против хода транскрипции γ -глобина или в других зонах β -глобинового локуса, может быть модифицирован таким обра-

зом, что регуляторы не могут взаимодействовать должным образом у γ -глобинового локуса и вырабатывается HbF, тем самым отменяя эффекты (т.е. SCD или β -талассемию), вызванные aberrантным β -глобиновым геном. Один из подходов дополнительно предполагает использование модификации стволовых клеток (например, гемопоэтических стволовых клеток или предшественников RBC), после чего такие стволовые клетки могут быть использованы для пересадки пациенту с целью лечения гемоглобинопатии.

В одном аспекте в настоящем документе раскрыт белок "цинковые пальцы" (ZFP), который связывается с сайтом-мишенью в целевой области (например, β -глобиновый, α -глобиновый ген или ген "безопасной гавани", или регуляторный ген или его ДНК-мишень, такие как BCL11A, γ -глобин или KLF1) в геноме, причем ZFP содержит один или более связывающих доменов "цинковых пальцев". В одном варианте реализации изобретения ZFP является нуклеазой "цинковые пальцы" (ZFN), которая расщепляет целевую геномную область-мишень, причем ZFN содержит один или более сконструированных сайтов связывания "цинковых пальцев" и нуклеазный домен расщепления или полудомен расщепления. Домены расщепления или полудомены расщепления могут быть получены, например, от разных рестрикционных эндонуклеаз и/или "хоминг"-эндонуклеаз. В одном варианте реализации изобретения полудомены расщепления получают из Типа IIS рестрикционной эндонуклеазы (например, Fok I). В определенных вариантах реализации изобретения домен "цинковых пальцев" распознает сайт-мишень в глобиновом гене или гене "безопасной гавани". В определенных вариантах реализации изобретения домен "цинковых пальцев" содержит 5 или 6 доменов "цинковых пальцев" и распознает сайт-мишень на глобиновом гене (например, белок "цинковые пальцы", содержащий 5 или 6 пальцев с распознающими спиральными областями, показанными в табл. 1A). В другом варианте реализации домен "цинковых пальцев" распознает сайт-мишень в BCL11A, KLF1, α -, β - или γ -глобиновом гене или их регуляторные элементы. В определенных вариантах реализации изобретения домен "цинковых пальцев" содержит 5 или 6 доменов "цинковых пальцев" и распознает сайт-мишень в BCL11A, KLF1, α -, β - и γ -глобиновом гене или в их регуляторных элементах (например, белок "цинковые пальцы", содержащий 5 или 6 пальцев с распознающими спиральными областями, показанными в табл. 1A). В другом аспекте настоящего документа раскрыт белок TALE (подобный активатору транскрипции), который связывает сайт-мишень в целевой области (например, α - или β -глобиновый ген или ген "безопасной гавани", или регуляторный ген или его ДНК-мишень, такую как BCL11A, γ -глобин или KLF1) в геноме, причем TALE содержит один или более сконструированных TALE связывающих доменов. В одном варианте реализации изобретения TALE является нуклеазой (TALEN), которая расщепляет целевую геномную область-мишень, где TALEN содержит один или более сконструированных TALE ДНК-связывающих доменов и нуклеазный домен расщепления или полудомен расщепления. Домены расщепления и полудомены расщепления могут быть получены, например, от разных рестрикционных эндонуклеаз и/или "хоминг"-эндонуклеаз. В одном варианте реализации изобретения полудомены расщепления получены от Типа IIS рестрикционной эндонуклеазы (например, Fok I). В определенных вариантах реализации изобретения ДНК-связывающий домен TALE распознает сайт-мишень в глобиновом гене или гене "безопасной гавани". В других вариантах реализации ДНК-связывающий домен TALE распознает сайт-мишень в BCL11A, KLF1, α -, β - или γ -глобиновом гене или в их регуляторных элементах (например, белок TALEN, представленный в табл. 3).

В другом аспекте настоящего документа раскрыта система CRISPR/Cas, которая связывает сайт-мишень в целевой области (например, ген с высоким уровнем экспрессии, связанный с болезнью ген или ген "безопасной гавани") в геноме, где система CRISPR/Cas содержит нуклеазу CRISPR/Cas и сконструированную crPHK/tracrPHK (или единичную гидовую PHK). В определенных вариантах реализации изобретения система CRISPR/Cas распознает сайт-мишень в высоко экспрессионном гене, связанном с болезнью гене и гене "безопасной гавани". В определенных вариантах реализации изобретения система CRISPR/Cas распознает сайт-мишень в глобиновом, альбуминовом, CCR5, CXCR4, AAVS1, Rosa или HPRT гене. ZFN, TALEN и/или система CRISPR/Cas, как раскрыто в настоящем документе, могут связываться с и/или расщеплять целевую область в кодирующем или не кодирующем участке в пределах гена или смежных с ним областях, такую как, например, лидерная последовательность, трейлерная последовательность или интрон, или в пределах нетранскрибируемой области, расположенной против хода транскрипции или по ходу транскрипции кодирующей области. В определенных вариантах реализации изобретения ZFN, TALEN и/или система CRISPR/Cas связывается с и/или расщепляет глобиновый ген. В других вариантах реализации ZFN, TALEN и/или система CRISPR/Cas связывается с и/или расщепляет ген "безопасной гавани", например ген CCR5, ген CXCR4, ген PPP1R12C (также известный как AAVS1), альбуминовый ген или ген Rosa. См., например, патентные публикации США №№ 20080299580; 20080159996; 201000218264; 20110301073; 20130177983 и 20130177960 и предварительную заявку на патент США № 61/823689. Кроме того, с целью селекции может быть использован locus HPRT (см. патентную публикацию США № 20130122591). В другом аспекте настоящего документа раскрыты композиции, содержащие одну или более нуклеаз "цинковые пальцы" и/или TALE или систему CRISPR/Cas, как раскрыто в настоящем документе. В некоторых вариантах реализации изобретения ZFN, TALEN и/или система CRISPR/Cas связывается с и расщепляет BCL11A, KLF1, α -, β - или γ -глобиновый ген или

расщепляет их регуляторные элементы. В другом аспекте в настоящем документе раскрыты композиции, содержащие одну или более нуклеаз "цинковые пальцы", TALE или Cas, как раскрыто в настоящей заявке. В другом аспекте настоящего документа раскрыт полинуклеотид, кодирующий одну или более ZFN, TALEN и/или систему CRISPR/Cas, как раскрыто в настоящем документе. В некоторых аспектах мПНК может быть химически модифицированной (см., например, Kormann et al., (2011) *Nature Biotechnology* 29(2): 154-157).

В другом аспекте настоящего документа раскрыт вектор экспрессии ZFN, TALEN и/или системы CRISPR/Cas, содержащий полинуклеотид, кодирующий одну или более ZFN, TALEN и/или систему CRISPR/Cas, раскрытые в настоящем документе, функционально связанный с промотором. В одном варианте реализации изобретения вектором экспрессии является вирусный вектор.

В одном аспекте настоящего документа описан белок ZFN, TALEN и/или системы CRISPR/Cas, который используется для расщепления ДНК-мишени.

В других аспектах генетически модифицированные предшественники RBC (кроветворные стволовые клетки, известные как "HSC") вводятся в трансплантат костного мозга, а RBC дифференцируют и созревают *in vivo*. В некоторых вариантах реализации изобретения HSC выделяют после G-CSF-индуцированной мобилизации, а в других клетки выделяют из костного мозга человека или пуповины. В некоторых аспектах HSC редактируют с помощью обработки нуклеазой, разработанной для нокаутирования регулятора экспрессии глобина (например, BCL11A or KLF1). В других аспектах HSC модифицируют с помощью сконструированной нуклеазы, а донорскую нуклеиновую кислоту, такую как ген дикого типа (например, глобиновый ген), вставляют и экспрессируют и/или корректируют эндогенный абберрантный ген. В некоторых случаях генная последовательность дикого типа для вставки кодирует β -глобин дикого типа или α -глобин дикого типа. В других случаях эндогенный абберрантный ген является β -глобиновым или α -глобиновым геном. В некоторых вариантах реализации изобретения модифицированные HSC вводят пациенту после мягкого миелоаблативного кондиционирования. В других аспектах HSC вводят после полной миелоабляции, таким образом, что после приживления 100% кроветворных клеток происходят от модифицированных HSC.

В другом аспекте настоящего документа раскрыт способ расщепления эндогенного гена (например, гена, инактивация которого приводит к повышению экспрессии γ -глобина, такого как BCL11A или KLF1) в клетке-предшественнице RBC, включающий введение в клетку одного или более полинуклеотидов, кодирующих ZFN, TALEN и/или систему CRISPR/Cas, которые связывают сайт-мишень в одном или более эндогенных генах таким образом, что экспрессируется ZFN(ы), TALEN(ы) и/или система CRISPR/Cas, а один или более генов расщепляются. В другом аспекте настоящего документа раскрыт способ расщепления гена BCL11A или KLF1 в клетке, включающий введение в клетку одного или более полинуклеотидов, кодирующих ZFN, TALEN и/или системы CRISPR/Cas, которые связывает сайт-мишень в одном или более генах BCL11A или KLF1 таким образом, что экспрессируется ZFN(ы), TALEN(ы) и/или система CRISPR/Cas, а один или более генов BCL11A или KLF1 расщепляются. В определенных вариантах реализации изобретения домен "цинковых пальцев" содержит 5 или 6 "цинковых пальцев" и распознает сайт-мишень в глобиновом гене (например, белок "цинковые пальцы", содержащий 5-6 пальцев с распознающими спиральными областями, показанными в табл. 1A). В других вариантах реализации изобретения TALEN распознает сайт-мишень в β -глобиновой, α -глобиновой, γ -глобиновой, KLF или BCL11A последовательности (представленные в табл. 3). В еще других вариантах реализации изобретения система CRISPR/Cas распознает сайт-мишень в β -глобиновой, α -глобиновой, γ -глобиновой, KLF или BCL11A последовательности, в которой единичная гидовая РНК сконструирована для распознавания целевого сайта-мишени в представляющем интерес гене-мишени. Расщепленный ген(ы) может быть инактивированным (нокаутированным), например, нокаут одного или более генов, чьи продукты могут ингибировать экспрессию гена (например, глобинового гена), или содержать нарушение в регуляторном сайте-мишени на ДНК таких белков. В некоторых вариантах реализации изобретения инактивированный ген(ы) или его последовательности-мишени являются такими, которые участвуют в ингибировании экспрессии фетального гемоглобина. Клетки (например, стволовые клетки) после дифференцировки содержат фетальный гемоглобин и могут быть введены пациенту, который в этом нуждается. В некоторых вариантах реализации изобретения глобиновый ген является нокаутированным. Например, α -глобиновый ген может быть нокаутирован для восстановления соотношения α -глобина и β -глобина, если β -глобин слабо экспрессируется, или HbS, кодирующий β -глобиновый ген, может нокаутироваться одновременно с введением β -глобина дикого типа. Клетки (например, стволовые клетки) после дифференцировки будут содержать гемоглобин HbA и могут быть введены пациентам, нуждающимся в этом. В другом аспекте настоящего документа раскрыт способ вставки последовательности в эндогенный ген (например, β -глобиновый, α -глобиновый ген и/или ген "безопасной гавани") в клетке (например, стволовой клетке), включающий расщепление эндогенного гена с использованием одной или более нуклеаз и вставку последовательности в сайт расщепления. В определенных вариантах реализации изобретения геномная последовательность в любом гене-мишени замещается, например, с использованием пары ZFN или TALEN или системы CRISPR/Cas (или вектора, кодирующего указанные ZFN, TALEN

и/или систему CRISPR/Cas), как раскрыто в настоящем документе, "донорской" последовательностью (также известной как "трансген"), которая вставляется в ген после намеченного расщепления ZFN, TALEN и/или системой CRISPR/Cas. Донорская последовательность может присутствовать в векторе ZFN или TALEN, присутствовать в отдельном векторе (например, векторе Ad, AAV или LV) или в качестве альтернативы может вводиться в клетку с использованием различных механизмов доставки нуклеиновых кислот. Такая вставка донорской нуклеотидной последовательности в локус-мишень (например, глобиновый ген, другой ген "безопасной гавани" и т.п.) приводит к экспрессии трансгена под контролем генетических контрольных элементов локуса-мишени (например, глобина). В некоторых вариантах реализации изобретения трансген кодирует некодирующую РНК (например, кшРНК).

Экспрессия трансгена перед созреванием RBC может давать RBC, содержащие целевую некодирующую РНК.

В других вариантах реализации изобретения трансген содержит функциональный белок, например глобиновый белок (например, β дикого типа и/или γ дикого типа). В некоторых вариантах реализации изобретения вставка целевого трансгена в эндогенный ген (например, глобиновый ген), приводит к экспрессии интактной экзогенной белковой последовательности и отсутствию любой последовательности, кодируемой эндогенным геном. В других вариантах реализации изобретения экспрессированный эндогенный белок является слитым белком и содержит аминокислоты, кодируемые трансгеном и глобиновым геном (например, из эндогенного локуса-мишени или, в качестве альтернативы, из глобин-кодирующих последовательностей на трансгене). В некоторых случаях глобиновым геном является β -глобиновый ген, тогда как в других случаях глобиновым геном является α -глобиновый ген. В других случаях глобиновым геном является γ -глобиновый ген. Если они присутствуют, эндогенные глобиновые последовательности могут присутствовать на аминоконцевой (N)-концевой части экзогенного белка и/или карбокси (C)-концевой части экзогенного белка. Глобиновые последовательности могут включать непротрансформированные глобиновые последовательности дикого типа или мутантные глобиновые последовательности или в качестве альтернативы могут включать частичные кодирующие глобин последовательности. В некоторых вариантах реализации слияние глобин-трансген локализуется у эндогенного локуса в пределах клетки, тогда как в других вариантах реализации глобин-трансген кодирующая последовательность вставляется в "безопасную гавань" в пределах генома. В некоторых аспектах "безопасную гавань" выбирают из гена CCR5, гена CXCR4, гена PPP1R12C (также известного как AAVS1), альбуминового гена или гена Rosa. См., например, патентные публикации США №№ 20080299580; 20080159996; 201000218264; 20110301073; 20130177983 и 20130177960 и предварительную заявку на патент США № 61/823689. Кроме того, с целью селекции может быть использован локус HPRT (см. патентную публикацию США № 20130122591). В еще другом аспекте в настоящем документе предложены клеточные линии и/или трансгенные животные модели (системы). В некоторых вариантах реализации изобретения трансгенная клетка содержит и/или животное несет трансген, который кодирует человеческий ген. В некоторых случаях трансгенное животное содержит нокаут у эндогенного локуса, соответствующего экзогенному трансгену (например, мышинный глобиновый ген нокаутируется, человеческий глобиновый ген вставляется в мышинный), тем самым делая возможным развитие *in vivo* системы, в которой человеческий белок может быть изучен в отдельности. Такие трансгенные модели могут быть использованы с целью скрининга для идентификации небольших молекул или больших биомолекул или других объектов, которые могут взаимодействовать с или модифицировать представляющий интерес человеческий белок. В некоторых аспектах трансген интегрируется в выбранный локус (например, глобиновый или "безопасной гавани") в стволовой клетке (например, эмбриональной стволовой клетке, индуцированной плюрипотентной клетке, кроветворной стволовой клетке и т.п.) или эмбрион животного, полученный любым способом, раскрытым в настоящем документе, а затем эмбрион имплантируется таким образом, чтобы родилось живое животное. В других аспектах стволовые клетки содержат геномные альтерации у эндогенных локусов, таких как гены BCL11A, KLF1 или γ -глобина или их комбинации, таким образом, что повышается экспрессия γ -глобина. В некоторых вариантах реализации изобретения повышение экспрессии γ -глобина изменяет соотношение γ -глобина к β -глобину в клетке по сравнению с неотредактированной стволовой клеткой. Животное затем выращивают до половой зрелости и позволяют производить потомство, в котором по меньшей мере некоторые из потомков несут отредактированную эндогенную генную последовательность или интегрированный трансген.

В еще дополнительном аспекте настоящего документа предложен способ сайт-специфичной интеграции последовательности нуклеиновой кислоты в эндогенный локус (например, глобиновый или "безопасной гавани") хромосомы, например в хромосому эмбриона. В определенных вариантах реализации изобретения способ включает: (а) инъекцию эмбриона (i) по меньшей мере одним ДНК-вектором, где ДНК-вектор содержит "апстрим"-последовательность и "даунстрим"-последовательность, фланкирующие последовательность нуклеиновой кислоты, которая будет интегрироваться, и (ii) по меньшей мере одной молекулой РНК, кодирующей нуклеазу "цинковых пальцев", TALE или Cas9. В случае использования белка Cas9 также вводится sgРНК. Нуклеаза или нуклеазная система распознает сайт-мишень в локусе-мишени (например, глобиновом локусе или локусе "безопасной гавани"), а затем (б) культивиру-

ют эмбрион, чтобы сделать возможной экспрессию нуклеазы "цинковые пальцы" или TALE, или системы CRISPR/Cas, причем двухцепочечный разрыв, введенный в мишень с помощью нуклеазы "цинковые пальцы", TALEN или системы CRISPR/Cas, затем восстанавливают путем гомологичной рекомбинации с ДНК-вектором, с тем, чтобы интегрировать последовательность нуклеиновой кислоты в хромосому. В любом из способов, раскрытых в настоящем документе, полинуклеотид, кодирующий нуклеазу(ы) "цинковые пальцы", TALEN(ы) и/или систему CRISPR/Cas, может содержать ДНК, РНК или их комбинации. В определенных вариантах реализации полинуклеотид содержит плазмиду. В других вариантах реализации полинуклеотид, кодирующий нуклеазу, содержит мРНК.

Также предложен набор, содержащий ZFP, TALEN и/или систему CRISPR/Cas согласно изобретению. Набор может содержать нуклеиновые кислоты, кодирующие ZFP, TALEN и/или систему CRISPR/Cas (например, молекулы РНК или гены, кодирующие ZFP, TALEN или Cas9, содержащиеся в пригодных векторах экспрессии), и сконструированную sgРНК, если необходимо, или аликваты нуклеазных белков, молекул-доноров, подходящих линий клеток-хозяев, инструкции по выполнению способов согласно изобретению и т.п.

Эти и другие аспекты будут очевидны специалисту в данной области в свете раскрытия в целом.

Краткое описание графических материалов

На фиг. 1 проиллюстрировано выравнивание геновой последовательности β -подобного глобина в области, окружающей мутацию серповидно-клеточной анемии (указана на фигуре). Проиллюстрированы (от верхней строки до нижней строки) последовательности β -гемоглобина с серповидно-клеточной мутацией (HBB-серп, SEQ ID NO: 1); гемоглобина β (HBB, SEQ ID NO: 2); гемоглобина δ (HBD, SEQ ID NO: 3); псевдогена β -гемоглобина (HBBP1, SEQ ID NO: 4); гемоглобина ϵ (HBE1, SEQ ID NO: 5), гемоглобина γ -1 (HBG1, SEQ ID NO: 6) и гемоглобина γ -2 (HBG2, SEQ ID NO: 7). Результаты анализа активности Cel I приведены ниже выравнивания для указанных пяти пар ZFN.

На фиг. 2 панели А и В, представлены гели, иллюстрирующие вставку последовательности, указанной β -глобиновым донором в клетках CD34+. На фиг. 2А (RFLP) проиллюстрирована вставка последовательности, указанной β -глобиновым донором, где вставка проверяется присутствием нового сайта рестрикции, присутствующем на донорской ДНК. На фиг. 2В проиллюстрированы результаты анализа несоответствия Cel-1 (Surveyor™, Transgenomic), подтверждающие наличие новых последовательностей, которые создают несоответствие. Процент аллелей, несущих мутацию (% NHEJ), указан в тексте справа от гелей (первый столбец, соответствующий номеру полосы на геле). Цифры относятся к комбинациям ZFN; unt: нетрансфицированный контроль.

Фиг. 3 представляет собой графическую иллюстрацию роли, которую KLF1 и BCL11A играют в регуляции экспрессии генов β - и γ -глобина. Экспрессия KLF1 стимулирует экспрессию как генов BCL11a, так и β -глобина. Белок BCL11A подавляет экспрессию γ -глобина.

На фиг. 4, панели А-Е, проиллюстрированы гели, демонстрирующие результаты анализа Cel 1, как раскрыто выше, после обработки HSC указанными BCL11A- (фиг. 4А), KLF1- (фиг. 4С и 4D) или HPRT- (фиг. 4В) специфичными ZFN и обработки трансдуцированных клеток кратковременным гипотермическим шоком (30°) или при стандартных условиях (37°). ДНК получали через 3 дня после трансфекции. На фиг. 4Е представлен тот же самый тип анализа Cel 1, проведенный с образцами, полученными через 3 дня после трансфекции HSC или через 17 дней после эритроидной дифференцировки. Процент аллелей, несущих мутацию (% NHEJ), указывается внизу на дорожках, а идентичность использованных пар ZFN указывается на каждой фигуре.

На фиг. 5, панели А и В, проиллюстрирована экспрессия γ -глобина по сравнению с β -глобином (фиг. 5А) или мРНК γ -глобина, скорректированного со стандартом 18s РНК (фиг. 5В) на 7 или 17 день после дифференцировки, по данным анализа с применением методики TaqMan®. Процент мРНК γ -глобина по сравнению с мРНК γ + β -глобина приведен выше каждой полосы на фиг. 5А. На фиг. 5В относительный уровень γ -глобина, который нормирован по 18s РНК, приведен выше полос и демонстрирует, что уровень мРНК γ -глобина относительно 18S выше в клетках, которые были обработаны с помощью BCL11A-специфичных ZFN.

На фиг. 6 проиллюстрировано количество мРНК γ -глобина в метилцеллюлозных колониях, полученных от HSC, в зависимости от генотипа клеток. Клетки, в которых оба гена BCL11A принадлежат к дикому типу ("BB"), вырабатывают наименьшее количество мРНК γ -глобина по сравнению с клетками, которые содержали одну BCL11A-нокаутную аллель ("Bb") или содержали обе нокаутные аллели ("нокаут"). Цифры над полосами указывают процент γ -глобина, полученный из суммарного β -глобина.

На фиг. 7 показана серия последовательностей ДНК (SEQ ID NO: 140-148) в области, расположенной против хода транскрипции гена γ -глобина, после обработки γ -глобин-специфичными ZFN в клетках K562. Последовательности содержат ряд вставок и делеций, включая делецию 13 п.о. ("Δ13 п.о."), которые идентичны одному из человеческих генотипов, ассоциированных с HPFH. "Стандартная" последовательность сверху является последовательностью 5'-регуляторной области дикого типа для γ -глобина. Сайты связывания пары ZFN выделены красным цветом, делеция 13 п.о. природного происхождения

подчеркнута.

На фиг. 8, панели от А до С, проиллюстрирован Taqman-анализ эритроидных колоний, полученных от HSC, обработанных ZFN, нацеливающих γ -глобиновый промотор, после чего следует осаждение на метилцеллюлозных колониях. Количество колоний указано в нижней части каждой полосы, как и генотип. На фиг. 8А проиллюстрированы относительные соотношения мРНК γ/β -глобина; на фиг. 8В проиллюстрированы уровни мРНК γ -глобина, скорректированные по уровням 18s РНК, а на фиг. 8С проиллюстрирован соответствующий анализ уровней β -глобина, скорректированный по 18s. Сравнение средних значений соотношений для дикого типа и мутантных колоний указывает на то, что уровни γ -глобина в колониях с ZFN-индуцированными мутациями в γ -глобиновом промоторе повышены.

На фиг. 9 проиллюстрирована промоторная область γ -глобинового гена (SEQ ID NO: 149-152). Две γ -глобиновые аллели выровнены (HBG1 и HBG2). Различия в последовательностях двух аллелей обозначены серыми клетками. Кроме того, мутации, которые связаны с HPFH, обозначены черным контуром. Стартовый АТГ обозначен, как и границы экзона 1. Повышение уровней фетального глобина, ассоциированное с каждой мутацией, указано числом над ним.

На фиг. 10, панели А и В, проиллюстрировано количество NHEJ (т.е. нарушение намеченного локуса, что является следствием NHEJ-базирующегося восстановительного события для направляемого нуклеазой разрыва) и коррекция гена, обнаруженная для β -глобинового гена в клетках CD34+ с использованием указанной нуклеазы "цинковые пальцы" и олигонуклеотидных доноров.

На фиг. 11 проиллюстрировано количество NHEJ и намеченная интеграция донорского нуклеотида в клетки CD34+, причем гомологические плечи на доноре разные.

На фиг. 12 проиллюстрирована устойчивость генного редактирования в эритроидных производных стволовых клеток, которые были обработаны ZFN и олигонуклеотидным донором. Генную модификацию анализировали для четырех типов клеточных популяций, возникших вследствие дифференцировки, эритроидных колониеобразующих единиц ("CFU-E"), эритроидных взрыв-образующих единиц ("BFU-E"), колониеобразующих единиц гранулоцитов/макрофагов ("CFU-GM") и колониеобразующих единиц гранулоцитов/эритроцитов/моноцитов/ макрофагов ("CFU-GEMM").

На фиг. 13 проиллюстрирована стабильность генной модификации гена β -глобина с течением времени.

Подробное описание сущности изобретения

В настоящем документе раскрыты способы и композиции для изучения и лечения генетического заболевания, такого как гемоглобинопатия. В изобретении раскрыто геномное "редактирование" клетки-мишени, таким образом, что присутствует благоприятная модификация экспрессии одного или более глобиновых генов, которая в свою очередь приводит к лечению гемоглобинопатий, таких как серповидно-клеточная анемия или талассемия, у субъекта, нуждающегося в этом. Благоприятные модификации экспрессии глобинового гена включают обеспечение гена γ -глобина у субъекта с аберрантным β -глобином; и/или коррекцию последовательности аберрантного гена α - или β -глобина, но не ограничиваются ими. Кроме того, доставка модифицированных кроветворных стволовых клеток в трансплантат, модифицированный с целью экспрессии желательного белкового продукта, может быть полезной при лечении гемоглобинопатий, таких как серповидно-клеточная анемия или талассемия. Также раскрыты клеточные линии и животные с модифицированной экспрессией глобина.

Таким образом, способы и композиции согласно изобретению могут быть использованы для модификации экспрессии одного или более глобиновых генов (например, γ , α и/или β) в клетке (например, эритроидной клетке-предшественнике). Эти способы и композиции могут быть использованы для разрушения генов, участвующих в γ -глобиновой репрессии (например, BCL11A или KLF1), таким образом, что после редактирования клетки будут экспрессировать более высокие уровни γ -глобина, и может быть получен HbF. Кроме того, редактирование может быть использовано для разрушения сайта связывания на гене (например, разрушение связывания BCL11A в β -глобиновом локусе, разрушение связывания репрессора транскрипции γ -глобина у γ -глобинового промотора) для отключения репрессии гена. В качестве альтернативы или в дополнение к этим модификациям способы и композиции могут быть использованы для корректировки аберрантного эндогенного α - и/или β -глобинового гена или вставки гена дикого типа в заданное место локализации в геноме клетки (например, в HSC). Клетки-предшественники можно получить от нуждающихся субъектов, модифицировать *ex vivo*, а затем вернуть обратно в организм субъекта в трансплантате костного мозга.

