

(19)



**Евразийское
патентное
ведомство**

(11) **045868**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

- (45) Дата публикации и выдачи патента
2024.01.11
- (21) Номер заявки
201990568
- (22) Дата подачи заявки
2017.08.24
- (51) Int. Cl. **C07K 14/47** (2006.01)
C12N 9/16 (2006.01)
C12N 15/85 (2006.01)
A61K 35/12 (2015.01)
A61K 38/46 (2006.01)

(54) **РЕГУЛИРОВАНИЕ ЭКСПРЕССИИ ГЕНОВ С ИСПОЛЬЗОВАНИЕМ
СКОНСТРУИРОВАННЫХ НУКЛЕАЗ**

- (31) **62/378,978; 62/443,981; 62/545,778**
- (32) **2016.08.24; 2017.01.09; 2017.08.15**
- (33) **US**
- (43) **2019.07.31**
- (86) **PCT/US2017/048397**
- (87) **WO 2018/039440 2018.03.01**
- (71)(73) Заявитель и патентовладелец:
САНГАМО ТЕРАПЬЮТИКС, ИНК.
(US)
- (72) Изобретатель:
Миллер Джеффри К., Ребар Эдвард
Дж. (US)
- (74) Представитель:
Медведев В.Н. (RU)
- (56) **US-A1-20150132269**
US-A1-20130209426
WO-A2-2015168547
US-A1-20150322131
US-A1-20160090607

-
- (57) Данное изобретение относится к области геномной инженерии, в частности к направленной модификации генома гемопоэтической клетки.
-

B1

045868

045868
B1

Перекрестные ссылки на родственные заявки

Данная заявка испрашивает приоритет по предварительной заявке США № 62/378978, поданной 24 августа 2016 г.; предварительной заявке США № 62/443981, поданной 9 января 2017 г.; и предварительной заявке США № 62/545778, поданной 15 августа 2017 г., раскрытие которых включено в данный документ в качестве ссылки в полном объеме.

Область техники

Данное изобретение относится к области геномной инженерии, в частности к направленной модификации генома гемопоэтической клетки.

Уровень техники

Если учесть, что усилия по сиквенированию генома показали, что геном человека содержит от 20000 до 25000 генов, но менее 2000 регуляторов транскрипции, становится ясно, что многие факторы должны взаимодействовать для контроля экспрессии генов во всех ее различных проявлениях во времени и процессе развития и тканеспецифических проявлениях. Экспрессия генов контролируется очень сложной смесью общих и специфических регуляторов транскрипции, которые взаимодействуют с ДНК-элементами. Эти ДНК-элементы включают как локальные ДНК-элементы, такие как основной промотор и связанные с ним сайты связывания транскрипционных факторов, так и дистальные элементы, такие как энхансеры, сайленсеры, инсуляторы и области контроля локуса (LCR) (см. Matson et al. (2006) *Ann Rev Genome Hum Genet* 7: 29-50).

Энхансерные элементы были сначала идентифицированы в вирусном геноме SV40, а затем обнаружены в локусе тяжелой цепи человеческого иммуноглобулина. В настоящее время известно, что они играют регуляторную роль в экспрессии многих генов, и, по-видимому, энхансеры в основном влияют на временные и пространственные характеристики экспрессии генов. Также было обнаружено, что энхансеры могут функционировать для регуляции экспрессии на больших расстояниях от основного промотора целевого гена и не зависят от какой-либо конкретной ориентации последовательности по отношению к промотору. Энхансеры могут быть расположены в нескольких сотнях килобаз выше или ниже ядра промоторной области, где они могут быть расположены в последовательности интрона или даже за 3'-концом гена.

Были описаны различные способы и композиции для целевого разрезания геномной ДНК. Такие целевые разрезы можно использовать, например, для индукции целевого мутагенеза, индукции целевых делеций последовательностей клеточной ДНК и облегчения направленной рекомбинации в заранее определенном локусе хромосомы. См., например, патенты США №№ 9255250; 9200266; 9045763; 9005973; 9150847; 8956828; 8945868; 8703489; 8586526; 6534261; 6599692; 6503717; 6689558; 7067317; 7262054; 7888121; 7972854; 7914796; 7951925; 8110379; 8409861; патентные публикации США №№ 2003/0232410; 2005/0208489; 2005/0026157; 2005/0064474; 2006/0063231; 2008/0159996; 2010/00218264; 2012/0017290; 2011/02 65198; 2013/0137104; 2013/0122591; 2013/0177983; 2013/0196373; 2015/0056705 и 2015/0335708, раскрытие которых включено в качестве ссылки во всей их полноте.

Эти методы часто включают использование сконструированных систем разрезания для индукции двойного разрыва цепи (DSB) или надреза в последовательности ДНК-мишени таким образом, что восстановление разрывов вызванных ошибкой, такой как негомологичное соединение концов (NHEJ), негомологично направленный концевой захват доноров или репарация с использованием репаративной матрицы (направляемое гомологией восстановление или HDR) может привести к выключению (нокауту) гена или вставке последовательности представляющей интерес (целевая интеграция). См., например, патенты США №№ 9045763; 9200266; 9005973 и 8703489. Эти методы могут также использоваться для введения сайт-специфических изменений в последовательность генома посредством использования донорного олигонуклеотида, включая введение специфических делеций геномных областей или специфических точечных мутаций или локализованных изменений (также известных как коррекция генов). Разрезание может происходить посредством использования специфических нуклеаз, таких как сконструированные нуклеазы цинкового пальца (ZFN), эффекторные нуклеазы, подобные активатору транскрипции (TALEN), или с помощью системы CRISPR/Cas с сконструированной crPHK/tracrPHK ("одиночная направляющая РНК") для направления конкретного разрезания. Кроме того, целевые нуклеазы разрабатываются на основе системы Аргонавт (Argonaute) (например, из *T. thermophilus*, известной как TtAgo, см. Swarts et al. (2014) *Nature* 507 (7491): 258-261), которая также может иметь потенциал для использования в редактировании генома и генной терапии.

Красные кровяные клетки (ККК, RBC), или эритроциты, являются основным клеточным компонентом крови и составляют четверть всех клеток человека. У зрелых эритроцитов отсутствует ядро и многие другие органеллы, и они полны гемоглобина, металлопротеина, который выполняет функцию переноса кислорода из легких в ткани, а также выведения углекислого газа из тканей и обратно в легкие для удаления из организма. Этот белок составляет примерно 97% от сухой массы эритроцитов и увеличивает способность крови переносить кислород примерно в семьдесят раз. Гемоглобин представляет собой гетротетрамер, включающий две альфа (α)-подобные глобиновые цепи и две бета (β)-подобные глобино-

вые цепи и 4 гемовые группы. У взрослых, $\