Общая часть.

В практике способов, а также получения и применения композиций, раскрытых в настоящем документе, применяются, если не указано иное, обычные методы молекулярной биологии, биохимии, хроматиновой структуры и анализа, вычислительной химии, культуры клеток, рекомбинантных ДНК и смежных областей, известные специалистам в данной области техники. Эти методы полностью раскрыты в литературе. См., например, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Second edition, Cold Spring Harbor Laboratory Press, 1989 and Third edition, 2001; Ausubel et al., CURRENT

PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 и периодические обновления; серии METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Third edition, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman and A.P. Wolffe, eds.), Academic Press, San Diego, 1999; и METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Определения.

Термины "нуклеиновая кислота", "полинуклеотид" и "олигонуклеотид" используются взаимозаменяемо и относятся к дезоксирибонуклеотиду или рибонуклеотидному полимеру, линейной или кольцевой конформации, и одно- или двухцепочечной форме. Для целей настоящего изобретения эти термины не следует рассматривать как ограничивающие по отношению к длине полимера. Термины могут включать известные аналоги природных нуклеотидов, а также нуклеотиды, которые модифицированы в основном, сахарном и/или фосфатном фрагментах (например, фосфоротиоатные скелеты). В общем, аналог определенного нуклеотида обладает такой же специфичностью спаривания оснований; т.е. аналог А будет спариваться с Т.

Термины "полипептид", "пептид" или "белок" используются взаимозаменяемо для обозначения полимера аминокислотных остатков. Термин также применяется к аминокислотным полимерам, в которых одна или более аминокислот являются химическими аналогами или модифицированными производными соответствующих аминокислот природного происхождения. "Связывание" относится к специфичному относительно последовательности, нековалентному взаимодействию между макромолекулами (например, между белком и нуклеиновой кислотой). Не все компоненты связывающего взаимодействия должны быть специфичными относительно последовательности (например, контакты с фосфатными остатками в скелете ДНК), поскольку взаимодействие в целом является специфичным относительно последовательности. Такие взаимодействия, как правило, характеризуются константой диссоциации (K_d) 10^{-6} М⁻¹ или ниже. "Аффинность" относится к силе связывания: повышенная аффинность связывания коррелирует со сниженной K_d .

"Связывающий белок" представляет собой белок, который способен к нековалентному связыванию с другой молекулой. Связывающий белок может связываться, например, с молекулой ДНК (ДНК-связывающий белок), молекулой РНК (РНК-связывающий белок) и/или белковой молекулой (белок-связывающий белок). В случае белок-связывающего белка он может связываться сам с собой (образуется гомодимер, гомотример, и т.д.) и/или он может связываться с одной или более молекулами другого белка или белков. Связывающий белок может обладать более чем одним типом связывающей активности. Например, белки "цинковые пальцы" обладают ДНК-связывающей, РНК-связывающей и белок-связывающей активностью.

"Белок цинковых пальцев, связывающий ДНК" (или связывающий домен) является белком или доменом в пределах большего белка, который связывает ДНК способом, специфичным относительно последовательности, посредством одного или более цинковых пальцев, которые являются областями аминокислотной последовательности, в пределах домена связывания, структура которого стабилизируется благодаря координации цинкового иона. Термин белок цинковых пальцев, связывающий ДНК, часто сокращают как белок "цинковые пальцы" или ZFP.

"ДНК-связывающий домен TALE" или "TALE" является полипептидом, содержащим один или более TALE повторных доменов/единиц. Повторные домены вовлечены в связывание TALE с его когнантной ДНК-последовательностью-мишенью. Одна "повторная единица" (также именуемая "повтором") обычно содержит 33-35 аминокислот и демонстрирует по меньшей мере некоторую гомологию последовательности с другими последовательностями TALE-повторов в пределах белка TALE природного происхождения. См., например, патентную публикацию США № 20110301073.

Связывающие домены "цинковых пальцев" и TALE могут быть "сконструированными" для связывания predetermined нуклеотидной последовательности, например, путем конструирования (модифицируя одну или более аминокислот) спиральной области распознавания белка "цинковые пальцы" или TALE природного происхождения. Поэтому сконструированные ДНК-связывающие белки ("цинковые пальцы" или TALE) представляют собой белки, которые не встречаются в природе. Неограничивающими примерами способов конструирования ДНК-связывающих белков являются разработка и селекция. Разработанный ДНК-связывающий белок представляет собой белок, который не встречается в природе и чья конструкция/состав является следствием, главным образом, рациональных критериев. Рациональные критерии для конструирования включают применение правил замещения и компьютерных алгоритмов для обработки информации в базе данных, хранящей информацию о существующих конструкциях ZFP и/или TALE и данных связывания. См., например, патенты США 6140081; 6453242 и 6534261; см. также WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 и WO 03/016496 и патентную публикацию США 20110301073. "Отобранный" белок "цинковые пальцы" или TALE представляет собой белок, не встречающийся в природе, выработка которого является следствием, главным образом, эмпирического процесса, такого как фаговый дисплей, интерактивная ловушка или гибридная селекция. См., например, US 5789538; US 5925523; US 6007988; US 6 013453; US 6200759; WO 95/19431; WO 96/06166;

WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084 и патентная публикация США 20110301073. "Рекомбинация" относится к процессу обмена генетической информацией между двумя полинуклеотидами. Для целей настоящего документа "гомологичная рекомбинация" (HR) относится к специализированной форме такого обмена, который происходит, например, во время восстановления двухцепочечных разрывов в клетке с помощью гомологично направленных восстановительных механизмов. Этот процесс требует гомологии нуклеотидных последовательностей, использования "донорской" молекулы для восстановительной матрицы молекулы-"мишени" (т.е. молекулы, испытавшей двухцепочечный разрыв), и известен как "некротингверная конверсия генов" или "краткопутевая конверсия генов", так как он приводит к передаче генетической информации от донора к мишени. Без связи с какой-либо конкретной теорией, такая передача может включать коррекцию несоответствия гетеродуплексной ДНК, которая образуется между разорванной мишенью и донором, и/или "синтез-зависимый отжиг нити", в котором донор используется для ресинтеза генетической информации, которая станет частью мишени, и/или родственные процессы. Такая специализированная HR часто приводит к модификации последовательности молекулы-мишени, таким образом, что часть или вся последовательность донорного полинуклеотида инкорпорируется в полинуклеотид-мишень. В способах согласно настоящему раскрытию одна или более нацеленных нуклеаз, как раскрыто в настоящем документе, создают двухцепочечный разрыв в последовательности-мишени (например, клеточном хроматине) в заданном сайте, а "донорский" полинуклеотид, обладающий гомологией с нуклеотидной последовательностью в области разрыва, может быть введен в клетку. Наличие двухцепочечного разрыва, как было показано, облегчает интеграцию донорской последовательности. Донорская последовательность может быть физически интегрирована или, в качестве альтернативы, донорский полинуклеотид используется в качестве матрицы для восстановления разрыва посредством гомологичной рекомбинации, в результате чего вся или часть нуклеотидной последовательности вводится в качестве донора в клеточный хроматин. Таким образом, первая последовательность в клеточном хроматине может изменяться и, в некоторых вариантах реализации, может превращаться в последовательность, присутствующую в донорском полинуклеотиде. Таким образом, использование терминов "заменить" или "замена" можно понимать как представление замены одной нуклеотидной последовательности другой (т.е. замена последовательности в информационном смысле), и не обязательно требует физической или химической замены одного полинуклеотида другим. В любом из способов, раскрытых в настоящем документе, дополнительные пары белков "цинковые пальцы" или TALEN могут быть использованы для дополнительного двухцепочечного расщепления дополнительных сайтов-мишеней в клетке. Кроме того, система CRISPR/Cas может быть аналогичным образом использована для индуцирования дополнительных двухцепочечных разрывов.

В некоторых вариантах реализации способов направленной рекомбинации, и/или замены, и/или модификации последовательности в целевой области в клеточном хроматине, хромосомная последовательность изменяется путем гомологичной рекомбинации с экзогенной "донорской" нуклеотидной последовательностью. Такая гомологичная рекомбинация стимулируется наличием двухцепочечного разрыва в клеточном хроматине, если присутствуют последовательности, гомологичные области разрыва. В любом из способов, раскрытых в настоящем документе, экзогенная нуклеотидная последовательность ("донорская последовательность" или "трансген") может содержать последовательности, которые гомологичны, но не идентичны геномным последовательностям в целевой области, тем самым стимулируя гомологичную рекомбинацию для вставки неидентичной последовательности в целевую область. Таким образом, в некоторых вариантах реализации участки последовательности донора, которые гомологичны последовательности в целевой области, демонстрируют идентичность приблизительно от 80 до 99% (или любое целое число между ними) с геномной последовательностью, которая заменяется. В других вариантах реализации гомология между донорской и геномной последовательностью составляет более 99%, например, если только один нуклеотид отличается между донорской и геномной последовательностями, состоящими из более чем 100 смежных пар оснований. В некоторых случаях негомологичная часть донорской последовательности может содержать последовательности, отсутствующие в целевой области, таким образом, что новые последовательности вводятся в целевую область. В этих случаях негомологичная последовательность, как правило, фланкирована последовательностями из 50-1000 пар оснований (или любого целочисленного значения между ними) или любого количества пар оснований, превышающего 1000, которые гомологичны или идентичны последовательности в целевой области. В других вариантах реализации донорская последовательность негомологична первой последовательности и вставляется в геном с помощью механизмов негомологичной рекомбинации. Любой из способов, раскрытых в настоящем документе, может быть использован для частичной или полной инактивации одной или более последовательностей-мишеней в клетке путем направленной интеграции донорской последовательности, которая нарушает экспрессию гена(ов), представляющего интерес. Также предложены клеточные линии с частично или полностью инактивированными генами.

Кроме того, способы направленной интеграции, как раскрыто в настоящем документе, также могут быть использованы для интеграции одной или нескольких экзогенных последовательностей. Экзогенная последовательность нуклеиновой кислоты может содержать, например, один или более генов или молекул кДНК или любой тип кодирующей или некодирующей последовательности, а также один или не-

сколько элементов управления (например, промоторы). Кроме того, экзогенная последовательность нуклеиновой кислоты может произвести одну или несколько молекул РНК (например, короткие шпилечные РНК (кшРНК), ингибирующие РНК (РНКис), микроРНК (миРНК), и т.д.).

"Расщепление" относится к разрыву ковалентного скелета молекулы ДНК. Расщепление может быть инициировано различными способами, включая ферментативный или химический гидролиз фосфодиэфирной связи, но не ограничиваясь им. Возможны как одноцепочечное расщепление, так и двухцепочечное расщепления, а двухцепочечное расщепление может происходить в результате двух различных событий одноцепочечного расщепления. Расщепление ДНК может привести к получению тупых концов или ступенчатых концов. В некоторых вариантах реализации слитые полипептиды используются для адресного двухцепочечного расщепления ДНК.

"Полудомен расщепления" представляет собой полипептидную последовательность, которая в сочетании со вторым полипептидом (одинаковым или отличающимся) образует комплекс, обладающий активностью расщепления (предпочтительно активностью двухцепочечного расщепления). Термины "первый и второй полудомены расщепления", "+ и - полудомены расщепления" и "правый и левый полудомены расщепления" используются как взаимозаменяемые для обозначения пар полудоменов расщепления, которые димеризуются.

"Сконструированный полудомен расщепления" является полудоменом расщепления, модифицированным таким образом, чтобы образовывать облигатные гетеродимеры с другим полудоменом расщепления (например, с другим сконструированным полудоменом расщепления). См. также патентные публикации США №№ 2005/0064474, 2007/0218528, 2008/0131962 и 2011/0201055, включенные в настоящий документ посредством ссылки в полном объеме.

Термин "последовательность" относится к нуклеотидной последовательности любой длины, которая может быть ДНК или РНК, может быть линейной, кольцевой или разветвленной и может быть одноцепочечной или двухцепочечной. Термин "донорская последовательность" относится к нуклеотидной последовательности, которая вставлена в геном. Донорская последовательность может быть любой длины, например от 2 до 10000 нуклеотидов (или любая целочисленная величина между ними или более), предпочтительно от 100 до 1000 нуклеотидов (или любая целочисленная величина между ними), более предпочтительно от приблизительно 200 до 500 нуклеотидов.

"Патогенный ген" является геном, каким-либо образом поврежденным при моногенном заболевании. Не ограничивающие примеры моногенных заболеваний включают тяжелый комбинированный иммунодефицит, кистозный фиброз, лизосомальные болезни накопления (например, болезнь Гоше, Херлера, Хантера, Фабри, Неймана-Пика, Тай-Саха и т.д.), серповидно-клеточную анемию и талассемию.

"Хроматин" представляет собой нуклеопротеиновую структуру, содержащую клеточный геном. Клеточный хроматин содержит нуклеиновую кислоту, в первую очередь ДНК, и белок, включая гистоны и негистоновые хромосомные белки. Большинство эукариотического клеточного хроматина существует в виде нуклеосом, в которых нуклеосомное ядро содержит приблизительно 150 пар оснований ДНК, связанных с октамером, содержащим два гистона H2A, H2B, H3 и H4; а линкерная ДНК (переменной длины, в зависимости от организма) простирается между нуклеосомными ядрами. Молекула гистона H1, как правило, связана с линкерной ДНК. Для целей настоящего изобретения термин "хроматин" предназначен охватывать все типы клеточного нуклеопротеина, как прокариотических, так и эукариотических. Клеточный хроматин включает как хромосомный, так и эписомный хроматин. "Хромосома" представляет собой хроматиновый комплекс, включающий весь или часть генома клетки. Геном клетки часто характеризуется кариотипом, который представляет собой совокупность всех хромосом, составляющих геном клетки. Геном клетки может содержать одну или несколько хромосом.

"Эписома" является реплицирующейся нуклеиновой кислотой, нуклеопротеиновым комплексом или другой структурой, содержащей нуклеиновую кислоту, которая не является частью хромосомного кариотипа клетки. Примеры эписом включают плазмиды и некоторые вирусные геномы. "Сайт-мишень" или "последовательность-мишень" представляет собой последовательность нуклеиновой кислоты, которая определяет часть нуклеиновой кислоты, с которой связывающаяся молекула будет связываться, и обеспечивает условия, достаточные для существования связывания. "Экзогенная" молекула представляет собой молекулу, которая обычно не присутствует в клетке, но может быть введена в клетку с помощью одного или нескольких генетических, биохимических или других методов. "Нормальное присутствие в клетке" определяется относительно конкретной стадии развития и окружающих условий клетки. Так, например, молекула, которая присутствует только во время эмбрионального развития мышцы, является экзогенной молекулой по отношению к взрослой мышечной клетке. Аналогично, молекула, индуцированная тепловым шоком, является экзогенной молекулой по отношению к клетке, не испытывающей теплового шока. Экзогенная молекула может содержать, например, функционирующую версию неправильно функционирующей эндогенной молекулы или неправильно функционирующую версию нормально функционирующей эндогенной молекулы.

Экзогенная молекула может быть, в частности, небольшой молекулой, например, генерируемой с помощью процесса комбинаторной химии, или макромолекулой, такой как белок, нуклеиновая кислота, углевод, липид, гликопротеин, липопротеин, полисахарид, любой модифицированной производной вы-

шеуказанных молекул или любым комплексом, включающим одну или более вышеуказанных молекул. Нуклеиновые кислоты включают ДНК и РНК, могут быть одно- или двухцепочечными; могут быть линейными, разветвленными или кольцевыми; и могут быть любой длины. Нуклеиновые кислоты включают кислоты, способные образовывать дуплексы, а также триплексообразующие нуклеиновые кислоты. См., например, патенты США №№ 5176996 и 5422251. Белки включают ДНК-связывающие белки, факторы транскрипции, факторы реконструкции хроматина, метилированные ДНК-связывающие белки, полимеразы, метилазы, деметилазы, ацетилазы, деацетилазы, киназы, фосфатазы, интегразы, рекомбиназы, лигазы, топоизомеразы, гиразы и геликазы, но не ограничиваются ими. Экзогенная молекула может быть того же типа молекулой, что и эндогенная молекула, например экзогенным белком или нуклеиновой кислотой. Например, экзогенная нуклеиновая кислота может содержать инфицирующий вирусный геном, плазмиду или эписому, вводимые в клетку, или хромосому, которые обычно не присутствуют в клетке. Способы введения экзогенных молекул в клетки известны специалистам в данной области и включают опосредованный липидами перенос (например, липосомы, включая нейтральные и катионные липиды), электропорацию, непосредственную инъекцию, слияние клеток, бомбардировку частицами, фосфатно-кальциевое соосаждение, диэтиламиноэтил (DEAE)-декстран-опосредованный перенос и опосредованный вирусным вектором перенос, но не ограничиваются ими. Кроме того, экзогенная молекула может быть молекулой того же типа, что и эндогенная молекула, но полученной из другого вида, чем клетка. Например, последовательность нуклеиновой кислоты человека может быть введена в линии клеток, первоначально полученных от мыши или хомяка. Наоборот, "эндогенная" молекула является молекулой, которая обычно присутствует в определенной клетке на определенном этапе развития в конкретных условиях окружающей среды. Например, эндогенная нуклеиновая кислота может содержать хромосому, геном митохондрии, хлоропласта или другой органеллы или быть эписомальной нуклеиновой кислотой природного происхождения. Дополнительные эндогенные молекулы могут включать белки, например, факторы транскрипции и ферменты.

"Слитая" молекула представляет собой молекулу, в которой соединены две или более субъединичных молекул, предпочтительно ковалентно. Молекулярные субъединицы могут быть одинакового химического типа или могут быть молекулами разных химических типов. Примеры первого типа слитой молекулы включают слитые белки (например, слияние между ZFP или TALE ДНК-связывающим доменом и одним или более доменов активации) и слитые нуклеиновые кислоты (например, нуклеиновую кислоту, кодирующую слитый белок, раскрытый выше), но не ограничиваются ими. Примеры второго типа слитой молекулы включают слияние между триплексообразующей нуклеиновой кислотой и полипептидом и слияние между белком, связывающимся с малой бороздой, и нуклеиновой кислотой, но не ограничиваются ими.

Экспрессия слитого белка в клетке может быть результатом доставки слитого белка в клетку или осуществляться путем доставки полинуклеотида, кодирующего слитый белок, в клетку, где полинуклеотид транскрибируется, а транскрипт транслируется, для получения слитого белка. Транс-сплайсинг, полипептидное расщепление и лигирование полипептида могут также вовлекаться в экспрессию белка в клетке. Методы доставки полинуклеотидов и полипептидов в клетки представлены в другом разделе настоящего документа.

"Ген" для целей настоящего изобретения включает область ДНК, кодирующую генный продукт (см. ниже), а также все участки ДНК, которые регулируют выработку генного продукта вне зависимости от того, являются ли такие регуляторные последовательности смежными с кодирующей последовательностью и/или транскрибирующейся последовательностью. Соответственно, ген включает, но не обязательно ограничивается ими, промоторные последовательности, терминаторы, последовательности, регулирующие трансляцию, такие как сайты связывания рибосом и сайты внутренней посадки рибосом, энхансеры, сайленсеры, инсуляторы, краевые элементы, точки начала репликации, сайты прикрепления матрицы и локусные контрольные участки.

"Экспрессия гена" относится к превращению информации, содержащейся в гене, в генный продукт. Генный продукт может быть прямым транскрипционным продуктом гена (например, мРНК, тРНК, рРНК, антисмысловой РНК, рибозимом, структурной РНК или любым другим типом РНК) или белком, полученным путем трансляции мРНК. Генные продукты также включают РНК, модифицированные с помощью таких процессов, как экзпиривание, полиаденилирование, метилирование и редактирование, и белки, модифицированные, например, метилированием, ацетилированием, фосфорилированием, убиквитинированием, аденозиндиосфат(ADP)-рибозилированием, миристилированием и гликозилированием. "Модуляция" экспрессии гена, относится к модификации активности гена. Модуляция экспрессии может включать активацию генов и репрессию генов, но не ограничиваться ими. Редактирование генома (например, расщепление, модификация, инактивация, случайная мутация) может быть использовано для модулирования экспрессии. Инактивация гена относится к любому снижению экспрессии гена в сравнении с клеткой, которая не содержит ZFP или TALEN, как раскрыто в настоящем документе. Таким образом, инактивация гена может быть частичной или полной.

"Целевая область" представляет собой любую область клеточного хроматина, такую как, например, ген или некодирующая последовательность в пределах или в непосредственной близости к гену, в кото-

рой желательным является связывание экзогенной молекулы. Связывание может осуществляться для целей намеченного расщепления ДНК и/или намеченной рекомбинации. Целевая область может присутствовать в хромосоме, эписоме, оргanelльном геноме (например, митохондрих, хлоропластов) или инфицирующем вирусном геноме, например. Целевая область может находиться в пределах кодирующей области гена, транскрибируемых некодирующих областей, таких как, например, лидерные последовательности, трейлерные последовательности или интроны, или в пределах нетранскрибируемых областей по ходу транскрипции или против хода транскрипции кодирующей области. Целевая область может быть небольшой, например, одна пара нуклеотидов, или ее длина может составлять до 2000 пар нуклеотидов, или любое целочисленное значение нуклеотидных пар.

"Эукариотические" клетки включают клетки грибов (например, дрожжей), растительные клетки, клетки животных, клетки млекопитающих и клетки человека (например, Т-клетки), но не ограничиваются ими. "Красные кровяные клетки" (RBC), или эритроциты, представляют собой терминально дифференцированные клетки, полученные из кроветворных стволовых клеток. Они не содержат нуклеазы и большинства клеточных органелл. Эритроциты содержат гемоглобин, чтобы переносить кислород от легких к периферическим тканям. В самом деле, 33% индивидуальных RBC составляет гемоглобин. Они также переносят CO₂, вырабатываемый клетками в процессе метаболизма, из тканей и обратно в легкие для высвобождения во время выдоха. Эритроциты образуются в костном мозге в ответ на гипоксию крови, которая опосредуется путем высвобождения эритропоэтина (EPO) в почках. EPO приводит к увеличению количества проэритробластов и сокращает время, необходимое для полного созревания эритроцитов. Через приблизительно 120 дней, поскольку RBC не содержит ядра или любых других средств восстановления, клетки удаляются из кровотока посредством фагоцитарной активности макрофагов в печени, селезенке и лимфатических узлах (~90%) или гемолиза в плазме (~10%). После поглощения макрофагами химические компоненты RBC разрушаются в вакуолях макрофагов под действием лизосомальных ферментов. "Секреторные ткани" являются тканями у животного, которые секретируют продукты из отдельной клетки в просвет какого-либо типа, которые обычно получают из эпителия. Примеры секреторных тканей, локализованных в желудочно-кишечном тракте, включают клетки, которые выстилают кишечник, поджелудочную железу и желчный пузырь. Другие секреторные ткани включают печень, ткани, связанные с глазом и слизистыми оболочками, например, слюнных желез, молочных желез, предстательной железы, гипофиза и других членов эндокринной системы. Кроме того, секреторные ткани включают отдельные клетки любого типа ткани, которые способны к секреции.

Термины "функциональное соединение" и "функционально связанный" (или "функционально соединенный") используются взаимозаменяемо со ссылкой на сопоставление двух или более компонентов (например, элементов последовательности), в которых компоненты расположены таким образом, что оба компонента функционируют нормально и допускают возможность того, что по меньшей мере один из компонентов может опосредовать функцию, которая влияет по меньшей мере на один из других компонентов. В качестве иллюстрации последовательность, регулирующая транскрипцию, такая как промотор, функционально связана с кодирующей последовательностью, если последовательность, регулирующая транскрипцию, контролирует уровень транскрипции кодирующей последовательности в ответ на присутствие или отсутствие одного или более регулирующих транскрипцию факторов. Регулирующая транскрипцию последовательность обычно функционально связана *in cis* с кодирующей последовательностью, но не обязательно должна находиться в непосредственной близости от нее. Например, энхансер является регуляторной последовательностью транскрипции, которая функционально связана с кодирующей последовательностью, даже если они не являются смежными.

Что касается слитых полипептидов, термин "функционально связанный" может относиться к тому факту, что каждый из компонентов выполняет такую же функцию в связи с другим компонентом, как если бы они не были связаны между собой. Например, относительно слитого полипептида, в котором ZFP, TALE или Cas ДНК-связывающий домен слит с доменом активации, ZFP, TALE или Cas ДНК-связывающий домен и домен активации связаны функционально если в слитом полипептиде часть ZFP, TALE или Cas ДНК-связывающего участка домена способна связывать свой сайт-мишень и/или свой сайт связывания, в то время как домен активации может повышать экспрессию генов. Когда слитый полипептид, в котором ZFP или TALE ДНК-связывающий домен слит с доменом расщепления, ZFP или TALE ДНК-связывающий домен и домен расщепления связаны функционально, если в слитом полипептиде ZFP или TALE ДНК-связывающая часть домена способна связывать сайт-мишень и/или его сайт связывания, а область расщепления может расщеплять ДНК в непосредственной близости от сайта-мишени. "Функциональный фрагмент" белка, полипептида или нуклеиновой кислоты представляет собой белок, полипептид или нуклеиновую кислоту, последовательность которых не совпадает с непроцессированным белком, полипептидом или нуклеиновой кислотой, но сохраняет ту же функцию, что у непроцессированного белка, полипептида или нуклеиновой кислоты. Функциональный фрагмент может содержать большее, меньшее или такое же количество остатков, что и соответствующая нативная молекула, и/или может содержать одну или более аминокислотных или нуклеотидных замен. Методы определения функции нуклеиновой кислоты (например, кодирующей функции, способности к гибридизации с другой нуклеиновой кислотой) хорошо известны в данной области техники. Точно так же методы опре-

деления функции белка хорошо известны. Например, ДНК-связывающая функция полипептида может быть определена, например, с помощью метода связывания на фильтрах, анализа изменения электрофоретической подвижности или иммунопреципитации. Расщепление ДНК может быть проанализировано с помощью гель-электрофореза. См. Ausubel et al., выше. Способность белка взаимодействовать с другим белком может быть определена, например, путем соиммунопреципитации, двухгибридных анализов или комплементации, как генетической, так и биохимической. См., например, Fields et al. (1989) *Nature* 340:245-246; патент США № 5585245 и РСТ WO 98/44350. "Вектор" способен переносить генные последовательности к клеткам-мишеням. Как правило, "векторный конструкт", "вектор экспрессии" и "вектор переноса гена" означает любую конструкцию нуклеиновой кислоты, способную регулировать экспрессию целевого гена, и которая может переносить генные последовательности в клетки-мишени. Таким образом, термин включает клонирование и экспрессию носителей, а также интегрирующих векторов.

"Ген-репортер" или "репортерная последовательность" относится к любой последовательности, которая продуцирует белковый продукт, легко поддающийся измерению, предпочтительно, хотя и не обязательно, в обычном анализе. Подходящие гены-репортеры включают последовательности, кодирующие белки, которые опосредуют резистентность к антибиотику (например, резистентность к ампициллину, резистентность к неомицину, резистентность к G418, резистентность к пурамицину), последовательности, кодирующие окрашенные или флуоресцентные, или люминесцентные белки (например, зеленый флуоресцентный белок, улучшенный зеленый флуоресцентный белок, красный флуоресцентный белок, люциферазу), и белки, которые опосредуют повышенный рост клеток и/или амплификацию гена (например, дигидрофолатредуктазу), но не ограничиваются ими. Эпитопные метки включают, например, одну или более копий FLAG, His, мус, Tap, HA или любую обнаруживаемую аминокислотную последовательность. "Экспрессионные метки" включают последовательности, кодирующие репортеры, которые могут быть функционально связаны с желательной генной последовательностью, чтобы контролировать экспрессию целевого гена. Термины "субъект" и "пациент" используются взаимозаменяемо и относятся к млекопитающим, таким как больные люди и приматы, кроме человека, а также к экспериментальным животным, таким как кролики, собаки, кошки, крысы, мыши и другие животные. Соответственно, термин "субъект" или "пациент" используемый в контексте настоящего документа, обозначает любого пациента-млекопитающее или субъект, которым могут быть введены модифицированные эритроциты (или стволовые клетки) в соответствии с настоящим изобретением. Субъекты в соответствии с настоящим изобретением, включают тех, которые были подвержены действию одного или более химических токсинов, включая, например, нейротоксин.

Нуклеазы.

В настоящем документе раскрыты композиции, в частности нуклеазы, которые являются пригодными для нацеливания гена с целью применения при гемоглобинопатиях. В определенных вариантах реализации нуклеазы имеют природное происхождение. В других вариантах реализации нуклеазы не являются природными, т.е. являются сконструированными в ДНК-связывающем домене и/или домене расщепления. Например, ДНК-связывающий домен нуклеазы природного происхождения может быть модифицирован с целью связывания выбранного сайта-мишени (например, мегануклеаза, которая была модифицирована для связывания с сайтом, отличающимся от когнантного сайта связывания). В других вариантах реализации нуклеаза содержит гетерологичные ДНК-связывающие сайты и сайты расщепления (например, нуклеазы "цинковые пальцы"; TAL-эффекторные нуклеазы; мегануклеазные ДНК-связывающие домены с гетерологичными доменами расщепления), или универсальная нуклеаза направляется специфичной гидовой РНК (например, CRPISR/Cas).

А. ДНК-связывающие домены.

В определенных вариантах реализации изобретения нуклеаза является мегануклеазой ("хоминг"-эндонуклеаза). Мегануклеазы природного происхождения распознают сайты расщепления из 15-40 пар оснований и обычно группируются в четыре семейства: семейство LAGLIDADG, семейство GIY-YIG, боксовое семейство His-Cyst и семейство HNH. Примеры "хоминг"-эндонуклеаз включают I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII и I-TevIII. Их последовательности распознавания известны. См. также патент США № 5420032; патент США № 6833252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82: 115-118; Verler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 и каталог New England Biolabs.

В определенных вариантах реализации изобретения нуклеаза содержит сконструированную (не природного происхождения) "хоминг"-эндонуклеазу (мегануклеазу). Последовательности распознавания "хоминг"-эндонуклеаз и мегануклеаз, такие как I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII и I-TevIII, известны. См. также патент США № 5420032; патент США № 6833252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 и каталог New England Biolabs. Кроме того, ДНК-связывающая специфичность "хоминг"-эндонуклеаз и мегануклеаз может быть сконструирована таким образом, чтобы связывать неприродные сайты-мишени. См., напри-

мер, Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66; патентную публикацию США № 20070117128. ДНК-связывающие домены "хоминг"-эндонуклеаз и мегануклеаз могут изменяться в контексте нуклеазы в общем (т.е. таким образом, что нуклеаза содержит когнанный сайт расщепления) или может быть слитой с гетерологичным доменом расщепления.

В других вариантах реализации изобретения ДНК-связывающий домен содержит природный или сконструированный (не встречающийся в природе) TAL эффекторный ДНК-связывающий домен. См., например, патентную публикацию № 20110301073, включенную в настоящий документ посредством ссылки в полном объеме. Фитопатогенные бактерии рода *Xanthomonas*, как известно, вызывают множество заболеваний у важных по значению сельскохозяйственных культур. Патогенность *Xanthomonas* зависит от консервативного типа III секреторной (T3 S) системы, которая инъецирует более чем 25 различных эффекторных белков в растительную клетку. Среди этих инъецированных белков находятся сходные с активатором транскрипции эффекторы (TALE), которые имитируют растительные активаторы транскрипции и манипулируют растительными транскриптомами (см. Kay et al. (2007) *Science* 318:648-651). Эти белки содержат ДНК-связывающий домен и домен активации транскрипции. Одним из наиболее хорошо охарактеризованных TALE является AvrBs3 из *Xanthomonas campestris* патовар *Vesicatoria* (см. Bonas et al. (1989) *Mol. Gen. Genet.* 218: 127-136 и WO 2010079430). TALE содержит централизованный домен тандемных повторов, каждый повтор содержит приблизительно 34 аминокислоты, которые являются ключом к ДНК-связывающей специфичности этих белков. Кроме того, они содержат последовательность внутриядерной локализации и кислотный домен активации транскрипции (обзор см. в Schor-nack S, et al. (2006) *J. Plant Physiol.* 163(3): 256-272). Кроме того, у фитопатогенной бактерии *Ralstonia solanacearum*, как было обнаружено, два гена, обозначенные *brg11* и *hpx17*, являются гомологичными семейству AvrBs3 из *Xanthomonas* у штамма GM1000 биовара 1 *R. solanacearum* и у штамма RS1000 биовара 4 (см. Heuer et al. (2007) *Appl and Envir Micro* 73(13): 4379-4384). Эти гены на 98,9% идентичны друг другу в нуклеотидной последовательности, но отличаются делецией 1575 п.о. в домене повтора *hpx17*. Однако оба генных продукта обладают менее чем 40% идентичностью последовательностей с белками семейства AvrBs3 из *Xanthomonas*.

Таким образом, в некоторых вариантах реализации изобретения ДНК-связывающий домен, который связывает сайт-мишень в локусе-мишени (например, глобиновом или "безопасной гавани") является сконструированным доменом из TAL-эффектора, подобного тем, что получены от фитопатогенов *Xanthomonas* (см. Boch et al., (2009) *Science* 326: 1509-1512 и Moscou and Bogdanove, (2009) *Science* 326: 1501) и *Ralstonia* (см. Heuer et al. (2007) *Applied and Environmental Microbiology* 73(13): 4379-4384); патент США № 8420782 и 8440431 и патентную публикацию США 20110301073. В определенных вариантах реализации изобретения ДНК-связывающий домен содержит белок "цинковые пальцы" (например, белок "цинковые пальцы", который связывает сайт-мишень в глобиновом гене или гене "безопасной гавани"). Предпочтительно белок "цинковые пальцы" не имеет природного происхождения, он сконструирован, чтобы связывать предпочтительный сайт-мишень. См., например, Beerli et al. (2002) *Nature Biotechnol.* 20:135-141; Pabo et al. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al. (2001) *Nature Biotechnol.* 19:656-660; Segal et al. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al. (2000) *Curr. Opin. Struct. Biol.* 10:411-416; патенты США № 6453242; 6534261; 6599692; 6503717; 6689558; 7030215; 6794136; 7067317; 7262054; 7070934; 7361635; 7253273; и патентные публикации США №№ 2005/0064474; 2007/0218528; 2005/0267061, все включены в настоящий документ посредством ссылки в полном объеме.

Сконструированное связывание "цинковых пальцев" или домен TALE может обладать новой специфичностью связывания по сравнению с природным белком "цинковые пальцы". Методы конструирования включают конструктивную разработку и различные типы отбора, но не ограничиваются ими. Конструктивная разработка включает, например, использование баз данных, содержащих триплетные (или квадруплетные) нуклеотидные последовательности и индивидуальные аминокислотные последовательности "цинковых пальцев", в которых каждая триплетная или квадруплетная нуклеотидная последовательность ассоциируется с одной или более аминокислотными последовательностями "цинковых пальцев", связывающими конкретную триплетную или квадруплетную последовательность. См., например, патенты США 6453242 и 6534261 того же заявителя, включенные в настоящий документ посредством ссылки в полном объеме. Примеры методов отбора, включая фаговый дисплей и двухгибридные системы, раскрыты в патентах США 5789538; 5925523; 6007988; 6013453; 6410248; 6140466; 6200759 и 6242568; а также WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 и GB 2338237. Кроме того, улучшение специфичности связывания доменов связывания "цинковых пальцев" раскрыты, например, в WO 02/077227.

Кроме того, как раскрыто в настоящем документе и других ссылках, ДНК-домены (например, мультипальцевый белок "цинковые пальцы" или домены TALE) могут соединяться вместе с использованием любых пригодных линкерных последовательностей, включая, например, линкеры длиной 5 или более аминокислот. См., также патенты США № 6479626; 6903185 и 7153949 касательно характерных линкерных последовательностей длиной 6 или более аминокислот. ДНК-связывающие белки, раскрытые в настоящем документе, могут включать любую комбинацию пригодных линкеров между отдельными "цин-

ковыми пальцами" белка. Кроме того, улучшение специфичности связывания доменов связывания "цинковых пальцев" раскрыты, например, в WO 02/077227.

Отбор сайтов-мишеней; ДНК-связывающих доменов и способов для разработки и конструирования слитых белков (и полинуклеотидов, их кодирующих) известны специалистам в данной области техники и подробно раскрыты в патентах США № 61400815; 789538; 6453242; 6534261; 5925523; 6007988; 6013453; 6200759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 и WO 03/016496 и патентной публикации США № 20110301073. Кроме того, как раскрыто в настоящем документе и других ссылках, ДНК-связывающие домены (например, мультипальцевый белок "цинковые пальцы") могут соединяться вместе с использованием любых пригодных линкерных последовательностей, включая, например, линкеры длиной 5 аминокислот. См. также патенты США № 6479626; 6903185 и 7153949 касательно примеров линкерных последовательностей длиной 6 или более аминокислот. Белки, раскрытые в настоящем документе, могут содержать любую комбинацию пригодных линкеров между отдельными "цинковыми пальцами" белка.

В. Домены расщепления.

Любой пригодный домен расщепления может функционально соединяться с ДНК-связывающим доменом для образования нуклеазы. Например, ZFP ДНК-связывающие домены сливаются с нуклеазными доменами для создания ZFN - функциональной единицы, которая способна распознать свою наметченную нуклеиновую кислоту-мишень посредством своего сконструированного (ZFP) ДНК-связывающего домена и вызывает разрезание ДНК около сайта связывания ZFP посредством нуклеазной активности. См., например, Kim et al. (1996) Proc. Nat'l. Acad. Sci USA 93(3): 1156-1160. Совсем недавно ZFN были использованы для модификации генома в различных организмах. См., например, патентные публикации США 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; и международную публикацию WO 07/014275. Аналогично, TALE ДНК-связывающие домены были слиты с нуклеазными доменами для создания TALEN. См., например, патентную публикацию США № 20110301073. Как отмечалось выше, домен расщепления может быть гетерологичным относительно ДНК-связывающего домена, например ДНК-связывающий домен "цинковых пальцев" и домен расщепления из нуклеазы или TALEN ДНК-связывающий домен и домен расщепления, или мегануклеазный ДНК-связывающий домен и домен расщепления из разных нуклеаз. Гетерологичные домены расщепления могут быть получены из любой эндонуклеазы или экзонуклеазы. Примеры эндонуклеаз, из которых может быть получен домен расщепления, включают рестрикционные эндонуклеазы и "хоминг" эндонуклеазы, но не ограничиваются ими. См., например, каталог 2002-2003, New England Biolabs, Беверли, Массачусетс; и Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388. Дополнительные ферменты, которые расщепляют ДНК, известны (например, нуклеаза S1; маш-нуклеаза; панкреатическая ДНКаза I; микрококковая нуклеаза; дрожжевая HO эндонуклеаза; см. также Linn et al. (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993). Один или более из этих ферментов (или их функциональных фрагментов) могут быть использованы в качестве источника доменов расщепления и полудоменов расщепления.

Аналогично, полудомен расщепления может быть получен из любой нуклеазы или ее части, как указано выше, что требует димеризации для расщепляющей активности. Обычно, для расщепления требуется два слитых белка, если слитые белки содержат полудомены расщепления. В качестве альтернативы может быть использован единственный белок, содержащий два полудомена расщепления. Два полудомена расщепления могут быть получены из одной эндонуклеазы (или ее функциональных фрагментов), или каждый полудомен расщепления может быть получен из разных эндонуклеаз (или их функциональных фрагментов). Кроме того, сайты-мишени для двух слитых белков предпочтительно расположены по отношению друг к другу таким образом, что связывание двух слитых белков с их соответствующими сайтами-мишенями помещает полудомены расщепления в пространственную ориентацию относительно друг друга, которая позволяет полудоменам расщепления образовывать функциональный домен расщепления, например, путем димеризации. Таким образом, в определенных вариантах реализации ближние края сайтов-мишеней разделены 5-8 нуклеотидами или 15-18 нуклеотидами. Однако любое целое число нуклеотидов или нуклеотидных пар может лежать между двух сайтов-мишеней (например, от 2 до 50 пар нуклеотидов и более). В общем, сайт расщепления находится между сайтами-мишенями.

Рестрикционные эндонуклеазы (рестрикционные ферменты) присутствуют у многих видов и способны специфично связываться с последовательностью ДНК (в сайте распознавания), и расщеплять ДНК у или вблизи сайта связывания. Некоторые рестрикционные ферменты (например, Тип IIS) расщепляют ДНК на сайтах, удаленных от сайта распознавания и содержат разделяемые домены связывания и расщепления. Например, тип IIS фермента Fok I катализирует двухцепочечное расщепление ДНК, у 9 нуклеотидов из ее сайта распознавания на одной цепи и 13 нуклеотидов из ее сайта распознавания на другой. См., например, патенты США 5356802; 5436150 и 5487994; а также Li et al. (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim et al. (1994b) J. Biol. Chem. 269:31, 978-31, 982. Таким образом, в одном варианте реализации слитые белки содержат домен расщепления (или полудомен расщепления) по меньшей мере из одного рестрикционного фермента типа IIS и один или более доменов связывания

"цинковых пальцев", которые могут быть сконструированными или несконструированными.

Примером типа IIS рестрикционного фермента, чей домен расщепления отделяется от связывающего домена, является Fok I. Этот конкретный фермент активный в качестве димера. Bitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 10, 570-10,575. Соответственно, для целей настоящего документа часть фермента Fok I, используемая в раскрытых слитых белках, считается полудоменом расщепления. Таким образом, для целенаправленного двухцепочечного расщепления и/или целенаправленной замены клеточных последовательностей с использованием слияний "цинковый палец"-Fok I, два слитых белка, каждый из которых содержит полудомен расщепления Fok I, могут быть использованы для восстановления каталитически активного домена расщепления. В качестве альтернативы также может быть использована одна молекула полипептида, содержащего ДНК-связывающий домен и два полудомена расщепления Fok I.

Домен расщепления или полудомен расщепления может быть любой частью белка, которая сохраняет активность расщепления или сохраняет способность к мультимеризации (например, димеризации) с образованием функционального домена расщепления.

Примеры рестрикционных ферментов типа IIS раскрыты в международной публикации WO 07/014275, включенной в настоящий документ в полном объеме. Дополнительные рестрикционные ферменты также содержат разделяемые домены связывания и расщепления и охватываются настоящим документом. См., например, Roberts et al. (2003) *Nucleic Acids Res.* 31:418-420. В определенных вариантах реализации домен расщепления содержит один или более сконструированных полудоменов расщепления (также упоминаются как мутантные домены димеризации), которые минимизируют или предотвращают гомодимеризацию, как раскрыто, например, в патентных публикациях США №№ 20050064474; 20060188987 и 20080131962, раскрытие всех из них включено в настоящий документ посредством ссылки в полном объеме. Аминокислотные остатки в положениях 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 и 538 в Fok I все являются мишенями для влияния на димеризацию полудоменов расщепления Fok I. Примеры сконструированных полудоменов расщепления Fok I, которые образуют облигатные гетеродимеры, включают пару, в которой первый полудомен расщепления включает мутации аминокислотных остатков в положениях 490 и 538 Fok I, а второй полудомен расщепления включает мутации аминокислотных остатков 486 и 499.

Таким образом, в одном варианте реализации изобретения мутация в 490 заменяет Glu (E) на Lys (K); мутация в 538 заменяет Iso (I) на Lys (K); мутация в 486 заменяет Gln (Q) на Glu (E); а мутация в положении 499 заменяет Iso (I) на Lys (K). В частности, сконструированные полудомены расщепления, раскрытые в настоящем документе, были получены путем мутации в положениях 490 (E → K) и 538 (I → K) в одном полудомене расщепления для получения сконструированного полудомена расщепления, обозначенного "E490K:1538K", и мутаций в положениях 486 (Q → E) и 499 (I → L) в другом полудомене расщепления для получения сконструированного полудомена расщепления, обозначенного "Q486E:I499L". Сконструированные полудомены расщепления, раскрытые в настоящем документе, являются облигатными гетеродимерными мутантами, в которых аберрантное расщепление минимизировано или упразднено. См., например, патентную публикацию США № 2008/0131962, раскрытие которой включено посредством ссылки в полном объеме для всех целей.

В определенных вариантах реализации сконструированный полудомен расщепления содержит мутации в положениях 486, 499 и 496 (пронумерованы относительно Fok I дикого типа), например, мутации, которые замещают остаток Gln (Q) дикого типа в положении 486 остатком Glu (E), остаток Iso (I) дикого типа в положении 499 остатком лейцина (L) и остаток Asn (N) дикого типа в положении 496 остатком Asp (D) или Glu (E) (также упоминаются как домены «ELD» и «ELE» соответственно). В других вариантах реализации сконструированный полудомен расщепления содержит мутации в положениях 490, 538 и 537 (пронумерованы относительно Fok I дикого типа), например, мутации, которые замещают остаток Glu (E) дикого типа в положении 490 остатком Lys (K), остаток Iso (I) дикого типа в положении 538 остатком Lys (K) и остаток His (H) дикого типа в положении 537 остатком Lys (K) или остатком Arg (R) (также упоминаются как домены "KKK" и "KKR" соответственно). В других вариантах реализации сконструированный полудомен расщепления содержит мутации в положениях 490 и 537 (пронумерованы относительно Fok I дикого типа), например, мутации, которые замещают остаток Glu (E) дикого типа в положении 490 остатком Lys (K), остаток His (H) дикого типа в положении 537 остатком Lys (K) или остатком Arg (R) (также упоминаются как домены "KIK" и "KIR" соответственно). (См. патентную публикацию США № 20110201055, включенную в настоящий документ посредством ссылки). Сконструированные полудомены расщепления, раскрытые в настоящем документе, могут быть получены с использованием любого пригодного метода, например с помощью сайт-направленного мутагенеза полудоменов расщепления дикого типа (Fok I), как раскрыто в патентных публикациях США №№ 20050064474; 20080131962 и 20110201055. В качестве альтернативы нуклеазы могут быть собраны *in vivo* в сайте-мишени нуклеиновой кислоты с использованием так называемой технологии "расщепление-фермент" (см., например, патентную публикацию США № 20090068164). Компоненты таких расщепляющих ферментов могут экспрессироваться в отдельных экспрессионных конструктах или могут быть соединены в одной открытой рамке считывания, в которой индивидуальные компоненты разделены, например, само-

расщепляющимся пептидом 2A или последовательностью IRES. Компоненты могут быть индивидуальными доменами связывания "цинковых пальцев" или доменами связывающего домена нуклеиновой кислоты мегануклеазы.

Нуклеазы могут быть подвергнуты скринингу на активность перед применением, например, в дрожжевой хромосомной системе, как раскрыто в WO 2009/042163 и 20090068164. Конструкты экспрессии нуклеаз могут быть легко разработаны с применением способов, известных в данной области техники. См., например, патентные публикации США 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231 и международную публикацию WO 07/014275. Экспрессия нуклеазы может находиться под контролем конститутивного промотора или индуцируемого промотора, например промотора галактокиназы, который активируется (дерепрессируется) в присутствии рафинозы и/или галактозы и подавляется в присутствии глюкозы.

Система CRISPR/Cas.

Убедительные доказательства в последнее время появились в пользу существования РНК-опосредованного генома защитного пути у архей и многих бактерий, гипотеза о котором была выдвинута параллельно эукариотическому пути РНКи (обзор см. в Godde and Bickerton, 2006. *J. Mol. Evol.* 62: 718-729; Lillestol et al., 2006. *Archaea* 2: 59-72; Makarova et al., 2006. *Biol. Direct* 1:7.; Sorek et al., 2008. *Nat. Rev. Microbiol.* 6: 181-186). Известный как система CRISPR-Cas или прокариотическая РНКи (пРНКи), путь, как предполагается, возник из двух эволюционно и часто физически связанных генных локусов: локуса CRISPR (кластерных коротких палиндромных повторов, разделённых регулярными промежутками), который кодирует РНК-компоненты системы, и локуса cas (CRISPR-ассоциированного), который кодирует белки (Jansen et al., 2002. *Mol. Microbiol.* 43: 1565-1575; Makarova et al., 2002. *Nucleic Acids Res.* 30: 482-496; Makarova et al., 2006. *Biol. Direct* 1:7; Haft et al., 2005. *PLoS Comput. Biol.* 1: e60). Локусы CRISPR в микробных хозяевах содержат комбинацию CRISPR-ассоциированных (Cas) генов, а также некодирующие РНК-элементы, способные программировать специфичность CRISPR-опосредованного расщепления нуклеиновых кислот. Индивидуальные белки Cas не разделяют значительного сходства с белковыми компонентами эукариотического РНКи комплекса, но обладают аналогичными предсказанными функциями (например, РНК-связывание, нуклеаза, хеликаза, и т.п.) (Makarova et al., 2006. *Biol. Direct* 1: 7). CRISPR-ассоциированные (cas) гены часто связаны с наборами CRISPR повтор-спейсер. Описано более чем сорок различных Cas-белковых семейств. В этих белковых семействах, Cas1, как представляется, является убиквитарным среди различных систем CRISPR/Cas. Отдельные комбинации генов cas и структур повторов используются для определения 8 подтипов CRISPR (*Ecoli*, *Ypest*, *Nmeni*, *Dvulg*, *Tneap*, *Hmari*, *Apern* и *Mtube*), некоторые из них ассоциируются с дополнительным генным модулем, кодирующим повтор-ассоциированные загадочные белки (RAMP). В одном геноме может находиться более чем один подтип CRISPR. Спорадическое распределение подтипов CRISPR/Cas наводит на мысль о том, что система является предметом для горизонтального генного переноса во время микробиологической эволюции. CRISPR типа II (представленный Cas9) является одной из наиболее хорошо охарактеризованных систем и осуществляет целенаправленный двухцепочечный разрыв ДНК в четыре последовательных этапа. На первом, две некодирующие РНК, набор пре-crРНК и tracrРНК, транскрибируются из локуса CRISPR. На втором, tracrРНК гибридизует с повторными областями пре-crРНК и опосредует процессинг пре-crРНК в зрелые crРНК, содержащие индивидуальные спейсерные последовательности. На третьем, комплекс зрелая crРНК:tracrРНК направляет Cas9 к ДНК-мишени путем спаривания на основе правила Уотсона-Крика между спейсером на crРНК и протоспейсером на ДНК-мишени рядом с протоспейсерным примыкающим мотивом (PAM), дополнительное требование для распознавания мишени. В конце, Cas9 опосредует расщепление ДНК-мишени для создания двухцепочечного разрыва в пределах протоспейсера. Активность системы CRISPR/Cas включает три этапа: (i) вставка чужих ДНК-последовательностей в набор CRISPR для предотвращения будущих атак, в ходе процесса, названного "адаптацией", (ii) экспрессия релевантных белков, а также экспрессия и процессинг набора, за которыми следует (iii) РНК-опосредованная интерференция с чужой нуклеиновой кислотой. Таким образом, в бактериальной клетке несколько так называемых белков "Cas" являются связанными с природной функцией системы CRISPR/Cas.

Первичные продукты локусов CRISPR, как представляется, являются короткими РНК, которые содержат внедряющиеся целенаправленные последовательности, и названы гидовыми РНК или прокариотическими сайленсинговыми РНК (psiРНК), исходя из их гипотетической роли в метаболическом пути (Makarova et al. (2006) *Biol. Direct* 1:7; Hale et al. (2008) *RNA* 14: 2572-2579). Анализ РНК указывает, что локусные транскрипты CRISPR расщепляются в пределах последовательностей повторов с высвобождением 60-70-нитронных промежуточных РНК-соединений, которые содержат индивидуальные внедряющие целенаправленные последовательности и фланкирующие фрагменты повторов (Tang et al. (2002) *Proc. Natl. Acad. Sci.* 99: 7536-7541; Tang et al. (2005) *Mol. Microbiol.* 55:469-481; Lillestol et al. (2006) *Archaea* 2:59-72; Brouns et al. (2008) *Science* 321: 960-964; Hale et al. (2008) *RNA* 14:2572-2579). У архайной *Rugosoccus furiosus*, эти промежуточные РНК дополнительно процессируются до часто встречающихся, стабильных 35-45-нитронных зрелых psiРНК (Hale et al. (2008) *RNA* 14: 2572-2579).

Белки Cas.

Полипептид "Cas1" относится к CRISPR ассоциированному (Cas) белку 1. Cas1 (COG1518 в кластерах ортологичной группы системы классификации белков), является лучшим маркером CRISPR-ассоциированных систем (CASS). Исходя из филогенетических сравнений, идентифицировано семь отдельных версий CRISPR-ассоциированной иммунной системы (CASS1-7). Полипептидом Cas1, используемым в способах, раскрытых в настоящем документе, может быть любой полипептид Cas1, присутствующий в любом прокариоте. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 архайичного микроорганизма. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 микроорганизма Euryarchaeota. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 микроорганизма Crenarchaeota. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 бактерии. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 грамотрицательной или грамположительной бактерии. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 *Pseudomonas aeruginosa*. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1 *Aquifexaolicus*. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1, который является членом одного из CASS1-7. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1, который является членом CASS3. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1, который является членом CASS7. В определенных вариантах реализации полипептидом Cas1 является полипептид Cas1, который является членом CASS3 или CASS7. В некоторых вариантах реализации полипептид Cas1 кодируется нуклеотидной последовательностью, обеспеченной GenBank, например, под номером GeneID: 2781520, 1006874, 9001811, 947228, 3169280, 2650014, 1175302, 3993120, 4380485, 906625, 3165126, 905808, 1454460, 1445886, 1485099, 4274010, 888506, 3169526, 997745, 897836 или 1193018, и/или представлен аминокислотной последовательностью, демонстрирующей гомологию (например, более чем 80%, от 90 до 99%, включая 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%) в отношении аминокислот, кодируемых указанными полинуклеотидами, причем полипептиды функционируют как полипептиды Cas1.

Cas6 является другим полипептидом Cas, и эндорибонуклеазная активность относится в настоящем документе к эндорибонуклеазной активности Cas6. Неограничивающие примеры пригодных полипептидов указаны в Genbank под регистрационным номером AAL81255. Полипептид Cas6 может быть обогащенным, выделенным или очищенным из микроорганизма, содержащего locus CRISPR и locus cas (CRISPR-ассоциированный), такой как *Rugococcus furiosus*, но не ограничиваться им, или может быть получен с использованием рекомбинантных технологий, или может быть химически или ферментативно синтезированным с использованием шаблонных методов. В некоторых аспектах полипептид Cas6 может быть обогащенным, выделенным или очищенным из микроорганизма, который не содержит locus CRISPR. Полипептид Cas6 содержит по меньшей мере один остаток, который может играть роль в катализе, или его консервативную замену. Полипептид Cas6 может содержать другие остатки, которые также могут играть роль в катализе, или их консервативные замены. Ожидается, что остаток(и), играющий роль в катализе, может располагаться возле богатой G петли, которая содержит опознавательный мотив Cas6 в 3D структуре белка. Полипептиды Cas6 могут содержать домены, присутствующие в базе данных TIGRFAM под регистрационными номерами TIGR01877 и PF01881. База данных TIGRFAM содержит семейства полипептидов, функция которых сохраняется (Haft et al. (2003) Nucl. Acids Res. 31:371-373, Bateman and Haft (2002) Briefings Bioinformatics, 3:236-245, и Haft et al. (2005) PLoS Computational Biol. 1(6):e60). Другие примеры полипептидов Cas6, предложенных в настоящем документе, включают полипептиды, присутствующие в прокариотических микроорганизмах, содержащих locus CRISPR и locus cas. Полипептиды Cas6 можно легко идентифицировать в любом микроорганизме, который содержит locus CRISPR. Кодированная область, которая кодирует полипептид Cas6, как правило, находится в локусе cas, расположенном в непосредственной близости к локусу CRISPR. Haft et al. (2005) PLoS Computational Biol. 1(6):e60) составлен обзор семейства белков Cas и созданы правила идентификации конкретных подтипов системы CRISPR/Cas. Haft et al. описывают кодирующую область, которая кодирует полипептиды Cas6, как обнаружено в связи по меньшей мере с четырьмя отдельными подтипами CRISPR/Cas (Tneap, Hmari, Aperi и Mtube), и как правило является cas-кодирующей областью, расположенной в наибольшем отдалении от локуса CRISPR. Полипептиды Cas6 можно идентифицировать, используя ресурсы, доступные от JCVI Comprehensive Microbial Resource. Таким образом, полипептиды Cas6, пригодные для способов, раскрытых в настоящем документе, могут быть идентифицированы специалистом в данной области с использованием шаблонных методов. Примеры прокариотических микробов с полностью известными геномными последовательностями, содержащими кодирующие области, которые, как ожидается, кодируют полипептид Cas6, включают *Thermotogamaritima* MSB8, *Campylobacter fetus* подвид *fetus* 82-40, *Fusobacterium nucleatum* ATCC 25586, *Streptococcus thermophilus* LMG 18311, *Thermoanaerobacter tengcongensis* MB4(T), *Moorellathermoacetica* ATCC 39073, *Desulfotobacterium hafniense* Y51, *Clostridium tetani* E88, *Clostridium perfringens* SM101, *Clostridium difficile* QCD-32g58, *Clostridium botulinum* Hall A Sanger, *Clostridium botulinum* F Langeland, *Clostridium botulinum* B1 штамм Okra, *Clostridium botulinum* A3 штамм Loch Maree, *Clostridium botulinum* A Hall, *Clostridium botulinum* A ATCC 19397, *Carboxydother-*

mushydrogenoformans Z-2901, *Staphylococcus epidermidis* RP62A, *Thermophilus* HB8, *Thermophilus* HB27, вид *Nostoc PCC 7120*, *Anabaena variabilis* ATCC 29413, вид *Synechococcus OS* тип В prime, вид *Synechococcus OS* тип А, *Porphyromonas gingivalis* W83, *Bacteroides fragilis* YCR46, *Bacteroides fragilis* NCTC9343, *Aquifexaerolicus* VF5, *Rubrobacter xylophilus* DSM 9941, *Mycobacterium tuberculosis* H37Rv (лабораторный штамм), *Mycobacterium tuberculosis* CDC1551, *Mycobacterium bovis* подвид *bovis AF2122/97*, *Frankia alni* ACN14a, *Thermoplasma volcanium* GSS1, *Picrophilus torridus* DSM 9790, *Thermococcus kodakarensis* KOD1, *Pyrococcus horikoshii* OT3, *Pyrococcus furiosus* DSM 3638, *Pyrococcus abyssi* GE5, *Methanosarcina barkeri* fusaro, *Methanosarcina acetivorans* C2A, *Methanococcoides burtonii* DSM 6242, *Methanococcus jannaschii* DSM2661, *Methanobacterium thermoautotrophicum* дельта Н, *Haloarcula marismortui* ATCC 43049, *Archaeoglobus fulgidus* DSM4304, *Pyrobaculum aerophilum* IM2, *Sulfolobus tokodaii* штамм 7, *Sulfolobus solfataricus* P2, *Sulfolobus acidocaldarius* DSM 639, *Aeropyrum pernix* K1. Другие примеры полипептидов Cas6 известны специалистам в данной области, см., например, члены группы полипептидов COG1583 (доступные из кластеров ортологических групп белков (COG) на веб-странице Интернет-сайта National Center for Biotechnology Information, см. также Tatusov et al. (1997) Science 278:631-637 и Tatusov et al. (2003) BMC Bioinformatics 4(1):41), члены семейства InterPro, входящий номер IPRO10156, Makarova et al. (2002) Nuc. Acids Res. 30:482-496 и Haft et al. (2005) PLoS Comput. Biol. 1(6):e60, 474-483).

Существует три типа систем CRISPR/Cas, все из которых включают crPHK и белки Cas. Тип I и III содержит Cas-эндонуклеазы, обрабатывающие пре-crPHK, которые, после полного превращения в crPHK, собирают поли-Cas белковый комплекс, который способен расщеплять нуклеиновые кислоты, комплементарные crPHK. В типе II систем CRISPR/Cas crPHK вырабатываются с использованием разных механизмов, где транс-активирующая РНК (tracrPHK), комплементарная повторным последовательностям в пре-crPHK, запускает процессинг двухцепочечно-специфичной рибонуклеазой III в присутствии белка Cas9. Затем Cas9 способен расщеплять ДНК-мишень, которая комплементарна зрелой crPHK, однако расщепление с помощью Cas9 зависит от спаривания оснований между crPHK и ДНК-мишенью и от присутствия короткого мотива в crPHK, названного последовательностью PAM (фотоспейсерный прилегающий мотив) (см. Qi et al. (2013) Cell 152:1173). Кроме того, tracrPHK должна также присутствовать в виде пар оснований с crPHK на своем 3'-конце, и эта ассоциация запускает активность Cas9.

Белок Cas9 содержит по меньшей мере два нуклеазных домена: один нуклеазный домен сходен с эндонуклеазой HNH, тогда как другой сходен с эндонуклеазным доменом Ruv. Как обнаружено, домен типа HNH ответственен за расщепление нити ДНК, которая комплементарна crPHK, тогда как домен Ruv расщепляет некомплементарную нить. Потребности в комплексе crPHK-tracrPHK можно избежать, используя сконструированную "однонаправляющую РНК" (sgPHK), которая содержит шпильку, обычно образованную путем отжига crPHK и tracrPHK (см. Jinek et al. (2012) Science 337:816 и Cong et al. (2013) Scienceexpress/10.1126/science. 1231143). В *S. pyogenes*, сконструированное слияние tracrPHK :crPHK, или sgRNA, направляет Cas9 на расщепление ДНК-мишени, в ходе которого двухцепочечный гетеродимер РНК:ДНК образуется между Cas-ассоциированными РНК и ДНК-мишенью. Эта система, содержащая белок Cas9, и сконструированная sgPHK, содержащая последовательность PAM, была использована для РНК-направленного редактирования генома (см. Ramalingam, там же) и была подходящей для редактирования генома эмбриона рыбок данио in vivo (см. Hwang et al. (2013) Nature Biotechnology 31(3):227) с эффективностью редактирования, подобной ZFN и TALEN. В определенных вариантах реализации белок Cas может быть "функциональным производным" белка Cas природного происхождения. "Функциональное производное" природной последовательности полипептида представляет собой соединение, обладающее качественным биологическим свойством, общим с природной последовательностью полипептида. "Функциональные производные" включают фрагменты природной последовательности и производные природной последовательности полипептида и их фрагменты, при условии, что они обладают биологической активностью, общей с соответствующей природной последовательностью полипептида, но не ограничиваются ими. Биологическая активность, предложенная настоящим документом, является способностью функционального производного гидролизовать субстрат ДНК на фрагменты. Термин "производное" охватывает варианты аминокислотной последовательности полипептида, ковалентные модификации и их слияния. "Полипептид Cas" охватывает непротессированный полипептид Cas, ферментативно активный фрагмент полипептида Cas и ферментативно активные производные полипептида Cas и их фрагменты. Пригодные производные полипептида Cas или их фрагменты включают мутанты, слияния, ковалентные модификации белка Cas или их фрагменты, но не ограничиваются ими.

Белки Cas и полипептиды Cas могут быть получены из клеток или синтезированы химически или с помощью комбинации этих двух процедур. Клеткой может быть клетка, которая природным образом вырабатывает белок Cas, или клетка, которая природным образом вырабатывает белок Cas и создана методом генной инженерии для выработки эндогенного белка Cas с более высокими уровнями экспрессии или для выработки белка Cas из эндогенно введенной нуклеиновой кислоты, причем нуклеиновая кислота кодирует Cas, такой же как и эндогенный Cas, или отличающийся. В некоторых случаях клетка не вырабатывает белок Cas природным образом, а является созданной методом генной инженерии для выработки белка Cas.

Дополнительно, система CRISPR/Cas может быть использована для ингибирования генной экспрес-

сии. Lei et al. (2013) Cell 152(5): 1173-1183) показали, что каталитически "мертвый" Cas9, не обладающий эндонуклеазной активностью, в случае совместной экспрессии с направляющей РНК, создает ДНК-распознающий комплекс, который позволяет целенаправленно вмешиваться в транскрипционную элонгацию, связывание РНК полимеразы или связывание фактора транскрипции. Эта система, названная CRISPR интерференцией (CRISPRi), может эффективно подавлять экспрессию генов-мишеней.

Кроме того, были разработаны белки Cas, которые содержат мутации в доменах расщепления, чтобы сделать их неспособными вызывать DSB, вместо чего они вводят одноцепочечный разрыв в ДНК-мишень ("Cas9 никующий фермент", см. Cong et al., там же).

Белки Cas согласно изобретению могут подвергаться мутагенезу с целью модификации функциональности. Примеры методов селекции, включая фаговый дисплей и двухгибридные системы, раскрыты в патентах США 5789538; 5925523; 6007988; 6013453; 6410248; 6140466; 6200759 и 6242568; а также WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 и GB 2338237.

РНК-компоненты CRISPR/Cas.

Cas9 связанная система CRISPR/Cas содержит два некодирующих РНК-компонента: tracrPНК и массив пре-crPНК, содержащие нуклеазные направляющие последовательности (спейсеры), разделенные промежутками одинаковых прямых повторов (DR). Чтобы использовать систему CRISPR/Cas для осуществления геномной инженерии, должны присутствовать обе функции этих РНК (см. Cong et al. (2013) Scienceexpress 1/10.1126/science 1231143). В некоторых вариантах реализации tracrPНК и пре-crPНК поставляются отдельными экспрессирующими конструктами или в виде отдельных РНК. В других вариантах реализации конструируют химерную РНК, в которой генетически сконструированную зрелую crPНК (обеспечивающую специфичность в отношении мишени) сливают с tracrPНК (обеспечивающую взаимодействие с Cas9) для создания гибрида химерная cr-РНК-tracrPНК (также называемого однонаправляющей РНК). (См. Jinek, там же и Cong, там же). Химерные или sgPНК можно сконструировать таким образом, чтобы они содержали последовательность, комплементарную любой желательной мишени. РНК содержат 22 основания, комплементарные мишени и в виде G[n19], за которыми следует фотоспейсерный прилегающий мотив (PAM) в виде NGG. Таким образом, в одном способе sgPНК могут быть разработаны с использованием известных ZFN, нацеленных в целевой ген путем (i) выравнивания последовательности распознавания гетеродимера ZFN с эталонной последовательностью релевантного генома (человеческой, мышиной или определенного вида растений); (ii) идентификации спейсерной области между полусайтами ZFN; (iii) идентификации локализации мотива G[N20]GG, ближайшего к спейсерной области который (если более чем один такой мотив перекрывает спейсер, мотив, который центрируется относительно выбранного спейсера); (iv) использования мотива в качестве ядерной sgRNA. Этот метод преимущественно опирается на испытанные нуклеазные мишени. В качестве альтернативы sgPНК могут быть разработаны для нацеливания на любую целевую область просто путем идентификации пригодной последовательности-мишени, которая соответствует формуле G[n20]GG.

Сайты-мишени.

Как подробно раскрыто выше, ДНК-связывающие домены могут быть сконструированы для связывания любой предпочтительной последовательности в локусе, например, глобинового гена или гена "безопасной гавани". Сконструированный ДНК-связывающий домен может обладать новой специфичностью связывания по сравнению с ДНК-связывающим доменом природного происхождения. Способы конструирования включают конструктивную разработку и различные типы отбора, но не ограничиваются ими. Конструктивная разработка включает, например, использование баз данных, содержащих триплетные (или квадруплетные) нуклеотидные последовательности и индивидуальные аминокислотные последовательности (например, "цинковых пальцев"), в которых каждая триплетная или квадруплетная нуклеотидная последовательность ассоциируется с одной или более аминокислотными последовательностями ДНК-связывающего домена, которые связывают отдельную триплетную или квадруплетную последовательность. См., например, патенты США 6453242 и 6534261 того же заявителя, включенные в настоящий документ посредством ссылки в полном объеме. Конструктивная разработка TAL-эффекторных доменов также может проводиться. См., например, патентную публикацию США № 20110301073. Примеры методов отбора, применимые для ДНК-связывающих доменов, включая фаговый дисплей и двухгибридные системы, раскрыты в патентах США 5789538; 5925523; 6007988; 6013453; 6410248; 6140466; 600 759 и 6242568; а также WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 и GB 2,338,237.

Отбор сайтов-мишеней; нуклеазы и методы разработки и конструирования слитых белков (и полинуклеотидов, их кодирующих) известны специалистам в данной области техники и подробно раскрыты в патентных публикациях США №№ 20050064474 и 20060188987, включенных в настоящий документ посредством ссылки в полном объеме.

Кроме того, как раскрыто в настоящем документе и других ссылках, ДНК-связывающие домены (например, мультипальцевых белков "цинковые пальцы") могут соединяться вместе с использованием любых пригодных линкерных последовательностей, включая, например, линкеры длиной 5 аминокислот. См., например, патенты США № 6479626; 6903185 и 7153949 касательно примеров линкерных последовательностей длиной 6 или более аминокислот. Белки, раскрытые в настоящем документе, могут содер-

жать любую комбинацию пригодных линкеров между отдельными ДНК-связывающими доменами белка. См. также патентную публикацию № 20110287512.

Доноры.

Как отмечалось выше, вставка экзогенной последовательности (также именуемой "донорской последовательностью" или "донором", или "трансгеном"), например, для коррекции мутантного гена или для повышения экспрессии гена дикого типа. Совершенно очевидно, что донорская последовательность, как правило, не идентична геномной последовательности, на место которой она помещается. Донорская последовательность может содержать негомологичную последовательность, фланкированную двумя участками гомологии, чтобы обеспечить эффективную HDR в целевом местоположении. Кроме того, донорские последовательности могут содержать векторную молекулу, содержащую последовательности, которые не являются гомологичными к целевой области в клеточном хроматине. Донорская молекула может содержать несколько прерывающихся участков гомологии клеточному хроматину. Например, для целенаправленной вставки последовательностей, которые обычно не присутствуют в целевой области, указанные последовательности могут присутствовать в молекуле донорской нуклеиновой кислоты и фланкироваться участками гомологии последовательности в целевой области.

В настоящем документе раскрыты способы целенаправленной вставки любых полинуклеотидов для вставки в выбранное местоположение. Полинуклеотиды для вставки также могут быть названы "экзогенными" полинуклеотидами, "донорскими" полинуклеотидами или молекулами или "трансгенами". Донорский полинуклеотид может представлять собой ДНК или РНК, одноцепочечную и/или двухцепочечную, и может вводиться в клетку в линейной или кольцевой форме. См., например, патентные публикации США №№ 20100047805, 20110281361, 20110207221 и заявку США № 13/889162. Донорская последовательность(и) может содержаться в пределах ДНК МС, которая может быть введена в молекулу в линейной или кольцевой форме. При введении в линейной форме концы донорской последовательности могут быть защищенными (например, от экзонуклеолитического разложения) с помощью методов, известных специалистам в данной области техники. Например, один или более дидезоксинуклеотидных остатков добавляют к 3'-концу линейной молекулы и/или самокомплементарные олигонуклеотиды лигируются к одному или двум концам. См., например, Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls et al. (1996) Science 272:886-889. Дополнительные методы защиты экзогенных полинуклеотидов от разложения включают добавление концевой аминокислоты(аминокислот) и использование модифицированных межнуклеотидных связей, таких как, например, фосфотиоаты, фосфорамидаты и остатки О-метилрибозы или дезоксирибозы, но не ограничиваются ими. Полинуклеотид может вводиться в клетку в виде части векторной молекулы, содержащей дополнительные последовательности, такие как, например, точки начала репликации, промоторы и гены, кодирующие резистентность к антибиотикам. Более того, донорские полинуклеотиды могут вводиться в виде "оголенной" нуклеиновой кислоты, в виде нуклеиновой кислоты в комплексе с агентом, таким как липосома или полонексамер, или могут доставляться с помощью вирусов (например, аденовируса, аденоассоциированного вируса (AAV), герпесвируса, ретровируса, лентивируса и интегразодефектного лентивируса (IDLV)).

В определенных вариантах реализации двухцепочечный донор включает последовательности (например, кодирующие последовательности, также именуемые трансгенами) размером более 1 т.н., например от 2 до 200 т.н., от 2 до 10 т.н. (или любое целочисленное значение между ними). Двухцепочечный донор также содержит, например, по меньшей мере один нуклеазный сайт-мишень. В определенных вариантах реализации донор содержит по меньшей мере 1 сайт-мишень, например, для использования с CRISPR/Cas, или 2 сайта-мишени, например, для пары ZFN или TALEN. Обычно, нуклеазные сайты-мишени находятся вне трансгенных последовательностей, например, в направлении 5' и/или 3' относительно трансгенных последовательностей для расщепления трансгена. Сайт(ы) расщепления нуклеазой может быть подходящим для любой нуклеазы. В определенных вариантах реализации сайт(ы)-мишень для расщепления нуклеазой, содержащийся в двухцепочечном доноре, предназначен для той же нуклеазы(нуклеаз), что используется для расщепления эндогенной мишени, в которую расщепленный донор интегрируется путем гомологонезависимых методов.

Донор обычно вставляется таким образом, что его экспрессия управляется эндогенным промотором в сайте вставки, а именно промотором, который управляет экспрессией эндогенного гена, в который вставляется донор (например, глобиновый, AAVS1 и т.д.). Однако очевидно, что донор может содержать промотор и/или энхансер, например конститутивный промотор или индуцибельный или тканеспецифичный промотор.

Донорская молекула может вставляться в эндогенный ген таким образом, что экспрессируются все, некоторые или ни один из эндогенных генов. Например, трансген, как раскрыто в настоящем документе, может вставляться в глобиновый локус, таким образом, что экспрессируются некоторые или ни одна из эндогенных глобиновых последовательностей, например, в виде слияния с трансгеном. В других вариантах реализации трансген (например, с кодирующими глобин последовательностями или без них) интегрируется в любой эндогенный локус, например, локус "безопасной гавани". См., например, патентные публикации США 20080299580; 20080159996 и 201000218264.

Если дополнительные последовательности (например, глобиновые последовательности, эндогенные

или часть трансгена) экспрессируются с трансгеном, то дополнительные (например, глобиновые) последовательности могут быть непроцессированными последовательностями (дикого типа или мутантными) или неполными последовательностями. Неограничивающие примеры функции этих непроцессированных и неполных последовательностей, например, кодирующих глобин последовательностей, включают увеличение периода полувыведения из сыворотки полипептида, экспрессируемого с помощью трансгена (например, терапевтического гена) и/или действующего в качестве носителя.

Кроме того, хотя это и не требуется для экспрессии, экзогенные последовательности могут также включать последовательности, регулирующие транскрипцию или трансляцию, например, промоторы, энхансеры, инсуляторы, участок внутренней посадки рибосомы, последовательности, кодирующие пептиды 2А, и/или сигналы полиаденилирования. Трансгены, внесенные в донорские последовательности, раскрытые в настоящем документе, могут быть выделены из плазмид, клеток или других источников с использованием стандартных методик, известных из уровня техники, таких как полимеразная цепная реакция (PCR). Используемые доноры могут включать различные типы топологии, включая кольцевые сверхспиральные, кольцевые расслабленные, линейные, и т.п. В качестве альтернативы они могут быть химически синтезированными с использованием стандартных методик синтеза олигонуклеотидов. Кроме того, доноры могут быть метилированными или неметилированными. Доноры могут находиться в виде бактериальных или дрожжевых искусственных хромосом (BAC или YAC). Двухспиральные донорские полинуклеотиды, раскрытые в настоящем документе, могут включать одно или более неприродных оснований и/или скелетов. В частности, вставку донорской молекулы с метилированными остатками цитозина можно осуществить, применяя методы, раскрытые в настоящем документе, для достижения состояния транскрипционного покоя в целевой области.

Экзогенный (донорский) полинуклеотид может содержать любую целевую последовательность (экзогенную последовательность). Примеры экзогенных последовательностей включают любую кодирующую полипептид последовательность (например, кДНК), промоторные последовательности, энхансерные последовательности, эпитопные метки, маркерные гены, сайты распознавания расщепляющих ферментов и различные типы экспрессионных конструкторов, но не ограничиваются ими. Маркерные гены включают последовательности, кодирующие белки, которые опосредуют резистентность к антибиотикам (например, резистентность к ампициллину, резистентность к неомицину, резистентность к G418, резистентность к пурамицину), последовательности, кодирующие окрашенные или флуоресцентные, или люминесцентные белки (например, зеленый флуоресцентный белок, усиленный зеленый флуоресцентный белок, красный флуоресцентный белок, люциферазу), и белки, которые опосредуют усиленный клеточный рост и/или амплификацию генов (например, дигидрофолатредуктазу), но не ограничиваются ими.

Эпитопные метки включают, например, одну или более копий FLAG, His, мус, Tap, HA или любую обнаружимую аминокислотную последовательность. В предпочтительном варианте реализации экзогенная последовательность (трансген) содержит полинуклеотид, кодирующий любой полипептид, экспрессия которого желательна в клетке, включая антитела, антигены, ферменты, рецепторы (клеточной поверхности или ядерные), гормоны, лимфокины, цитокины, репортерные полипептиды, факторы роста и функциональные фрагменты любого из вышеперечисленных, но не ограничиваясь ими. Кодирующими последовательностями могут быть, например, кДНК. В определенных вариантах реализации, экзогенные последовательности могут содержать маркерный ген (раскрытый выше), позволяющий отбирать клетки, в которых произошла намеченная интеграция, и присоединенную последовательность, кодирующую дополнительную функциональность. Неограничивающие примеры маркерных генов включают зеленый флуоресцентный белок (GFP), маркеры селекции с помощью лекарственных средств, и т.п.

Дополнительные генные последовательности, которые могут быть вставлены, могут включать, например, гены дикого типа для замены мутированных последовательностей. Например, последовательность гена β -глобина дикого типа может быть вставлена в геном стволовой клетки, в которой произошла мутация эндогенной копии гена. Копия дикого типа может быть вставлена в эндогенный локус или, в качестве альтернативы, может быть направлена в локус "безопасной гавани".

При конструировании таких кассет экспрессии, вслед за методиками согласно настоящему документу, используют методологии, хорошо известные из уровня техники в области молекулярной биологии (см., например, Ausubel или Maniatis). Перед использованием кассет экспрессии для получения трансгенного животного, реактивность кассеты экспрессии на стрессовый индуктор, связанный с выбранными элементами контроля, может быть протестирована с помощью введения кассеты экспрессии в пригодную клеточную линию (например, первичные клетки, трансформированные клетки или иммортализованные клеточные линии).

Дополнительно, несмотря на то, что для экспрессии этого не требуется, экзогенные последовательности могут быть последовательностями, регулирующими транскрипцию или трансляцию, например, промоторами, энхансерами, инсуляторами, участками внутренней посадки рибосомы, последовательностями, кодирующими пептиды 2А, и/или сигналами полиаденилирования. Кроме того, контрольные элементы целевых генов могут быть функционально связанными с генами-репортерами для создания химерных генов (например, репортерных кассет экспрессии).

Также может быть достигнута целевая вставка не кодирующей последовательности нуклеиновой

кислоты. Последовательности, кодирующие антисмысловые РНК, РНКи, кшRNA и микроРНК (миРНК), также могут быть использованы для целевых вставок.

В дополнительных вариантах реализации донорская нуклеиновая кислота может содержать некодирующие последовательности, которые являются специфичными сайтами-мишенями для дополнительных разработок нуклеаз. Впоследствии, дополнительные нуклеазы могут экспрессироваться в клетках, таким образом, что оригинальная донорская молекула расщепляется и модифицируется путем вставки другой целевой донорской молекулы. Таким образом, могут быть получены повторяющиеся интеграции донорских молекул, допускающие стекинг признаков у определенного целевого локуса или у локуса "безопасной гавани".

Доставка.

Нуклеазы, полинуклеотиды, кодирующие эти нуклеазы, донорские полинуклеотиды и композиции, содержащие белки и/или полинуклеотиды, раскрытые в настоящем документе, могут быть доставлены *in vivo* или *ex vivo* любым подходящим способом.

Способы доставки нуклеаз, как раскрыто в настоящем документе, раскрыты, например, в патентах США №№ 6453242; 6503717; 6534261; 6599692; 6607882; 6689558; 6824978; 6933113; 6979539; 7013219 и 7163824, раскрытие которых включено в настоящий документ посредством ссылки в полном объеме.

Нуклеазы и/или донорские конструкторы, как раскрыто в настоящем документе, могут также быть доставлены, используя векторы, содержащие последовательности, кодирующие один или более белков "цинковые пальцы" или TALEN. Может быть использована любая векторная система, включая плазмидные векторы, ретровирусные векторы, лентивирусные векторы, аденовирусные векторы, поксивирусные векторы, герпесвирусные векторы и аденоассоциированные вирусные векторы, и т.п. См. также патенты США №№ 6534261; 6607882; 6824978; 6933113; 6979539; 7013219 и 7163824, включенные в настоящий документ посредством ссылки в полном объеме, но не ограничиваясь ими. Кроме того, очевидно, что любой из этих векторов может содержать одну или более последовательностей, необходимых для лечения. Таким образом, если в клетку вводится одна или более нуклеаз и донорский конструктор, то нуклеазы и/или донорский полинуклеотид могут переноситься одним и тем же вектором или разными векторами. При использовании нескольких векторов, каждый вектор может содержать последовательность, кодирующую одну или несколько нуклеаз и/или донорских конструкторов. Обычные вирусные и невирусные способы переноса генов могут быть использованы для введения нуклеиновых кислот, кодирующих нуклеазы и донорские конструкторы, в клетки (например, клетки млекопитающих) и ткани-мишени. Невирусные векторные системы доставки включают ДНК-плазмиды, "оголенную" нуклеиновую кислоту и нуклеиновую кислоту в комплексе с носителем, таким как липосома или поллоксамер. Вирусные векторные системы доставки включают ДНК и РНК вирусы с эписомальными или интегрированными геномами после доставки в клетку. Обзор процедур генной терапии см. в Anderson, *Science* 256:808-813 (1992); Nabel & Feigner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al. в *Current Topics in Microbiology and Immunology* Doerfler and Bohm (eds.) (1995); и Yu et al., *Gene Therapy* 1:13-26 (1994).

Способы невирусной доставки нуклеиновых кислот включают электропорацию, липофекцию, микроинъекцию, баллистическую трансфекцию, виросомы, липосомы, иммунолипосомы, поликатионные или липидные конъюгаты нуклеиновых кислот, "оголенные ДНК", искусственные вирионы и усиленное агентами поглощение ДНК. Кроме того, для доставки нуклеиновых кислот может быть применена сонопорация с использованием, например, системы Sonitron 2000 (Rich-Mar).

Дополнительные примеры систем доставки нуклеиновых кислот включают системы, предоставленные Amaxa Biosystems (Кельн, Германия), Maxcyte, Inc. (Роквилл, Мэриленд), BTX Molecular Delivery Systems (Холлистон, Массачусетс) и Copernicus Therapeutics Inc, (см., например, патент США № 6008336). Липофекция раскрыта, например, в патентах США № 5049386; 4946787 и 4897355), а реагенты для липофекции имеются в продаже (например, Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые пригодны для эффективной рецептор-распознающей липофекции полинуклеотидов, включают липиды Feigner, WO 91/17424, WO 91/16024. Получение комплексов липид: нуклеиновая кислота, включая направленные липосомы, такие как иммунолипидные комплексы, хорошо известны из уровня техники (см., например, Crystal, *Science* 270:404-410 (1995); Blaese et al., *Cancer Gene Ther.* 2:291-297 (1995); Behr et al., *Bioconjugate Chem.* 5:382-389 (1994); Remy et al., *Bioconjugate Chem.* 5:647-654 (1994); Gao et al., *Gene Therapy* 2:710-722 (1995); Ahmad et al., *Cancer Res.* 52:4817-4820 (1992); патенты США № 4186183, 4217344, 4235871, 4261975, 4485054, 4501728, 4774085, 4837028 и 4946787).

Дополнительные способы доставки включают использование упаковывания нуклеиновых кислот, предназначенных для доставки, в носители EnGeneIC (EDV). Эти EDV специально доставляются в ткани-мишени с использованием биспецифичных антител, в которых одно плечо антитела обладает специфичностью в отношении ткани-мишени, а другое - специфичностью в отношении EDV. Антитело доставляет EDV к целевой клеточной поверхности, а затем EDV доставляется в клетку с помощью эндоцитоза. Оказавшись в клетке, содержимое высвобождается (см. MacDiarmid et al. (2009) *Nature Biotechnol-*

ogy 27(7): 643).

Использование РНК или ДНК вирусных систем для доставки нуклеиновых кислот, кодирующих сконструированные ZFP, обеспечивает преимущество высокоразвитых процессов нацеливания вирусов в определенные клетки в организме и направленной миграции полезной вирусной нагрузки в ядро. Вирусные векторы могут вводиться непосредственно субъектам (*in vivo*), или они могут быть использованы для обработки клеток *in vitro*, с введением модифицированных клеток субъектам (*ex vivo*). Обычные вирусные системы доставки ZFP включают ретровирусные, лентивирусные, аденовирусные, аденоассоциированные, осповакцинные и герпетические вирусные векторы для генного переноса, но не ограничиваются ими. Интеграция в геном хозяина возможна с помощью методов генного переноса ретровирусами, лентивирусами и аденоассоциированными вирусами, часто приводя к длительной экспрессии вставленного трансгена. Кроме того, высокая эффективность трансдукции наблюдается во множестве различных типов клеток и тканей-мишеней. Тропизм ретровируса может быть модифицирован путем введения чужеродных оболочечных белков, что расширяет потенциальную целевую популяцию клеток-мишеней. Лентивирусные векторы являются ретровирусными векторами, которые способны трансдуцировать или инфицировать неделящиеся клетки и, как правило, продуцируют высокие вирусные титры. Выбор ретровирусной системы переноса генов зависит от ткани-мишени. Ретровирусные векторы состоят из цис-действующих длинных концевых повторов с упаковкой вместимостью до 6-10 т.н. чужеродной последовательности. Минимальные цис-действующие LTR являются достаточными для репликации и упаковки векторов, которые затем используются для интеграции терапевтического гена в клетку-мишень, чтобы обеспечить постоянную экспрессию трансгена. Широко используемые ретровирусные векторы включают те, которые основаны на вирусе мышинного лейкоза (MuLV), вирусе лейкоза гиббонов (Galv), вирусе иммунодефицита обезьян (SIV), вирусе иммунодефицита человека (ВИЧ), а также их комбинации (см., например, Buchscher et al., J. Virol. 66:2731-2739 (1992); Johann et al., J. Virol. 66:1635-1640 (1992); Sommerfelt et al., Virol. 176:58-59 (1990); Wilson et al., J. Virol. 63:2374-2378 (1989); Miller et al., J. Virol. 65:2220-2224 (1991); PCT/US 94/05700).

В сферах применения, для которых предпочтительна временная экспрессия, могут быть использованы аденовирусные системы. Аденовирусные векторы способны с очень высокой эффективностью трансдуцировать многие типы клеток и не требуют деления клеток. С такими векторами могут быть достигнуты высокие титры и высокие уровни экспрессии. Этот вектор может быть получен в больших количествах в относительно простой системе. Векторы аденоассоциированных вирусов ("AAV") также используются для трансдукции клеток целевыми нуклеиновыми кислотами, например, при *in vitro* продуцировании нуклеиновых кислот и пептидов, и процедурах генной терапии *in vivo* и *ex vivo* (см., например, West et al., Virology 160:38-47 (1987); U.S. Patent № 4797368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). Конструирование рекомбинантных векторов AAV раскрыто в ряде публикаций, включая патент США № 5173414; Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); и Samulski et al., J. Virol. 63:03822-3828 (1989).

В настоящее время доступны по меньшей мере шесть подходов на основе вирусных векторов для переноса генов в клинических испытаниях, в которых используются подходы, включающие комплементацию дефектных векторов с помощью генов, вставленных в линии клеток-помощников для создания трансдукционного агента.

pLASN и MFG-S являются примерами ретровирусных векторов, которые были использованы в клинических испытаниях (Dunbar et al., Blood 85:3048-305 (1995); Kohn et al., Nat. Med. 1:1017-102 (1995); Malech et al., PNAS 94:22 12133-12138 (1997)). PA317/pLASN был первым терапевтическим вектором, используемым в испытании генной терапии. (Blaese et al., Science 270:475-480 (1995)). Эффективность трансдукции 50% или более была обнаружена для упакованных векторов MFG-S. (Ellem et al., Immunol Immunother. 44(1):10-20 (1997); Dranoff et al., Hum. Gene Ther. 1:111-2 (1997). Рекомбинантные аденоассоциированные вирусные векторы (rAAV) являются перспективными альтернативными системами доставки генов, базирующимися на дефектном и непатогенном типе 2 парвовирусного аденоассоциированного вируса. Все векторы получены из плазмиды, которая сохраняет только 145 п.о. инвертированных концевых повторов AAV, фланкирующих кассету экспрессии трансгена. Эффективная передача генов и стабильная доставка трансгена благодаря интеграции в геномы трансдуцированной клетки являются основными характеристиками этой векторной системы. (Wagner et al., Lancet 351:9117 1702-3 (1998), Kearns et al., Gene Ther. 9:748-55 (1996)). Кроме того, другие серотипы AAV, включая AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV9 и AAVrh10 и все их варианты, могут быть использованы в соответствии с настоящим изобретением. Репликационно-дефицитные рекомбинантные аденовирусные векторы (Ad) могут быть получены с высокими титрами и легко инфицируют несколько различных типов клеток. Большинство аденовирусных векторов разработано таким образом, что трансген заменяет гены Ad Ela, Elb, и/или E3; затем репликация дефектного вектора распространяется в человеческих клетках 293, которые восполняют функцию удаленных генов в *trans*. Векторы Ad могут трансдуцировать различные типы тканей в живом организме, включая неделящиеся, дифференцированные клетки, такие как найдены в печени, почках и мышцах. Обычным векторам Ad присущ большой переносимый объем. Пример исполь-

зования вектора Ad в клинических испытаниях включает полинуклеотидную терапию для противоопухолевой иммунизации внутримышечной инъекцией (Sternan et al., Hum. Gene Ther. 7:1083-9 (1998)). Дополнительные примеры использования аденовирусных векторов для переноса генов в клинических испытаниях включают Rosenecker et al., Infection 24:1 5-10 (1996); Sternan et al., Hum. Gene Ther. 9:7 1083-1089 (1998); Welsh et al., Hum. Gene Ther. 2:205-18 (1995); Alvarez et al., Hum. Gene Ther. 5:597-613 (1997); Topf et al., Gene Ther. 5:507-513 (1998); Sternan et al., Hum. Gene Ther. 7:1083-1089 (1998).

Упаковка клеток используется для формирования вирусных частиц, способных инфицировать клетку-хозяина. Такие клетки включают клетки 293, которые упаковывают аденовирус, и клетки ψ 2 или клетки PA317, которые упаковывают ретровирус. Вирусные векторы, используемые в генной терапии, как правило, продуцируются клеточной линией, которая упаковывает вектор нуклеиновой кислоты в вирусную частицу. Векторы обычно содержат минимальные вирусные последовательности, необходимые для упаковки и последующей интеграции в хозяина (если доступен), другие вирусные последовательности заменяются кассетой экспрессии, кодирующей белок, подлежащий экспрессии. Недостающие вирусные функции поставляются в транс-пакующей клеточной линией. Например, векторы AAV, используемые в генной терапии, как правило, обладают только последовательностями инвертированного концевой повтора (ITR) из генома AAV, которые необходимы для упаковки и интеграции в геном хозяина. Вирусная ДНК упаковывается в клеточную линию, которая содержит вспомогательную плазмиду, кодирующую другие гены AAV, а именно *тер* и *сар*, но лишённые последовательности ITR. Клеточная линия также инфицируется аденовирусом в качестве помощника. Вспомогательный вирус способствует репликации вектора AAV и экспрессии генов AAV из плазмиды-помощника. Плазида-помощник не упаковывается в значительных количествах из-за отсутствия последовательностей ITR. Контаминация аденовирусом может быть уменьшена, например, тепловой обработкой, к которой аденовирус более чувствителен, чем AAV.

Во многих сферах применения генной терапии, желательно, чтобы вектор для генной терапии доставлялся с высокой степенью специфичности к конкретному типу ткани. Соответственно, вирусный вектор может быть модифицирован таким образом, чтобы обладать специфичностью в отношении данного типа клеток, путем экспрессии лиганда в виде слитого белка с вирусным белком оболочки на внешней поверхности вируса. Лиганд выбирают таким образом, чтобы он обладал аффинностью к рецептору, который, как известно, присутствует в клеточном типе, представляющем интерес. Например, Han et al., Proc. Natl. Acad. Sci. USA 92:9747-9751 (1995), сообщали, что вирус мышинного лейкоза Молони может быть модифицирован таким образом, чтобы экспрессировать человеческий херегулин, слитый с gp70, причем рекомбинантный вирус инфицирует определенные раковые клетки молочной железы человека, экспрессирующие человеческий рецептор эпидермального фактора роста. Этот принцип может быть распространен на другие пары клеток-мишеней вирусов, в которых клетка-мишень экспрессирует рецептор, а вирус экспрессирует слитый белок, содержащий лиганд для рецептора клеточной поверхности. Например, нитчатый фаг может быть сконструирован таким образом, чтобы воспроизводить фрагменты антител (например, Fab или Fv), обладающие аффинностью специфичного связывания, с практически любым выбранным клеточным рецептором. Хотя приведенное выше раскрытие относится, прежде всего, к вирусным векторам, те же самые принципы могут быть применены к невирусным векторам. Такие векторы могут быть сконструированы таким образом, чтобы содержать последовательности специфичного поглощения, которые способствуют поглощению определенными клетками-мишенями.

Векторы генной терапии могут доставляться *in vivo* путем введения отдельному субъекту, как правило, путем системного введения (например, внутривенного, внутривентриального, внутримышечного, подкожного или внутричерепной инфузии) или местного применения, как раскрыто ниже. С другой стороны, векторы могут доставляться в клетки *ex vivo*, такие как клетки, эксплантированные от индивидуального пациента (например, лимфоциты, пунктаты костного мозга, ткани биопсии) или универсальных доноров кроветворных стволовых клеток, с последующей реимплантацией клеток пациенту, как правило, после отбора клеток, которые содержат инкорпорированный вектор.

Кроме того, векторы (например, ретровирусы, аденовирусы, липосомы и т.п.), содержащие нуклеазы и/или донорские конструкции, могут быть введены непосредственно в организм для трансдукции клеток в живом организме. Кроме того, может быть введена "оголенная" ДНК. Введение осуществляют любым путем, обычно используемым для приведения молекулы в максимальный контакт с клетками крови или ткани, включая инъекцию, инфузию, местное применение и электропорацию, но не ограничиваясь ими. Подходящие способы введения таких нуклеиновых кислот доступны и хорошо известны специалистам в данной области техники, и, хотя может быть использован более чем один путь для введения конкретной композиции, конкретный путь часто может обеспечить в большей степени немедленную и в большей степени эффективную реакцию, чем другой путь. Векторы, пригодные для введения полинуклеотидов, раскрытых в настоящем документе, включают неинтегрирующие лентивирусные векторы (IDLV). См., например, Ory et al. (1996) Proc. Natl. Acad. Sci. USA 93:11382-11388; Dull et al. (1998) J. Virol. 72:8463-8471; Zuffery et al. (1998) J. Virol. 72:9873-9880; Follenzi et al. (2000) Nature Genetics 25:217-222; патентную публикацию № 2009/054985.

Фармацевтически приемлемые носители частично определяется конкретной композицией, подле-

жащей введению, а также конкретным способом, используемым для введения композиции. Таким образом, существует большое разнообразие подходящих составов доступных фармацевтических композиций, как раскрыто ниже (см., например, Remington's Pharmaceutical Sciences, 17th ed., 1989).

Очевидно, что кодирующие нуклеазу последовательности и донорские конструкции могут быть доставлены с использованием одних и тех же или разных систем. Например, донорский полинуклеотид может доставляться с помощью плазмиды, в то время как одна или более нуклеаз могут доставляться с помощью вектора AAV. Кроме того, различные векторы могут быть введены одним и тем же или разными путями (внутримышечно, инъекцией в хвостовую вену, другим видом внутривенной инъекции, внутрибрюшинным введением и/или внутримышечной инъекцией). Векторы могут быть доставлены одновременно или в любом последовательном порядке. Таким образом, настоящее изобретение включает *in vivo* или *ex vivo* лечение заболеваний и состояний, которые восприимчивы к вставке трансгенов, кодирующих терапевтический белок, например, лечение гемоглобинопатии путем нуклеазопосредованной интеграции гена, кодирующего белок глобина. Композиции вводят пациенту-человеку в количестве, эффективном для получения желаемой концентрации терапевтического полипептида в сыворотке или органе-мишени или клетках. Введение может осуществляться любыми способами, которыми полинуклеотиды доставляются в желательные клетки-мишени. Например, предполагаются *in vivo* и *ex vivo* способы. Внутривенное введение в воротную вену является предпочтительным способом введения. Другие способы введения *in vivo* включают, например, непосредственную инъекцию в доли печени или желчный проток и внутривенную инъекцию дистальнее печени, в том числе через печеночную артерию, непосредственную инъекцию в паренхиму печени, инъекцию в печеночную артерию и/или ретроградную инъекцию через желчные протоки. *Ex vivo* способы введения включают трансдукцию *in vitro* резецированных гепатоцитов или других клеток печени, с последующей инфузией трансдуцированных, резецированных гепатоцитов обратно в портальную васкулатуру, паренхиму печени или желчные протоки пациента-человека, см., например, Grossman et al., (1994) *Nature Genetics*, 6:335-341. Эффективное количество нуклеазы(нуклеаз) и донора, подлежащее введению, будет варьировать от пациента к пациенту и в зависимости от терапевтического полипептида, представляющего интерес. Соответственно, эффективные количества будут наилучшим образом определяться врачом, вводящим композиции, и соответствующие дозировки могут быть легко определены специалистом в данной области техники. После предоставления достаточного времени для интеграции и экспрессии (обычно 4-15 дней, например), анализ уровней терапевтического полипептида в сыворотке или других тканях и сравнение с исходным уровнем до введения будет определять, является ли введенное количество слишком низким, в целевом диапазоне или слишком высоким. Подходящие схемы для начального введения и последующего введения также варьируют, но типичным образом для них служит начальное введение с последующим введением, если это необходимо. Последующее введение может осуществляться через разные интервалы, начиная с ежедневного до ежегодного или каждые несколько лет. Специалисту в данной области техники будет понятно, что подходящие иммуносупрессивные методы могут быть рекомендованы во избежание ингибирования или блокирования трансдукции путем иммуносупрессии векторов доставки, см., например, Vilquin et al., (1995) *Human Gene Ther.* 6:1391-1401.

Препараты для *ex vivo* и *in vivo* введения включают жидкие суспензии или эмульгированные жидкости. Активные ингредиенты часто смешивают с эксципиентами, которые являются фармацевтически приемлемыми и совместимыми с активным ингредиентом. Подходящие наполнители включают, например, воду, физиологический раствор, декстрозу, глицерин, этанол или т.п. и их комбинации. Кроме того, композиция может содержать небольшие количества вспомогательных веществ, таких как, увлажняющие или эмульгирующие агенты, pH буферные агенты, стабилизирующие агенты или другие реагенты, которые повышают эффективность фармацевтической композиции.

Применение.

Способы и композиции, раскрытые в настоящем документе, предназначены для модификации экспрессии белка или коррекции aberrантной геновой последовательности, которая кодирует белок, экспрессирующийся при генетическом заболевании, таком как серповидно-клеточная анемия или талассемия. Таким образом, предложены способы и композиции для лечения и/или профилактики таких генетических заболеваний. Геномное редактирование, например, стволовых клеток, используется для коррекции aberrантного гена, вставки гена дикого типа или модификации экспрессии эндогенного гена. В качестве неограничивающего примера ген дикого типа, например, кодирующий по меньшей мере один глобин (например, α - и/или β -глобин), может быть вставлен в клетки, чтобы обеспечить глобиновые белки, недостающие и/или отсутствующие в клетке, и, таким образом, лечить генетическое заболевание, например гемоглобинопатию, вызванную патологической экспрессией глобина. Альтернативно или в дополнение, геномное редактирование, с введением соответствующего донора или без него, может исправить патологический эндогенный ген, например, корректируя точечную мутацию в α - или β -гемоглобине, восстановить экспрессию гена и/или лечить генетическое заболевание, например серповидно-клеточную анемию, и/или нокаутировать модификацию (избыточную экспрессию или репрессию) любого прямого или косвенного регуляторного гена глобина (например, инактивация регулирующего γ -глобин гена BCL11A или

BCL11A-регулятора KLF1).

Способы и композиции согласно изобретению также могут быть использованы в любых обстоятельствах, где желательно обеспечить трансген, кодирующий один или более терапевтических агентов, причем терапевтический агент продуцируется в RBC и/или кроветворной стволовой клетке таким образом, что зрелые эритроциты на основе этих клеток содержат терапевтический агент.

Следующие примеры относятся к примерным вариантам реализации настоящего изобретения, в которых нуклеаза содержит нуклеазу "цинковых пальцев" (ZFN) или TALEN. Следует иметь в виду, что эти примеры приведены только для иллюстративных целей и что могут быть использованы другие нуклеазы, например "хоминг"-эндонуклеазы (мегануклеазы) со сконструированными ДНК-связывающими доменами и/или слияния природных или сконструированных "хоминг"-эндонуклеазных (мегануклеазных) ДНК-связывающих доменов и гетерологичных доменов расщепления и/или системы CRISPR/Cas, содержащей сконструированную единичную гидовую РНК.

Примеры

Пример 1. Разработка, конструирование и общая характеристика нуклеаз белков "цинковые пальцы" (ZFN).

Белки "цинковые пальцы" были разработаны и введены в плазмиды, AAV или аденовирусные векторы, в основном, как раскрыто Urnov et al. (2005) Nature 435 (7042):646-651, Perez et al. (2008) Nature Biotechnology 26(7):808-816, и как раскрыто в патенте США № 6534261. Для ZFN и TALEN, специфичных в отношении локуса человеческого β -глобина и локуса человеческого HPRT, см. патент США № 7888121 того же заявителя и патентные публикации США № 20130137104 и 20130122591. Относительно нуклеаз, специфичных для человеческого AAVS1, см. патент США № 8110379. Относительно нуклеаз, специфичных для CCR5, см. патент США № 7951925. Относительно нуклеаз, специфичных для альбумина, см. патентные публикации США № 20130177983 и 20139177968.

Пример 2. Активность глобин-специфичных ZFN.

Пары ZFN, ориентированные на локус человеческого глобина или регуляторы экспрессии гена β -подобного глобина, использовали для проверки способности этих ZFN индуцировать DSB в определенном сайте-мишени. Аминокислотные последовательности спиральных областей распознавания каждого пальца указанных ZFN представлены в табл. 1А вместе с полноразмерными сайтами-мишенями (сайты-мишени ДНК указаны заглавными буквами; неконтактные нуклеотиды указаны в нижнем регистре).

Таблица 1А

Нуклеазы "цинковых пальцев"

SBS #,мишень	Конструкция					
ZFN, специфичные для человеческого В-гемоглобина						
	F1	F2	F3	F4	F5	F6
SBS#33511 <u>ggGCAGTAACGGC</u> <u>AGACttctcctca</u> <u>gg</u> (SEQ ID NO:8)	DRSNLSR (SEQ ID NO:9)	QSSDLRR (SEQ ID NO:10)	RSDTLSA (SEQ ID NO:11)	QSGALAR (SEQ ID NO:12)	QSGDLTR (SEQ ID NO:13)	нет данных
SBS#33533 <u>tgGGGCAAGGTGA</u> <u>ACGTGGAtgaagt</u> <u>tg</u> (SEQ ID NO:14)	QSAHRKN (SEQ ID NO:15)	LKHHLTD (SEQ ID NO:16)	QRSNLVR (SEQ ID NO:17)	TSGHLNR (SEQ ID NO:18)	QSNHLTE (SEQ ID NO:19)	RSHHLKA (SEQ ID NO:20)

SBS#35256 agAGTCAGGTGCA CCATggtgtctgt tt (SEQ ID NO:21)	TNQNRIT (SEQ ID NO:22)	DRSNRRT (SEQ ID NO:23)	RNASRTR (SEQ ID NO:24)	RSDNLSE (SEQ ID NO:25)	RSQHRKT (SEQ ID NO:26)	нет данных
SBS#35263 gtGGAGAAAGTctG CCGTtactgccct gt(SEQ ID NO:27)	TSGSLSR (SEQ ID NO:28)	DRSDLSR (SEQ ID NO:29)	DRSALAR (SEQ ID NO:30)	QSSNLAR (SEQ ID NO:31)	QSGHLSR (SEQ ID NO:32)	нет данных
SBS#34770 acAGGAGTCAGGT GCACcatggtgtc tg(SEQ ID NO:33)	DQSNLRA (SEQ ID NO:34)	RNASRTR (SEQ ID NO:24)	RSDNLSE (SEQ ID NO:25)	RSQHRKT (SEQ ID NO:26)	RSDHLTQ (SEQ ID NO:35)	нет данных
SBS#34791 gaGAAGTctGCCG TTACTgccctgtg gg(SEQ ID NO:36)	ARSTRTN (SEQ ID NO:37)	TSGSLSR (SEQ ID NO:28)	DRSDLSR (SEQ ID NO:29)	DRSARTR (SEQ ID NO:38)	QSGNLAR (SEQ ID NO:39)	нет данных
SBS#34805 taACGGCAGActT CTCCAcaggagtc ag (SEQ ID NO:40)	QSGDLTR (SEQ ID NO:13)	SSSDRKK (SEQ ID NO:41)	DRSNLSR (SEQ ID NO:9)	QSADRTK (SEQ ID NO:42)	RSDTLTA (SEQ ID NO:11)	нет данных
SBS#34826 gcCCTGTGGGGCA AGGTgaacgtgga tg(SEQ ID NO:43)	LRHHLTR (SEQ ID NO:44)	QSGNLHV (SEQ ID NO:45)	RSAHLSR (SEQ ID NO:46)	RSDVLST (SEQ ID NO:47)	RKQDLRT (SEQ ID NO:48)	нет данных
SBS#35301 ggGCAGTAACGGC AGACTtctctcctca gg(SEQ ID NO:8)	DRSNLSR (SEQ ID NO:9)	QSGDLTR (SEQ ID NO:13)	RSDTLTA (SEQ ID NO:11)	QSGALAR (SEQ ID NO:12)	QSGDLTR(SEQ ID NO:13)	нет данных
SBS#35328 tgGGGCAAGGTGA ACGTggatgaagt tg(SEQ ID NO:14)	MSHHLRD (SEQ ID NO:49)	QRSNLVR (SEQ ID NO:17)	TSGHLSR (SEQ ID NO:18)	QSNHLTE (SEQ ID NO:19)	RSHHLKA (SEQ ID NO:20)	нет данных
SBS#35497 caCAGGGCAGTAA CgGCAGACTtctc ct(SEQ ID NO:50)	DRSNLSR (SEQ ID NO:9)	QSGDLTR (SEQ ID NO:13)	DRSNLSR (SEQ ID NO:9)	LKHHLTD (SEQ ID NO:16)	DRSHLTR (SEQ ID NO:51)	RSDNLRE (SEQ ID NO:52)
SBS#35506 ggCAAGGTGAACG TGGAtgaagttgg tg(SEQ ID NO:53)	QSGHLAR (SEQ ID NO:54)	VSHHLRD (SEQ ID NO:55)	QSGNLAR (SEQ ID NO:39)	LRHHLTR (SEQ ID NO:44)	QSGNLHV (SEQ ID NO:45)	нет данных
бета-глобин IVS.1						
SBS#43545 atCAAGGTTACAA GACAGGTTtaagg ag(SEQ ID NO:158)	LRHHLTR (SEQ ID NO:44)	QSGTRKT (SEQ ID NO:153)	RSDNLST (SEQ ID NO:154)	DSANRIK (SEQ ID NO:155)	LRHHLTR (SEQ ID NO:44)	QSGNLHV (SEQ ID NO:45)

SBS#43544 aaTCTGCCCAGGG CCTCaccaccaac tt (SEQ ID NO:159)	AMQTLRV (SEQ ID NO:156)	DRSHLAR (SEQ ID NO:76)	RSDNLSE (SEQ ID NO:25)	ASKTRKN (SEQ ID NO:77)	RNSDRTK (SEQ ID NO:157)	нет данных
ZFN, специфичные для человеческого BCL11A						
Экзон 2						
SBS#39172 ctCCAGAAGGGGA TCATGACctcctc ac (SEQ ID NO:160)	DRSNLSR (SEQ ID NO:9)	LRQNLIM (SEQ ID NO:161)	TSANLTV (SEQ ID NO:162)	RSDHLSR (SEQ ID NO:94)	QSGNLAR (SEQ ID NO:39)	QRNDRKS (SEQ ID NO:163)
SBS#43490 ctCCAGAAGGGGA TCATGACctcctc ac (SEQ ID NO:160)	DRSNLSR (SEQ ID NO:9)	LRQNLIM (SEQ ID NO:161)	LQSQLNR (SEQ ID NO:164)	RSDHLSR (SEQ ID NO:94)	QSGNLAR (SEQ ID NO:39)	QRNDRKS (SEQ ID NO:163)
SBS#44642 ctCCAGAAGGGGA TCATGACctcctc ac (SEQ ID NO:160)	DRANLSR (SEQ ID NO:165)	LRQNLIM (SEQ ID NO:161)	LQSQLNR (SEQ ID NO:164)	RSDHLSR (SEQ ID NO:94)	QSGNLAR (SEQ ID NO:39)	QRNDRKS (SEQ ID NO:163)
SBS#45148 ctCCAGAAGGGGA TCATGACctcctc ac (SEQ ID NO:160)	DRSNLSR (SEQ ID NO:9)	TSSNRNH (SEQ ID NO:166)	HSGNLTK (SEQ ID NO:167)	RSDHLSR (SEQ ID NO:94)	QSGNLAR (SEQ ID NO:39)	QKVDLSR (SEQ ID NO:168)
SBS#45147 ctCCAGAAGGGGA TCATGACctcctc ac (SEQ ID NO:160)	DRSNLSR (SEQ ID NO:9)	TSSNRNH (SEQ ID NO:166)	QANNLKV (SEQ ID NO:169)	RSDHLSR (SEQ ID NO:94)	QSGNLAR (SEQ ID NO:39)	QKVDLSR (SEQ ID NO:168)
SBS#39145 ccCAACGGGCCGT GGTCTGGttcatc at (SEQ ID NO:170)	RSDHLSA (SEQ ID NO:59)	DRSALAR (SEQ ID NO:30)	RSDSLSR (SEQ ID NO:171)	DRSVRTK (SEQ ID NO:172)	RSDHLSA (SEQ ID NO:59)	QRSNLKV (SEQ ID NO:173)
SBS#44490 ccCAACGGGCCGT GGTCTGGttcatc at (SEQ ID NO:170)	RSDHLTQ (SEQ ID NO:35)	DRSALAR (SEQ ID NO:30)	RSDSLSR (SEQ ID NO:171)	DRSVRTK (SEQ ID NO:172)	RSDHLSA (SEQ ID NO:59)	QRSNLKV (SEQ ID NO:173)
SBS#44489 ccCAACGGGCCGT GGTCTGGttcatc at (SEQ ID NO:170)	RSDHLTT (SEQ ID NO:174)	DRSALAR (SEQ ID NO:30)	RSDSLSR (SEQ ID NO:171)	DRSVRTK (SEQ ID NO:172)	RSDHLSA (SEQ ID NO:59)	QRSNLKV (SEQ ID NO:173)
SBS#45081 ccCAACGGGCCGT GGTCTGGttcatc at (SEQ ID NO:170)	RSDHLSA (SEQ ID NO:59)	WATARDR (SEQ ID NO:175)	RSDSLSR (SEQ ID NO:171)	HTKSLSR (SEQ ID NO:176)	RSDHLSA (SEQ ID NO:59)	QRSNLKV (SEQ ID NO:173)
SBS#44493 ccCAACGGGCCGT GGTCTGGttcatc at (SEQ ID NO:170)	RSAHLTQ (SEQ ID NO:177)	DRSVLRR (SEQ ID NO:178)	RSDSLSR (SEQ ID NO:171)	DRSVRTK (SEQ ID NO:172)	RSDHLSA (SEQ ID NO:59)	QRSNLKV (SEQ ID NO:173)
SBS#29527 atCCCATGGAGAG GTGGCTGggaagg	RSDVLSE (SEQ ID NO:57)	RNQHRRKT (SEQ ID NO:58)	RSDHLSA (SEQ ID NO:59)	RSANLTR (SEQ ID NO:60)	RSDVLSN (SEQ ID NO:61)	DRSTRIT (SEQ ID NO:62)

ac (SEQ ID NO:56)						
SBS#29528 atATTGCAGACAA TAACccctttaac ct (SEQ ID NO:63)	DRSNLSR (SEQ ID NO:9)	HRQHLVT (SEQ ID NO:64)	DRSNLTR (SEQ ID NO:65)	QSGDLTR (SEQ ID NO:13)	HRSSLN (SEQ ID NO:66)	нет данных
SBS# 29525 caTCCAGGCGTG GGGAttagagctc ca (SEQ ID NO:66)	QSGHLSR (SEQ ID NO:32)	RSDHLST (SEQ ID NO:67)	RSADLSR (SEQ ID NO:68)	RSDNLSQ (SEQ ID NO:69)	ASNDRKK (SEQ ID NO:70)	нет данных
SBS#29526 gtGCAGAATATGC CCCGCAGgtatt tg (SEQ ID NO:71)	RSDNLSA (SEQ ID NO:72)	RNDRKT (SEQ ID NO:73)	DRSDLSR (SEQ ID NO:29)	TSSNRTK (SEQ ID NO:74)	QSGNLAR (SEQ ID NO:39)	QSGDLTR (SEQ ID NO:13)
Экзон 4						
SBS#34678 atATTGCAGACAA TAACccctttaac ct (SEQ ID NO:179)	DRSNLSR (SEQ ID NO:9)	HRQHLVT (SEQ ID NO:64)	DRSNLTR (SEQ ID NO:65)	QSGDLTR (SEQ ID NO:13)	HRWLRSN (SEQ ID NO:180)	нет данных
SBS#34642 atCCCATGgAGAG GTGGCTGGgaagg ac (SEQ ID NO:56)	RSDHLSQ (SEQ ID NO:99)	DSSHRTR (SEQ ID NO:181)	LRHHLTR (SEQ ID NO:44)	QSAHLKA (SEQ ID NO:182)	RSDVLSN (SEQ ID NO:61)	DRSTRIT (SEQ ID NO:62)
Bcl11a-XL						
SBS#44889 ctCACTGTCCACA GGAGaagccacac gg (SEQ ID NO:183)	RSANLAR (SEQ ID NO:184)	RLDNRTA (SEQ ID NO:185)	QSNLNS (SEQ ID NO:186)	WRSSLKT (SEQ ID NO:187)	DRSNRKT (SEQ ID NO:188)	нет данных
SBS#44888 ttGCTACAGTTCT TGAAGACTttccc ac (SEQ ID NO:189)	DRSNLSR (SEQ ID NO:9)	QSGNLAR (SEQ ID NO:39)	YKHLSD (SEQ ID NO:190)	TSGSLTR (SEQ ID NO:191)	QSGDLTR (SEQ ID NO:13)	LKDTLRR (SEQ ID NO:192)
SBS#44905* gaGAAGCCACACG GGCGAAAgcctt at (SEQ ID NO:193)	QSGNLDS (SEQ ID NO:194)	RSADLSR (SEQ ID NO:68)	RSDHLS (SEQ ID NO:78)	QNATRIN (SEQ ID NO:195)	WNSDLRK (SEQ ID NO:196)	QSGNLAR (SEQ ID NO:39)
SBS#44904* tgGACAGTGAGAT TGCTacagttctt ga (SEQ ID NO:197)	QSSDLR (SEQ ID NO:88)	YKWTLRN (SEQ ID NO:198)	RSANLTR (SEQ ID NO:60)	TSTKLRT (SEQ ID NO:199)	DRSNLTR (SEQ ID NO:65)	нет данных
SBS#44911 gcCACACGGGCGA AaGGCCTTataaa tg (SEQ ID NO:200)	AMQTLRV (SEQ ID NO:156)	DRSHLAR (SEQ ID NO:76)	QRSNLVR (SEQ ID NO:17)	DRSHLAR (SEQ ID NO:76)	RSDTLST (SEQ ID NO:201)	DSSNRIN (SEQ ID NO:202)
SBS#44910 ctCCTGTGGACAG TGAGATTgctaca gt (SEQ ID NO:203)	NDLFLYL (SEQ ID NO:204)	RSANLTR (SEQ ID NO:60)	TSTKLRT (SEQ ID NO:199)	DRSNLTR (SEQ ID NO:65)	RSDSLSV (SEQ ID NO:205)	HNSRKN (SEQ ID NO:206)

SBS#44945 aaGCTCACCAGGC ACATGAAaacgca tg (SEQ ID NO:207)	QSGNLR (SEQ ID NO:39)	CRQNLAR (SEQ ID NO:208)	YQGVLTR (SEQ ID NO:209)	RSDNLRE (SEQ ID NO:52)	DRSNRTT (SEQ ID NO:23)	HRSSLRR (SEQ ID NO:210)
SBS#44944 ctACTCTGgGCAC AGGCATAGttgca ca (SEQ ID NO:211)	RSDNLST (SEQ ID NO:154)	QSSDLRR (SEQ ID NO:10)	RSDALSE (SEQ ID NO:212)	QNATRTK (SEQ ID NO:115)	RSDTLSE (SEQ ID NO:84)	ARSTRTN (SEQ ID NO:37)
SBS#44947** ctCACCAGGCACA TGAAACGcatgg cc (SEQ ID NO:213)	GSSALTQ (SEQ ID NO:214)	QSGNLR (SEQ ID NO:39)	TASHLKE (SEQ ID NO:215)	QNATRTK (SEQ ID NO:115)	RSDNLSE (SEQ ID NO:25)	SSRNLAS (SEQ ID NO:216)
SBS#44946** ctACTCTGgGCAC AGGCATAGttgca ca (SEQ ID NO:211)	RSDNLST (SEQ ID NO:154)	QSSDLRR (SEQ ID NO:10)	RSDALSE (SEQ ID NO:212)	QNATRTK (SEQ ID NO:115)	RSDTLSE (SEQ ID NO:84)	ARSTRTN (SEQ ID NO:37)
ZFN, специфичные для человеческого KLF1						
KLF- Экзон1						
SBS#36004 ggGAAGGGGCCCA GGGCGGTcagtgt gc (SEQ ID NO:75)	TSGHLR (SEQ ID NO:18)	DRSHLR (SEQ ID NO:76)	RSDNLSQ (SEQ ID NO:69)	ASNDRKK (SEQ ID NO:70)	RSDHLS (SEQ ID NO:78)	QSGNLR (SEQ ID NO:39)
SBS#36021 acACACAGGATGA Cttcctcaagggtg gg (SEQ ID NO:79)	DRSNLTR (SEQ ID NO:65)	TSANLSR (SEQ ID NO:217)	RSDHLS (SEQ ID NO:78)	QSASRKN (SEQ ID NO:81)	нет данных	нет данных
SBS#33237 <u>ggGAAGGGGCCCA</u> <u>GGGCGGTcagtgt</u> gc (SEQ ID NO:75)	TSGHLR (SEQ ID NO:18)	DRSHLR (SEQ ID NO:76)	RSDNLSE (SEQ ID NO:25)	ASKTRKN (SEQ ID NO:77)	RSDHLS (SEQ ID NO:78)	QSGNLR (SEQ ID NO:39)
SBS#33238 <u>acACACAGGATGA</u> <u>Cttcctcaagggtg</u> gg (SEQ ID NO:79)	DRSNLTR (SEQ ID NO:9)	TSGNLTR (SEQ ID NO:80)	RSDHLS (SEQ ID NO:78)	QSASRKN (SEQ ID NO:81)	нет данных	нет данных
SBS#33257 <u>cgCCACCGGCTC</u> <u>CGGGccccgagaag</u> tt (SEQ ID NO:82)	RSALSR (SEQ ID NO:46)	DSSDRKK (SEQ ID NO:83)	DRSHLR (SEQ ID NO:76)	RSDTLSE (SEQ ID NO:84)	QSGDLTR (SEQ ID NO:13)	нет данных
SBS#33258 <u>cccCAGACcTGCG</u> <u>CTCTGCGCccag</u> cg (SEQ ID NO:85)	RSDSLR (SEQ ID NO:86)	RLDWLPV (SEQ ID NO:87)	QSSDLR (SEQ ID NO:88)	AASNRSK (SEQ ID NO:89)	DRSNLTR (SEQ ID NO:9)	QSGDLTR (SEQ ID NO:13)
SBS#33269 <u>ggCTCGGGgGCCG</u> <u>GGGCTGGAgccag</u> gg	QSSHLTR (SEQ ID NO:91)	QSSDLTR (SEQ ID NO:92)	RSDHLS (SEQ ID NO:78)	HSRTRTK (SEQ ID NO:93)	RSDHLR (SEQ ID NO:94)	DRSARNS (SEQ ID NO:95)

(SEQ ID NO:90)						
SBS#33270 aaGGCGCTGGCGC <u>TgCAACCGgtgta</u> cc (SEQ ID NO:96)	RSDTLSE (SEQ ID NO:84)	QSHNRK (SEQ ID NO:97)	QSSDLSR (SEQ ID NO:88)	DRSHLAR (SEQ ID NO:76)	QSSDLSR (SEQ ID NO:88)	DRSHLAR (SEQ ID NO:76)
SBS#33271 ttGCAGCGCCAGC <u>GCCTTGGgctcgg</u> gg (SEQ ID NO:98)	RSDHLSQ (SEQ ID NO:99)	HRSSLGD (SEQ ID NO:100)	RSDDLTR (SEQ ID NO:101)	QRSTLSS (SEQ ID NO:102)	RSADLTR (SEQ ID NO:103)	QSGDLTR (SEQ ID NO:13)
SBS#33272 cgGTGTACCCGGG <u>GCCCGgcgcgcgc</u> tc (SEQ ID NO:104)	DRSDLSR (SEQ ID NO:29)	RSTHLVR (SEQ ID NO:105)	RSDSLST (SEQ ID NO:106)	DSSDRTK (SEQ ID NO:107)	RSAALAR (SEQ ID NO:108)	нет данных
KLF-Экзон 2						
SBS#36071 ggTGAGGAGGAGA TCCAggtcccagg tg (SEQ ID NO:218)	NNRDLIN (SEQ ID NO:219)	TSSNLSR (SEQ ID NO:220)	QSGHLSR (SEQ ID NO:32)	QSGHLAR (SEQ ID NO:54)	QRTHLNS (SEQ ID NO:221)	нет данных
SBS#36085 ctTCTCGGGCCCG GaGCCCGGtggcg cg (SEQ ID NO:222)	RSDHLSQ (SEQ ID NO:78)	HSRTRTK (SEQ ID NO:93)	RSDHLSQ (SEQ ID NO:78)	HSRTRTK (SEQ ID NO:93)	RSDHLSQ (SEQ ID NO:78)	RKSDRIK (SEQ ID NO:223)
ZFN 5' регуляторной области человеческого гамма-глобина						
регуляторная область (-175)						
SBS#34360 ttGCATTGAGATA GTGTGGGgaagg gc (SEQ ID NO:109)	RSDHLSV (SEQ ID NO:110)	RSDVRKT (SEQ ID NO:111)	RSDYLSK (SEQ ID NO:112)	TSSVVRT (SEQ ID NO:113)	RPYTLRL (SEQ ID NO:114)	QNATRTK (SEQ ID NO:115)
SBS#34363 atCTGTCTGAAAC GGTCcctggctaa ac (SEQ ID NO:116)	DRSALAR (SEQ ID NO:30)	RRDILHQ (SEQ ID NO:117)	QSGNLAR (SEQ ID NO:39)	LAYDRRK (SEQ ID NO:118)	RSDVLSE (SEQ ID NO:57)	нет данных
SBS#34398 ttTGCAATTGAGAT AGTGtggggaagg gg (SEQ ID NO:119)	RSDSLLR (SEQ ID NO:86)	QSCARNV (SEQ ID NO:120)	RSDNLAR (SEQ ID NO:121)	HRNTLLG (SEQ ID NO:122)	MRNRLNR (SEQ ID NO:123)	нет данных
SBS#34400 ctGTCTGAaACGG TCcCTGGCTaaac tc (SEQ ID NO:124)	QSSDLSR (SEQ ID NO:88)	RRDALLM (SEQ ID NO:125)	DRSALAR (SEQ ID NO:30)	RRDILHQ (SEQ ID NO:117)	QNAHRKT (SEQ ID NO:126)	DRSALAR (SEQ ID NO:30)
SBS#31160 taTTTGCAtTGAG ATAGTGTGgggaa gg (SEQ ID NO:127)	RSDSLLR (SEQ ID NO:86)	LQHHLLTD (SEQ ID NO:128)	TSGNLTR (SEQ ID NO:80)	TSTHLHI (SEQ ID NO:129)	QSGDLTR (SEQ ID NO:13)	HKWVLRQ (SEQ ID NO:130)

SBS#34365 ctGTCTGAaACGG TCcCTGGCTaaac tc (SEQ ID NO:124)	QSSDLR (SEQ ID NO:88)	RRDALM(SEQ ID NO:131)	DRSALAR (SEQ ID NO:30)	RRDILHQ (SEQ ID NO:117)	QNAHRKT(SEQ ID NO:126)	DRSALAR (SEQ ID NO:30)
регуляторная область (-110)						
SBS#34539 tgGTCAAGGCAAG GCTGgccaaccca tg (SEQ ID NO:224)	RSDVLSE (SEQ ID NO:57)	RNQHRKT (SEQ ID NO:58)	QSGDLTR (SEQ ID NO:13)	RSDHLST (SEQ ID NO:67)	DRSALAR(SEQ ID NO:30)	нет данных
SBS#34574 gcCTTGACAAGGC AAACttgaccaat ag (SEQ ID NO:225)	DRSNRTT (SEQ ID NO:23)	QSGSLTR (SEQ ID NO:226)	RSDNLSV (SEQ ID NO:227)	DRSNLSR (SEQ ID NO:9)	LKFALAN(SEQ ID NO:228)	нет данных
SBS#43865 gcCTTGACAAGGC AAACttgaccaat ag (SEQ ID NO:225)	NPANLTR (SEQ ID NO:229)	QNATRTK (SEQ ID NO:115)	RSDNLSV (SEQ ID NO:227)	DRSNLSR (SEQ ID NO:9)	LKFALAN(SEQ ID NO:228)	нет данных
SBS#43852 tgGTCAAGGCAAG GCTGgccaaccca tg (SEQ ID NO:224)	RSDVLSE (SEQ ID NO:57)	RNQHRKT (SEQ ID NO:58)	QSGDLTR (SEQ ID NO:13)	RSDNLSV (SEQ ID NO:154)	DSSARKK(SEQ ID NO:230)	нет данных

Примечание: BCL11A XL-специфичные пары ZFN, обозначенные одной звездочкой (*) или двумя звездочками (**), содержат новые линкеры L7a и L8p, соответственно. См. пример 6.

Анализ Cel-I (Surveyor™, Transgenomics), как раскрыто Perez et al. (2008) Nat. Biotechnol. 26: 808-816 и Guschin et al. (2010) Methods Mol. Biol. 649:247-56), был использован для выявления ZFN-индуцированных модификаций гена-мишени в K562 или HSC. В этом анализе PCR-амплификация сайта-мишени сопровождалась количественной оценкой вставок и делеций (инсерционно-делеционные мутации) с использованием фермента Cel-I, выявляющего ошибки (Yang et al. (2000) Biochemistry 39: 3533-3541), который обеспечивает оценку нижнего предела частоты DSB. Через три дня после трансфекции вектора экспрессии ZFN при стандартных условиях (37°C) или с использованием гипотермического шока (30°C, см. патентную публикацию США № 20110041195 того же заявителя), выделили геномную ДНК из клеток K562, используя набор DNeasy (Qiagen).

Результаты анализа Cel-I демонстрируют, что ZFN были способны индуцировать расщепление в их соответствующих сайтах-мишенях (см. также предварительную патентную заявку США № 61/556691 того же заявителя).

Результаты проиллюстрированы на фиг. 1 и демонстрируют, что активные белки были найдены у большинства локусов-мишеней в гене β-глобина.

Пример 3. Редактирование β-глобинового локуса.

ZFN, специфичные для гена человеческого β-глобина (HBB) (табл. 1), были использованы для введения донорской ДНК в β-глобиновый локус следующим образом. Донорская ДНК была разработана таким образом, что последовательность, кодирующая генные последовательности HBB, фланкировалась последовательностями, которые были гомологичными (гомологичные плечи) области, окружающей сайт расщепления ZFN в β-глобиновом гене. Длина гомологичных плеч составляет приблизительно 500-600 пар оснований. Донорская последовательность HBB не содержит каких-либо некодирующих последовательностей, поэтому при вставке в β-глобиновый сайт-мишень экспрессия донора регулируется β-глобиновым промотором и любыми другими регуляторными последовательностями β-глобина. При вставке донор HBB сливается в рамке с эндогенными глобиновыми последовательностями и дает слитый белок. Кроме того, олигодонор HBB был разработан для захвата в расщепленном гене HBB после обработки ZFN. Олиго содержал сайт рестрикции, таким образом, что после вставки олиго был введен новый сайт рестрикции в HBB-ген, который впоследствии можно расщепить.

Как проиллюстрировано на фиг. 2A, β-глобиновый олигодонор был вставлен в надлежащий локус, что проверено наличием нового сайта рестрикции, присутствующего на донорской ДНК. Кроме того, как проиллюстрировано на фиг. 2B, анализ Cel-I продемонстрировал, что несколько пар ZFN были способны расщеплять ДНК, хотя олиго присутствовал только в образце полосы 8.

Использовали методы дифференцировки трансгенных клеток CD34⁺ в зрелые RBC, известные из уровня техники. Например, клетки SCD CD34⁺ очищали с помощью фиколл-пака (GE Healthcare) и микросфер CD34⁺ (Miltenyi Biotec) согласно инструкциям производителя. Клетки CD34⁺ культивировали в среде Дульбекко, модифицированной по способу Исков, с BIT 95000 (StemCell Technologies) в присутствии факторов роста. Клетки дифференцировали в эритроидную клеточную линию с использованием модели 2-фазной жидкой культуры. В течение первых 6 дней (первая фаза), клетки CD34⁺ выращивали с SCF (100 нг/мл), Flt3-L (100 нг/мл) и IL-3 (20 нг/мл). Выращенные клетки затем коммитировали и дифференцировали в эритроидную клеточную линию (вторая фаза) с помощью Epo (2 Ед/мл) и SCF (50

нг/мл). См., Giarratana et al. (2011) Blood 118(19):5071-9.

Пример 4. Генная коррекция мутаций в β -глобине.

Для корректировки серповидно-клеточных мутаций в гене серповидного β -глобина был осуществлен двухцепочечный разрыв в β -глобиновом локусе с помощью ZFN, с последующим восстановлением ДНК, с использованием экзогенного корректирующего олигонуклеотида в качестве матрицы ("олигодонор"). Чтобы избежать вероятности нуклеаз, расщепляющих исправленный ген глобина (ген, в котором олигодонор направил коррекцию серповидной мутации в эндогенном гене HBB), был разработан олигодонор для совместной вставки трансляционно молчащих мутаций в кодирующую последовательность HBB таким образом, чтобы откорректированным аллелям не хватало одной из последовательностей-мишеней ZFN. Таким образом, наблюдалось увеличение частоты откорректированных аллелей целевых генов. Для разработки оптимального олигонуклеотидного донора были исследованы некоторые мутации в последовательности-мишени ZFN, а также длина гомологичных плеч.

Ниже показана последовательность, окружающая серповидную мутацию, а разные мутации обозначены цифрами. Таким образом, мутация 1 = G на A, мутация 2 = G на A, мутация 3 = TCT на AGC, мутация 4 = C на T и мутация 5 = T на G. Были получены олигонуклеотиды, которые содержали различные комбинации мутаций. Последовательность дикого типа ("wt") указана вверх (SEQ ID NO: 231), а последовательность с мутациями ("mut") указана ниже (SEQ ID NO: 232).

Сайты-мишени для нуклеаз ("target") обозначены жирными линиями, а сайт серповидной мутации взят в прямоугольник. Олигонуклеотиды обозначены в соответствии с мутациями, таким образом, например, олигонуклеотид SMS1 содержит только сайт 1 молчащей мутации, тогда как SMS 124 содержит сайты 1, 2 и 4 молчащих мутаций.

```

target      _____
              |
ACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT wt
ACCATGGTGCATCTGACTCCTGAGGAAAAAGCGCTGTG mut
              1 2 3      4 5

```

Разные олиго были доставлены в клетки CD34⁺ в виде одноцепочечных молекул в качестве "смысловых" или прямых цепей (обозначается как "F") или "антисмысловых" или обратных цепей (обозначается как "R"). Олиго доставляются путем трансфекции с помощью устройства Square Wave BTX ECM 830 в присутствии или в отсутствие нуклеаз. Если не указано иное, то доставляется 3 мкг нуклеаз. Генное редактирование было измерено с помощью высокоэффективного секвенирования ДНК ампликонов PCR гена HBB. Процент генной модификации определяли путем негомологичного соединения концов ("NHEJ", вызванное заживлением двухцепочечного разрыва в ДНК после ZFN-индуцированного расщепления) или целенаправленной интеграции олиго после расщепления ZFN ("генная коррекция") (см. фиг. 9). Результаты указывают, что некоторые комбинации мутаций были способны повышать генную коррекцию в клетках таким образом, что до 20% клеток демонстрировали генную коррекцию в серповидном локусе. Для исследования влияния длины гомологичных плеч на процент генной коррекции были использованы олиго SMS12 и SMS124 с 41 и 46 нуклеотидами (88 п.о. олигодонора) или 50 и 50 нуклеотидами гомологии (101 п.о. олигодонора) на каждой стороне сайта серповидной мутации. Результаты (см. фиг. 10) указывают на то, что более длинные гомологичные плечи были более эффективными при вызове генной коррекции, причем до 40% аллелей содержали модификации, определенные с помощью олиго. Используемые олиго показаны ниже

SMS124, 88 п.о., R (SEQ ID NO: 233):

```

5' CGTTCACCTTGCCCCACAGGGCAGTAACAGCAGATTTTTCCTCAGGAGTCAGGTGCACCA
TGGTGTCTGTTTGAGGTTGCTAGTGAAC

```

SMS12, 88 п.о., R (SEQ ID NO: 234):

```

5' CGTTCACCTTGCCCCACAGGGCAGTAACGGCAGATTTTTCCTCAGGAGTCAGGTGCACCA
TGGTGTCTGTTTGAGGTTGCTAGTGAAC

```

SMS124, 101 п.о., R (SEQ ID NO: 235):

```

5' CTTTCATCCACGTTTCACCTTGCCCCACAGGGCAGTAACAGCAGATTTTTCCTCAGGAGTCA
GGTGCACCATGGTGTCTGTTTGAGGTTGCTAGTGAACACAG

```

Для исследования дифференцировочной способности и долговечности генной коррекции во время дифференцировки клеток CD34⁺, пулы ZFN-модифицированных клеток CD34⁺ побуждали дифференцировать, используя метилцеллюлозную среду Methocult Stemcell Technologies' согласно инструкциям производителя. Дифференцировку анализировали с помощью анализа типа колоний, возникающих вследствие дифференцировки, вызванной Methocult: колониеобразующие единицы, эритроид ("CFU-E"); бурстобразующие единицы, эритроид ("BFU-E"); колониеобразующие единицы, гранулоцит/макрофаг ("CFU-GM") и колониеобразующие единицы; гранулоцит/эритроцит/моноцит/макрофаг ("CFU-GEMM").

Результаты указывают на то, что ZFN-обработанные клетки сохраняют ту же способность к дифференцировке, что и мимотрансфектированные клетки. Индивидуальные колонии BFU-E были собраны с чашки и генотипированы по поводу HBB. Результаты указывают на то, что ZFN-индуцированные модификации сохранялись в течение дифференцировки колоний (см. фиг. 11). Кроме того, частота модифицированных BFU-E колоний была подобна частоте модифицированных аллелей в стартовом пуле, демонстрируя, что не наблюдается смещения относительно отредактированных клеток во время образования BFU-E. Сверх того, клеточная популяция в целом была исследована по поводу генной модификации в ходе дифференцировки красных кровяных клеток во время культивирования в жидкой среде *in vitro*. Модификации были стабильными в течение по меньшей мере 18 дней процесса дифференцировки красных кровяных клеток (см. фиг. 12). Другая общая мутация в β -глобиновом гене, которая ассоциируется с β -талассемией, известна как IVS 1.1. Эта G→A мутация локализуется в пределах первой пары оснований интрона 1 β -глобинового гена, а ее присутствие в гене приводит к патологическому сплайсингу β -глобиновой пре-мРНА. Таким образом, была сконструирована пара ZFN для распознавания и расщепления области, существенно рекапитулирующая эту мутацию для модельных целей. Тестирование этих ZFN обнаружило, что они способны расщеплять сайт на β -глобиновом гене, приводя к 52,63% NHEJ в клетках CD34⁺.

Пример 5. Вставка β -глобинового донора в локус "безопасной гавани".

Для вставки β -глобинового гена дикого типа в локус "безопасной гавани" таким образом, чтобы экспрессия из трансгена корректировала дефицит β -глобина в HSC, нуклеазы, специфичные для этого локуса "безопасной гавани", вводятся в клетку наряду с донорской нуклеиновой кислотой. Нуклеазы, специфичные для HPRT (см. патентные публикации США № 20130137104 и 20130122591 того же заявителя), AAVS1 (см. патент США 8110379), CCR5 (см. патент США 7951925) или β -глобину (см. табл. 1A) вводятся в полученные от пациента стволовые клетки CD34⁺. Введение может осуществляться любым методом, известным из уровня техники, таким как электропорация мРНК. Разрабатывается донорская ДНК, содержащая трансген, β -глобин дикого типа и области гомологии, фланкирующие трансген, с достаточной гомологией с областью, окружающей "безопасную гавань"-мишень, чтобы позволить HDR (как правило, 500 п.о. на каждой стороне). В качестве альтернативы может быть предложена донорская конструкция, таким образом, что, независимо от того содержит ли она области гомологии, она вводится в ZFN или TALEN-намеченный локус путем концевой захвата (см. заявку США № 13/889162). Донор совместно вводится в клетку CD34⁺ перед, во время или после введения ZFN. Модифицированные клетки CD34⁺ повторно вводятся пациенту, и после приживления вырабатывают достаточные уровни β -гемоглобина для обеспечения терапевтически релевантного количества гемоглобина, который будет вырабатываться.

Пример 6. Инактивация BCL11A и KLF1.

Нуклеазы, специфичные для BCL11A и KLF1 (например, ZFN, как показано в табл. 1A), вводили в HSC, как раскрыто выше, чтобы вызвать положительную регуляцию γ -глобиновой экспрессии (см. фиг. 3), а геном клеток анализировали с помощью анализа Cel, как раскрыто выше (Perez et al. (2008), там же).

Как показано на фиг. 4, после обработки HSC указанными KLF1-специфичными ZFN, ZFN успешно модифицировали локус KLF1 (фиг. 4C и 4D). Аналогично, BCL11A-специфичные ZFN модифицировали локус BCL11A (фиг. 4A). Пара ZFN, целенаправленно действующая на локус HPRT (см. предварительную заявку США № 61/552309 того же заявителя), была использована в качестве контроля и также продемонстрировала успешное расщепление (фиг. 4B). Сравнение сигнала в день 3 после CD34⁺ клеточной трансдукции с днем 17 дифференцировки культуры (фиг. 4E) продемонстрировало, что процент генетического редактирования (% NHEJ) является стабильным с течением времени. В каждом геле, показанном на фиг. 4E, полосы с отсутствующей идентификацией являются отрицательным контролем.

Подобным образом были проверены дополнительные пары ZFN, целенаправленно действующие на экзон 2 или экзон 4 BCL11A. Для этих исследований пары-кандидаты ZFN вводили в клетки K562 с помощью Амаха, как раскрыто выше, или вводили в клетки CD34⁺. Для трансдукции CD34⁺ было использовано устройство BTX ECM830 с кюветой с 2 мм зазором. мРНК из клеток получали с использованием mMessageMachine T7 Ultra Kit (# AM1345, Ambion). Человеческие клетки CD34⁺ выращивали в среде x-vivo10 (Lonza) с 1xCC110 (Stem cell Technology) в планшетах, обработанных нетканевой культурой. Клетки подсчитывали и собирали центрифугированием при 1200 оборотах в минуту в течение 10 мин при комнатной температуре. Клетки промывали 1-2 раза в фосфатно-солевом буфере (PBS) при комнатной температуре. 200000 клеток были использованы для каждой трансфекции, их ресуспендировали в 100 мкл раствора VTexpress. Добавляли по 2-4 мкг мРНК на трансфекцию, и смесь переносили в кювету. Сразу же после переноса смесь подвергали электропорации при 250 В в течение 5 мс. Предварительно подогретую среду добавляли в кювету, и среду плюс клетки переносили на 48-луночные планшеты, обработанные нетканевой культурой, а затем инкубировали при 37°C.

После указанного количества дней клетки повторно подвергались геномному анализу с использованием Illumina MiSeq. Для количественного определения процента отредактированных аллелей целевую геномную область подвергали PCR-амплификации с использованием праймеров, которые добавляют

стандартные Illumina секвенирующие адаптерные последовательности. Вторую группу из 13 раундов PCR проводили, чтобы добавить штрих-код и адаптерные мостовые последовательности к обоим концам. Секвенирование проводили на Illumina MiSeq в соответствии с протоколами производителя для секвенирования ампликонов. MiSeq генерирует соединенные парноконцевые считываемые фрагменты и подобранный адаптер с использованием стандартного программного обеспечения для выравнивания. Считываемые фрагменты затем демультиплексируются образцом с помощью пары последовательностей штрих-кодов с использованием обычных сценариев. Последовательности ампликонов затем повсеместно выравнивали с эталонной последовательностью с применением алгоритма Needleman-Wunsch (Needleman, Saul B.; и Wunsch, Christian D. (1970). *Jour Mol Bio* 48 (3): 443-53). Брешы или вставки в выравнивании подсчитывали как % NHEJ событий и сравнивали с необработанной контрольной последовательностью для определения скорости счета последовательность-специфичного фона.

Для расчета направленной интеграции последовательности ампликонов были выровнены повсеместно с эталонной последовательностью с помощью "биопитонного" применения алгоритма Needleman-Wunsch (Needleman, Saul B.; и Wunsch, Christian D., там же). Был осуществлен поиск, подсчет и сравнение модификаций последовательностей, генерируемых с помощью экспериментальной обработки, с подсчетами в контрольных образцах. Известные однопризнаковые полиморфизмы (SNP) могут маскироваться во время этого процесса и исключаться из дальнейшего подсчета (например, делеция 1 п.о. SNP, близких к сайту-мишени ZFN). Процент NHEJ (также упоминается как инсерционно-делеционная мутация) рассчитывали путем определения процентной доли последовательностей, которые содержат вставки или делеции. Образцы, обработанные только вектором GFP, были использованы для оценки ошибок PCR и секвенирования на основе фоновой частоты вставок и делеций. Наблюдались фоновые частоты, составляющие менее 1%. Набор репрезентативных данных показан ниже в табл. 1B, и демонстрирует, что эти нуклеазные белки принимают активное участие в расщеплении своих мишеней. Кроме того, экспрессия γ -глобина контролировалась в некоторых клетках, обработанных нуклеазой. Чтобы выполнить этот анализ, использовали количественную PCR с обратной транскриптазой (RT-qPCR) в режиме реального времени ("Taqman") в качестве стандартной процедуры (см. ниже). Результаты из набора репрезентативных данных отображаются как кратность повышения экспрессии γ -глобина по сравнению с контрольными клетками, обработанными GFP. Значения γ рассчитывали как отношение γ -глобина к α -глобину, таким образом, что любой наблюдаемый рост, показанный ниже, представляет собой увеличение соотношения γ к α в клетках, обработанных нуклеазой, по сравнению с соотношением γ к α в клетках, обработанных вектором GFP.

Таблица 1B

Активность BCL11A экзона 2 и экзона 4 пар ZFN

Мишень	Пара ZFN	% инсерционно-делеционных мутаций, K562	% инсерционно-делеционных мутаций, CD34+	Кратность увеличения содержания гамма мРНК
Экзон 2	39145/39172		69,78	3,65X
	39145/43490	19,88		не определяли
	39145/44642	38,52		не определяли
	39145/41548	42,26		не определяли
	39145/41547	35,63		не определяли
	44490/39172	29,38		не определяли
	44489/39172	24,34		не определяли
	45081/39172	27,80		не определяли
	44493/39172	25,68		не определяли
Экзон 4	34678/34642		82,24	3,52X

TALEN были также созданы как для области экзона 2, так и для области экзона 4 BCL11A. TALEN были сконструированы, как раскрыто выше, с помощью канонического кода TALE и скелета "+17" TALEN (см. патентную публикацию № 20110301073 того же заявителя). в табл. 1C приведена последовательность-мишень для TALEN, а также последовательность RVD в ДНК-связывающем домене.

Таблица 1C

Пары TALEN против BCL11A

Номер SBS (экзон)	Последовательность-мишень 5'→3'	Последовательность RVD (N→C)
101291 (экзон 2)	ctGTGGGCGAGTGCCAGATga (SEQ ID NO:236)	NN-NG-NN-NN-NN-HD-NI-NN-NG-NN-HD-HD-NI-NN-NI-NG (SEQ ID NO:237)
101292	ctCGATAAAAAATAAGAAATgt (SEQ ID NO:238)	HD-NN-NI-NG-NI-NI-NI-NI-NI-NG-NI-NI-NN-NI-NI-NG (SEQ ID NO:239)

(экзон 2)		
101301 (экзон 4)	atGTCCTTCCCAGCCACCTct (SEQ ID NO:240)	NN-NG-HD-HD-NG-NG-HD-HD-HD-NI-NN- HD-HD-NI-HD-HD-NG (SEQ ID NO:241)
101304 (экзон 4)	gtTAAAGGGTTATGTct (SEQ ID NO:242)	NG-NI-NI-NI-NN-NN-NN-NN-NG-NG-NI- NG-NG-NN-NG (SEQ ID NO:243)

Пары TALEN, показанные выше, были введены в клетки и продемонстрировали расщепляющую активность. Пара 101291/101292 дала 0,8% инсерционно-делеционных мутаций, по данным анализа Cel-1 в клетках K562. Пара TALEN 101301/101304 дала 35,7% инсерционно-делеционных мутаций в клетках CD34+, и по данным анализа RT-PCR, раскрытого выше, вызывала повышение экспрессии γ -глобиновой мРНК приблизительно в 2,31 раза.

Пары ZFN также были получены для целенаправленного воздействия на часть "XL" в сплайс-варианте BCL11A-XL. Эти белки проверялись на клетках K562, набор репрезентативных данных показан ниже в табл. 1D. Изоформа "XL" BCL11A содержит 3 дополнительных природных "цинковых пальца" (пальцы 4-6), таким образом, подход включал разрушение гена BCL11A в этой области, чтобы вызвать разворачивание потенциально "цинковых пальцев" 4, 5 и/или 6 и их комбинаций (номера от 1 до 3 в области XL). Также ZFN были разработаны, чтобы избежать расщепления родственной генной последовательности BCL11B. Одна пара ZFN, 44888/44889, направлена на четвертый "цинковый палец" BCL11A, тогда как две пары, 44904/44905 и 44910/44911, направлены вверх на четвертый палец (номер 1 в области XL), а две другие пары, 44946/44947 и 44945/44944, направлены на пятый палец (номер 2 в области XL). Эти белки проверяли на клетках K562, набор репрезентативных данных показан ниже в табл. 1D. Две из пар ZFN содержали новые линкерные последовательности между ДНК-связывающим доменом ZFP и нуклеазным доменом FokI. Пара 44904/44905 содержала линкерную последовательность L7a (патентную заявку США № 20090305419), а пара 44946/44947 содержала линкерную последовательность L8p, обе из которых приведены ниже. См. также предварительную заявку США № 61/871219:

L7a: HTKIHLRGSQVLVKSSEAAAR (SEQ ID NO: 244)

L8p: HTKIHLRGSYAPMPPLALASP (SEQ ID NO: 245).

Таблица 1D

Активность пар ZFN, специфичных для XL BCL11A

Пара ZFN	% инсерционно- делеционных мутаций, K562
44889/44888	35,14
44905/44904	25,45
44911/44910	36,43
44945/44944	24,03
44947/44946	34,22

Пары XL BCL11A, проверенные на клетках CD34+, являются активными. Измерение экспрессии γ -глобина показывает, что модификация XL BCL11A приводит к увеличению экспрессии γ -глобина относительно α -глобина.

Дополнительные пары KLF1-специфичных ZFN были проверены по поводу активности в клетках CD34+, а эти клетки были проанализированы по поводу каких-либо модификаций в экспрессии γ -глобина. Набор репрезентативных данных показан ниже в табл. 1E.

Таблица 1E

Активность KLF1-специфичных пар ZFN

Мишень	Пара ZFN	% инсерционно- делеционных мутаций, CD34+	Кратность увеличения содержания гамма мРНК
KLF экзон 1	36004/36021	44,4	2,2X
KLF экзон 2	36071/36085	22,6	3,17X

Соотношения мРНК, кодирующих γ -глобин и β -глобин, после обработки BCL11A- или KLF1-специфичными нуклеазами в HSC определяли в разные моменты времени до 17 дней после введения ZFN с помощью анализа Taqman, а уровни мРНК бетаподобного глобина также нормировали к уровню 18S rPFC. Уровни экспрессии γ -глобина повышались в тех клетках, которые были обработаны BCL11A- или KLF1-специфичными нуклеазами (фиг. 5). Анализ проводился согласно стандартному анализу Taqman в соответствии с протоколом, с использованием геноспецифичных анализов, поставляемых производителем (Applied Biosystems).

BCL11A ZFN-модифицированные клетки также анализировались для определения соотношения γ/β мРНК как между клеточными популяциями, в которых одна аллель была модифицирована с помощью ZFN ("Bb"), так и клетками, в которых обе аллели были модифицированы с помощью ZFN ("нокаутированные"), и диким типом ("BB").

Как показано на фиг. 6, соотношения γ/β мРНК отличаются между клетками, в которых нокаут BCL11A встречается только в одной аллели (Bb, полосы 6-10 слева) или когда нокаутированы обе аллели (нокаут, крайние справа 5 полос, полосы 11-15 слева), и оба пула клеток отличались от дикого типа (BB, первые 5 полос).

Пример 7. Модификация регуляторной области гена γ -глобина.

При другом подходе, чтобы увеличить экспрессию γ -глобина, мутации были введены в регуляторную область гена γ -глобина, чтобы имитировать мутации HPFH (см. фиг. 9). Ниже показана область от -202 до -102 относительно ATG в гене γ -глобина. На этой последовательности серые прямоугольники указывают области, для которых показана связь с HPFH, а подчеркнута последовательность, которая при удалении также была связана с HPFH (см. A Syllabus of Thalassemia Mutations (1997), Titus H.J. Huisman, Marianne F.H. Carver, и Erol Baysal, published by The Sickle Cell Anemia Foundation in Augusta, GA, USA. Copyright © 1997 by Titus H.J. Huisman): -202

CCCTTCCCCACACTATCTCAATGCCAAATATGTCTCTGAAACGGTCCCTGGCTAAACTCCACCCATGGGTT

GGCCTTGCCCTTGACCAATAGCCTTGAC (SEQ ID NO: 132)

-102

Нуклеазы были разработаны, как раскрыто в примере 1 и показано в табл. 1A, таким образом, чтобы связываться в области этих HPFH-ассоциированных мутаций и вызывать мутации в области дикого типа. Процент отредактированных аллелей, определенный (% NHEJ) в клетках K562 с помощью анализа Cel I (см. Perez et al. (2008), там же), показан ниже в табл. 2. Кроме того, некоторые пары были испытаны в клетках CD34+, как раскрыто выше, и проанализированы с помощью секвенирования MiSeq, как раскрыто выше. Для некоторых пар клетки анализировали по поводу любой модификации в экспрессии γ -глобина. В табл. 2 ниже показаны репрезентативные наборы данных.

Таблица 2

Редактирование пар ZFN, специфичных для γ -глобина

Пара ZFN (локализация)	% NHEJ K562	% NHEJ CD34+	Кратность увеличения содержания гамма мРНК
34360/34363 (-175)	39		
34398/34400 (-175)	54		
31160/34365 (-175)	53	45,22	1,63X
34539/34574 (-110)		45,71	5,38X
43865/43852 (-110)		56,13	

Первые три пары, проверенные в этом анализе, целенаправленно действовали на область вокруг -175 в γ -промоторной области, в то время как две последние целенаправленно действовали на область -110 в γ -глобиновом промоторе.

γ -Промоторную область в клетках K562, которая была отредактирована, секвенировали для анализа созданных мутаций. Область сначала PCR-амплифицировали, а затем PCR-продукты секвенировали, наблюдалось несколько различных мутаций, в том числе делеций и вставок (фиг. 8). В этом эксперименте 42% аллелей мутировало, а 20% несло делецию 13 п.о. от -114 до -102, ассоциированную с HPFH.

Две пары ZFN, направленных на γ -глобиновый промотор, также использовались для обработки клеток в комбинации с олигонуклеотидным донором, разработанным для воссоздания наиболее распространенных мутаций у субъекта с HPFH. Придерживались того же протокола, что раскрытый выше для применения с устройством ВТХ, с добавлением 3 мкл 100 мкМ раствора донорского олигонуклеотида. Последовательность олигонуклеотидных доноров показана ниже. Как правило, в этих экспериментах использовали прямой олигонуклеотидный донор, но и обратный работал хорошо:

HBG_d13 прямой:

acactatctcaatgcaaatactgtctgtgaaacgggtccctggctaaactccacccatg

ggttggccagccttgcccttgacaaggcaaaacttgaccaatagctcttagagtatccag

tgaggccagg (124mer, SEQ ID NO: 246)

HBG_d13_обратный:

cctggcctcactggatactctaagactatttggtcaagtttgccctgtcaaggcaagg

ctggccaacccatgggtggagtttagccagggaccggttcagacagatatttgcatg

gagatagtggt (124mer, SEQ ID NO: 247)

Для пары ZFN 34539/34574 в присутствии донора продуцирование мРНК из γ -глобинового гена повышается в 6,38 раз по сравнению с клетками, обработанными вектором GFP, тогда как для пары ZFN 31160/34365, содержание γ мРНК повышается в 6,13 раз по сравнению с клетками, обработанными вектором GFP.

Обработанные нуклеазами HSC высевали на метилцеллюлозу.

После генотипирования индивидуальных колоний с помощью PCR-секвенирования мы измеряли уровни мРНК для γ -глобина, β -глобина и 18s рРНК, контроля для дикого типа, а мутировавшие колонии - с помощью RT-PCR. (фиг. 8). В среднем, мутантам γ -глобинового промотора было свойственно более высокое соотношение γ -глобиновых мРНК к β -глобиновым мРНК, чем клеткам дикого типа, а коррекция со стороны сигнала 18s рРНК показывает, что увеличение соотношения γ -глобина/ β -глобина в мутировавших колониях вызвано повышением уровней γ -глобиновой мРНК в этих колониях, а не снижением уровней β -глобиновой мРНК.

Пример 8. Нуклеазы TALE, целенаправленно действующие на γ -глобиновый промотор.

Нуклеазы TALE также были получены для целенаправленного действия на область -200 или область -110 (раскрыты выше) промоторной области γ -глобина. TALEN были сконструированы, как раскрыто выше, с помощью канонического кода TALEN и скелета "+17" TALEN (см. патентную публикацию № 20110301073 того же заявителя).

Таблица 3

TALEN, специфичные для γ -глобинового промотора

Номер SBS	Последовательность 5'->3'	Последовательность RVD (N->C)
102314	gtATCCTCTTGGGGGcc (SEQ ID NO:133)	NI-NG-HD-HD-NG-HD-NG-NG-NN-NN-NN-NN-NK (SEQ ID NO:134)
102318	atATTGTCATTGAGATAGT gt (SEQ ID NO:135)	NI-NG-NG-NG-NN-HD-NI-NG-NG-NN-NI-NN-NI-NG-NI-NN-NG (SEQ ID NO:136)
102315	gtATCCTCTTGGGGGccc (SEQ ID NO:133)	NI-NG-HD-HD-NG-HD-NG-NG-NN-NN-NN-NN-NN-HD (SEQ ID NO:137)
102320	atATTGTCATTGAGATAg (SEQ ID NO:135)	NI-NG-NG-NG-NN-HD-NI-NG-NG-NN-NI-NN-NI-NG-NI (SEQ ID NO:136)
102316	gtATCCTCTTGGGGGCCcc (SEQ ID NO:133)	NI-NG-HD-HD-NG-HD-NG-NG-NN-NN-NN-NN-NN-HD-HD (SEQ ID NO:138)
102321	atATTGTCATTGAGATag (SEQ ID NO:135)	NI-NG-NG-NG-NN-HD-NI-NG-NG-NN-NI-NN-NI-NG (SEQ ID NO:139)
102566 (-110)	gtTGGCCAGCCTTGCCTTG ac (SEQ ID NO:248)	NG-NN-NN-HD-HD-NI-NN-HD-HD-NG-NG-NN-HD-HD-NG-NG-NK (SEQ ID NO:249)
102568 (-110)	ttGGTCAAGTTTGCCTTGT ca (SEQ ID NO:250)	NN-NN-NG-HD-NI-NI-NN-NG-NG-NG-NN-HD-HD-NG-NG-NN-NG (SEQ ID NO:251)

TALEN затем использовались в парах для тестирования расщепления в клетках K562 и анализировались с помощью анализа Cel 1, как раскрыто выше, а результаты пар показаны ниже в табл. 4. Кроме того, пара TALEN 102566/102568 тестировалась против клеток CD34+ и обнаружила 51,39% NHEJ, по данным анализа MiSeq.

Две пары TALEN также тестировались по поводу экспрессии γ -глобиновой мРНК, измеренной в виде соотношения гамма-глобиновой мРНК к альфа-глобиновой мРНК. Пара 102566/102568 увеличивала γ -глобиновую экспрессию в 6,25 раза по сравнению с клетками CD34+, обработанными вектором GFP, а пара 102318/102314 увеличивала γ -глобин в 2,14 раза по сравнению с клетками CD34+, обработанными вектором GFP. Пару 102566/102568 также проверяли с олигодонором, раскрытым выше, было обнаружено, что полученные клетки повышали γ -глобиновую экспрессию в 9,13 раз по сравнению с клетками CD34+, обработанными вектором GFP.

Таблица 4

Редактирование области γ -глобинового промотора с помощью TALEN

Пара TALEN	% NHEJ «+17»
102314:102318	41,6
102315:102320	47,9
102316:102321	46,6

Пример 9. γ -Глобиновое редактирование в стволовых клетках CD34+.

Нуклеазы, специфичные к γ -глобиновой промоторной области, затем использовали в клетках CD34+, полученных от пациента. Клетки обрабатывали нуклеазами, а затем анализировали по поводу

успешного редактирования с помощью анализа Cel 1. Клетки дополнительно были проанализированы по поводу соотношения γ -глобина к β -глобину и демонстрировали повышенную экспрессию γ -глобина. Репрезентативные данные, обнаруженные для повышенной экспрессии γ -глобина, находятся в экспериментальных параграфах для разных подходов выше.

Пример 10. Редактированный энграфмент CD34+ у мышей.

Обработанные нуклеазами клетки CD34+ (человеческие стволовые клетки-предшественники HSPC) сохранили способность приживляться у мышей NOD/SCID/IL2 γ (нуль) и приводить к появлению поликлонального мультилинейного потомства, у которого гены, участвующие в регуляции γ -глобина, постоянно разрушаются (см. Holt et al., (2010) Nat. Biotechnol. Aug; 28(8):839-47). Аналогично, CD34+ или HSPC, редактированные у локуса β -глобина, где мутация скорректирована, или донорский β -глобиновый ген вставлен в локус "безопасной гавани", или которые обработаны нуклеазами, чтобы изменить экспрессию γ -глобина, способны внедряться и приводить к мультилинейному потомству, несущему желательное геномное редактирование. Демонстрация того, что меньшинство отредактированных HSPC может заполнить отредактированное потомство животного, подтверждает использование модифицированных нуклеазой аутологичных кроветворных стволовых клеток в качестве клинического подхода к лечению гемоглобинопатии.

Все патенты, патентные заявки и публикации, упомянутые в настоящем документе, включены в настоящий документ посредством ссылки в полном объеме.

Хотя раскрытие изобретения дано более подробно путем иллюстрации и примера в целях ясности понимания, специалистам в данной области техники, будет понятно, что различные изменения и модификации могут быть осуществлены без отступления от сущности или объема раскрытия.

Соответственно, приведенные выше раскрытие и примеры не должны быть истолкованы в качестве ограничения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Белок "цинковые пальцы" для модулирования экспрессии гена BCL11A, содержащий 5 или 6 доменов "цинковых пальцев", каждый из которых содержит спиральную область распознавания, которая распознает сайт-мишень, причем домены "цинковых пальцев" расположены в порядке от F1 до F5 или от F1 до F6, и спиральные области распознавания этих доменов являются следующими:

(i) F1: DRSNLSR (SEQ ID NO: 9),
F2: LRQNLIM (SEQ ID NO: 161),
F3: TSANLTV (SEQ ID NO: 162),
F4: RSDHLSR (SEQ ID NO: 94),
F5: QSGNLAR (SEQ ID NO: 39) и
F6: QRNDRKS (SEQ ID NO: 163), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 160, или

(ii) F1: DRSNLSR (SEQ ID NO: 9),
F2: LRQNLIM (SEQ ID NO: 161),
F3: LQSQLNR (SEQ ID NO: 164),
F4: RSDHLSR (SEQ ID NO: 94),
F5: QSGNLAR (SEQ ID NO: 39) и
F6: QRNDRKS (SEQ ID NO: 163), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 160, или

(iii) F1: DRANLSR (SEQ ID NO: 165),
F2: LRQNLIM (SEQ ID NO: 161),
F3: LQSQLNR (SEQ ID NO: 164),
F4: RSDHLSR (SEQ ID NO: 94),
F5: QSGNLAR (SEQ ID NO: 39) и
F6: QRNDRKS (SEQ ID NO: 163), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 160, или

(iv) F1: DRSNLSR (SEQ ID NO: 9),
F2: TSSNRNH (SEQ ID NO: 166),
F3: HSGNLTK (SEQ ID NO: 167),
F4: RSDHLSR (SEQ ID NO: 94),
F5: QSGNLAR (SEQ ID NO: 39) и
F6: QKVDLSR (SEQ ID NO: 168), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 160, или

(v) F1: DRSNLSR (SEQ ID NO: 9),
F2: TSSNRNH (SEQ ID NO: 166),
F3: QANNLKV (SEQ ID NO: 169),
F4: RSDHLSR (SEQ ID NO: 94),

F5: QSGNLAR (SEQ ID NO: 39) и
 F6: QKVDLSR (SEQ ID NO: 168), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 160, или

(vi) F1: RSDHLSA (SEQ ID NO: 59),
 F2: DRSALAR (SEQ ID NO: 30),
 F3: RSDSLSR (SEQ ID NO: 171),
 F4: DRSVRTK (SEQ ID NO: 172),
 F5: RSDHLSA (SEQ ID NO: 59) и
 F6: QRSNLKV (SEQ ID NO: 173), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 170, или

(vii) F1: RSDHLTQ (SEQ ID NO: 35),
 F2: DRSALAR (SEQ ID NO: 30),
 F3: RSDSLSR (SEQ ID NO: 171),
 F4: DRSVRTK (SEQ ID NO: 172),
 F5: RSDHLSA (SEQ ID NO: 59) и
 F6: QRSNLKV (SEQ ID NO: 173), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 170, или

(viii) F1: RSDHLTT (SEQ ID NO: 174),
 F2: DRSALAR (SEQ ID NO: 30),
 F3: RSDSLSR (SEQ ID NO: 171),
 F4: DRSVRTK (SEQ ID NO: 172),
 F5: RSDHLSA (SEQ ID NO: 59) и
 F6: QRSNLKV (SEQ ID NO: 173), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 170, или

(ix) F1: RSDHLSA (SEQ ID NO: 59),
 F2: WATARDR (SEQ ID NO: 175),
 F3: RSDSLSR (SEQ ID NO: 171),
 F4: HTKSLSR (SEQ ID NO: 176),
 F5: RSDHLSA (SEQ ID NO: 59) и
 F6: QRSNLKV (SEQ ID NO: 173), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 170, или

(x) F1: RSAHLTQ (SEQ ID NO: 177),
 F2: DRSVLRR (SEQ ID NO: 178),
 F3: RSDSLSR (SEQ ID NO: 171),
 F4: DRSVRTK (SEQ ID NO: 172),
 F5: RSDHLSA (SEQ ID NO: 59) и
 F6: QRSNLKV (SEQ ID NO: 173), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 170, или

(xi) F1: RSDVLSE (SEQ ID NO: 57),
 F2: RNQHRKT (SEQ ID NO: 58),
 F3: RSDHLSA (SEQ ID NO: 59),
 F4: RSANLTR (SEQ ID NO: 60),
 F5: RSDVLSN (SEQ ID NO: 61) и
 F6: DRSTRIT (SEQ ID NO: 62), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 56, или

(xii) F1: DRSNLSR (SEQ ID NO: 9),
 F2: HRQHLVT (SEQ ID NO: 64),
 F3: DRSNLTR (SEQ ID NO: 65),
 F4: QSGDLTR (SEQ ID NO: 13) и
 F5: HRSSLLN (SEQ ID NO: 66), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 63, или

(xiii) F1: QSGHLSR (SEQ ID NO: 32),
 F2: RSDHLST (SEQ ID NO: 67),
 F3: RSADLSR (SEQ ID NO: 68),
 F4: RSDNLSQ (SEQ ID NO: 69) и
 F5: ASNDRKK (SEQ ID NO: 70), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 66, или

(xiv) F1: RSDNLSA (SEQ ID NO: 72),
 F2: RNNDRKT (SEQ ID NO: 73),
 F3: DRSDLSR (SEQ ID NO: 29),
 F4: TSSNRTK (SEQ ID NO: 74),
 F5: QSGNLAR (SEQ ID NO: 39) и

F6: QSGDLTR (SEQ ID NO: 13), причем белок "цинковые пальцы" связывается с сайтом-мишенью, представленным в SEQ ID NO: 71.

2. Слитый белок для инактивации гена BCL11A, содержащий белок "цинковые пальцы" по п.1 и нуклеазный домен расщепления дикого типа, или сконструированный домен расщепления, или полудомен расщепления дикого типа, или сконструированный полудомен расщепления.

3. Полинуклеотид, кодирующий белок по п.1 или 2.

4. Выделенная клетка, в которой ген BCL11A инактивирован, содержащая белок по п.2 или полинуклеотид по п.3.

5. Клетка по п.4, выбранная из кроветворной клетки-предшественника или красной кровяной клетки (RBC).

6. Клетка по п.5, отличающаяся тем, что кроветворной клеткой-предшественником является кроветворная стволовая клетка CD34+.

7. Набор для модулирования экспрессии гена BCL11A, содержащий белок по п.1 или 2 либо полинуклеотид по п.3 и подходящие клетки-хозяева.

8. Способ модулирования экспрессии глобинового гена в клетке, включающий введение в клетку белка по п.1 или 2 либо полинуклеотида по п.3 и обеспечение условий, в которых продуцируется белок по п.1 или 2 и модулируется экспрессия гена BCL11A, в результате чего модулируется экспрессия глобинового гена.

9. Способ по п.8, отличающийся тем, что экспрессия глобинового гена повышается.

10. Способ по п.8, отличающийся тем, что глобиновым геном является ген γ -глобина или β -глобина.

11. Способ по любому из пп.8-10, отличающийся тем, что клеткой является кроветворная клетка-предшественник красной кровяной клетки (RBC).

12. Способ п.11, отличающийся тем, что клеткой-предшественником RBC является кроветворная стволовая клетка CD34+.

13. Способ модулирования экспрессии глобинового гена в клетке *in vitro*, включающий введение в клетку белка по п.1 или 2 либо полинуклеотида по п.3 и обеспечение условий, в которых продуцируется белок по п.1 или 2 и модулируется экспрессия гена BCL11A, в результате чего модулируется экспрессия глобинового гена.

14. Кроветворная клетка-предшественник красной кровяной клетки (RBC), в которой экспрессия глобинового гена изменена способом согласно любому из пп.8-13.

15. Клетка-предшественник по п.14, содержащая одну или более инсерций и/или делеций.

16. Клетка-предшественник по п.14, которая представляет собой кроветворную стволовую клетку.

17. Клетка-предшественник по п.16, которая дифференцируется в красную кровяную клетку (RBC).

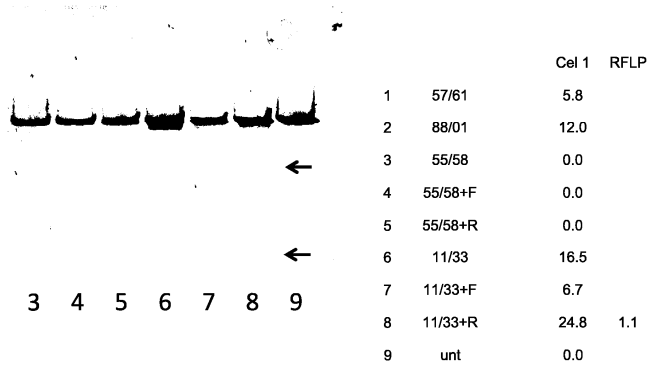
18. Клетка, происходящая от кроветворной клетки-предшественника RBC по п.14.



Лocus	% инсерционно-делеционных мутаций, исследования CD34	
	Эксперимент 1	Эксперимент 2
3	13,3	0,0
4	0,0	0,0
5	36,8	25,6
6	42,0	23,7
7	23,2	19,9

Фиг. 1

A. RFLP

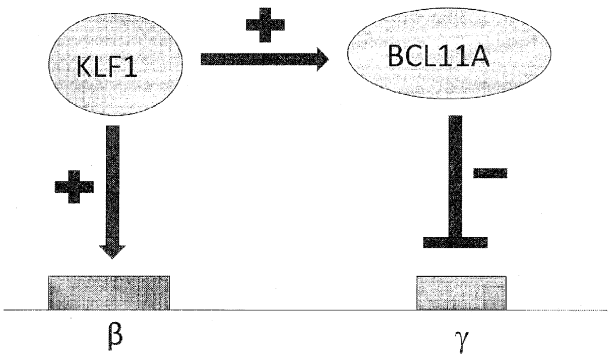


Фиг. 2

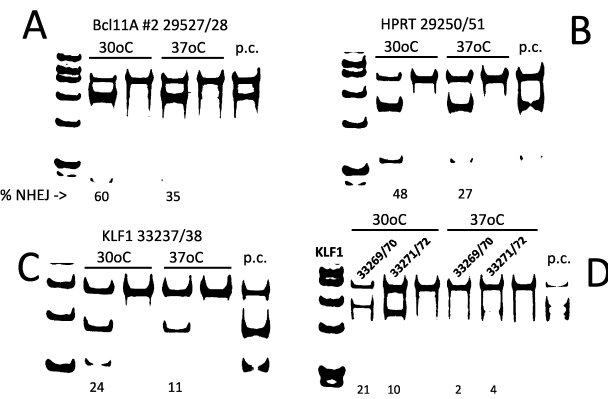
B. CEL-I



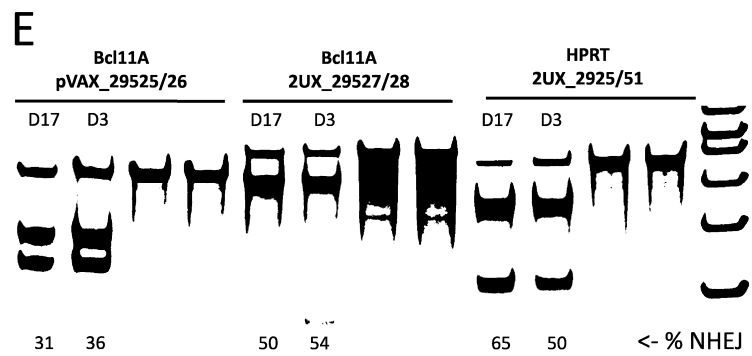
Фиг. 2



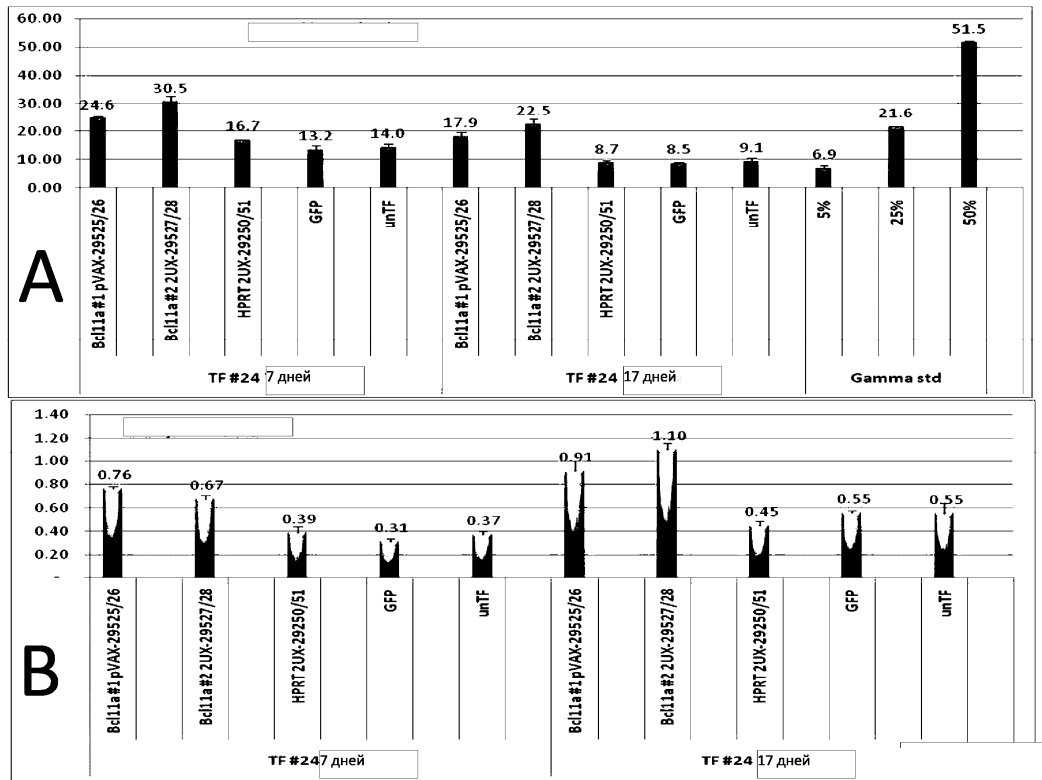
Фиг. 3



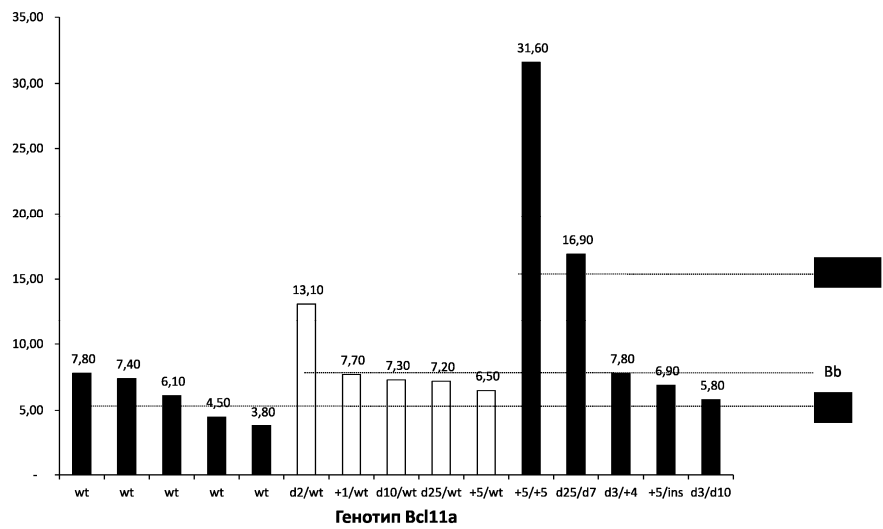
Фиг. 4



Фиг. 4



Фиг. 5



Фиг. 6

Эталон

GGCCAGCCTTGCCCTTGACCAATAGCCTTGACAAGGCAAACCTTGACCAATAG

wt (26)

GGCCAGCCTTGCCCTTGACCAATAGCCTTGACAAGGCAAACCTTGACCAATAG

 $\Delta 1bp$ (1)

GGCCAGCCTTGCCCTTGACC-ATAGCCTTGACAAGGCAAACCTTGACCAATAG

 $\Delta 4bp$ (1)

GGCCAGCCTTGCCCTT-----AATAGCCTTGACAAGGCAAACCTTGACCAATAG

 $\Delta 6bp$ (2)

GGCCAGCCTTGCCCTTGAC-----CCTTGACAAGGCAAACCTTGACCAATAG

 $\Delta 6bp$ (1)

GGCCAGCCTTGCCCTTGACCAATAGC-----AAGGCAAACCTTGACCAATAG

 $\Delta 13bp$ (9)

GGCCAGCCTTGCCCTTGAC-----AAGGCAAACCTTGACCAATAG

+1bp (1)

GGCCAGCCTTGCCCTTGACCAATAGCCTTGACAAGGCAAACCTTGACCAATA

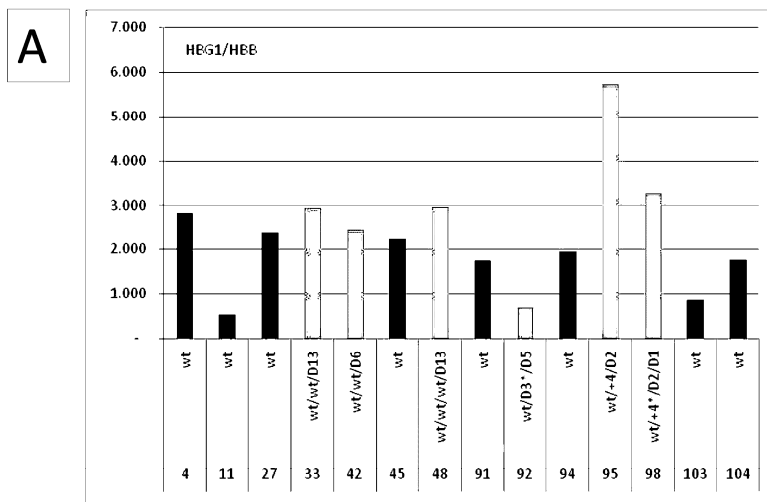
+3bp (1)

GGCCAGCCTTGCCCTTGACCAATAGCAGCCTTGACAAGGCAAACCTTGACCAA

+5bp (3)

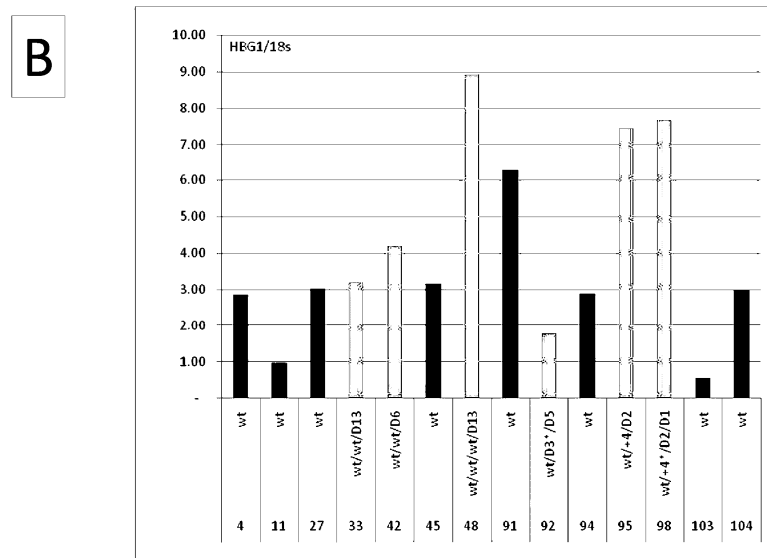
GGCCAGCCTTGCCCTTGACCAATAGAAATAGCCTTGACAAGGCAAACCTTGACC

Фиг. 7



Дикий тип, среднее 1.79

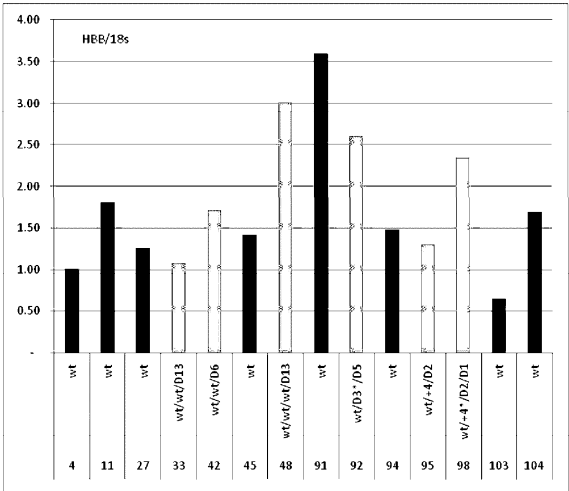
Мутация, среднее 3.01



Дикий тип, среднее 2.85

Мутация, среднее 5.55

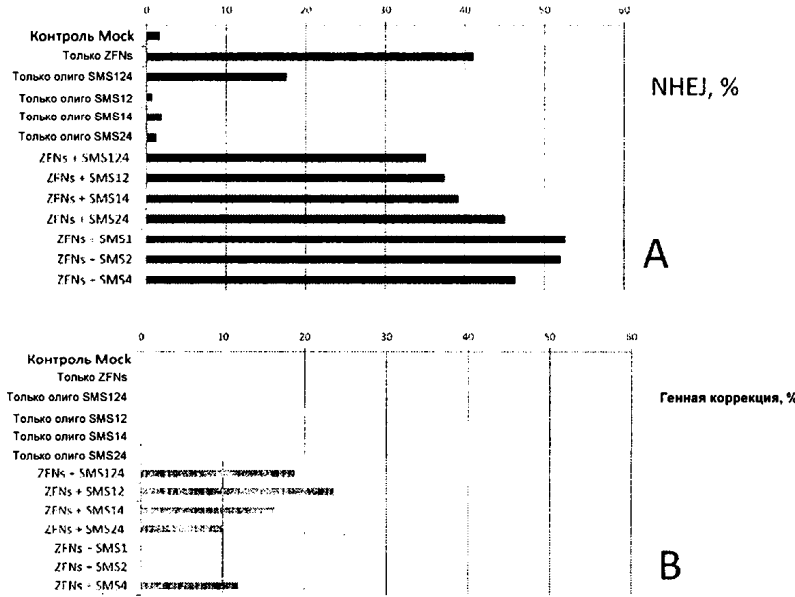
C



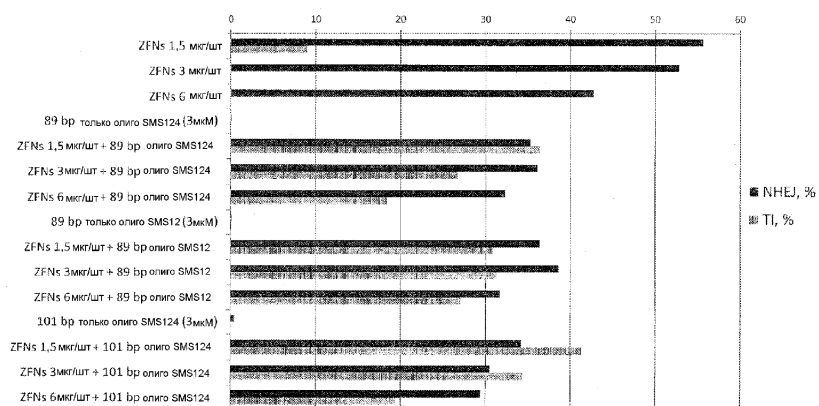
Дикий тип, среднее 1.62
Мутация, среднее 2.01
Фиг. 8



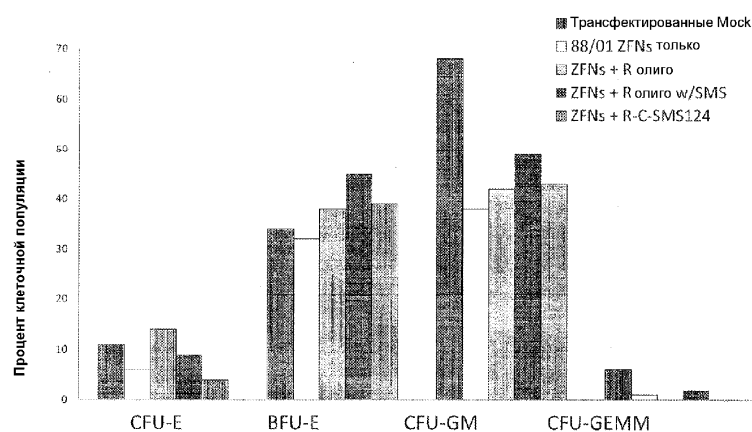
Фиг. 9



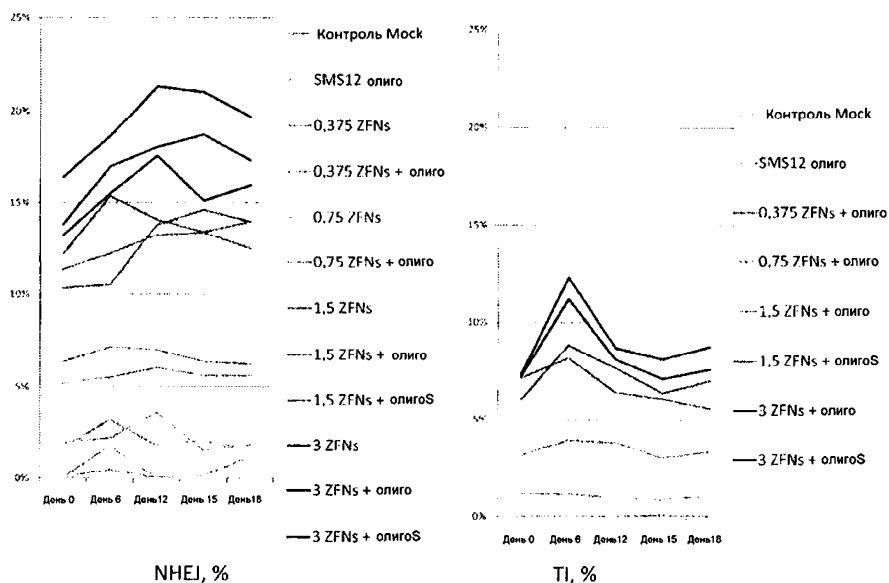
Фиг. 10



Фиг. 11



Фиг. 12



Фиг. 13



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2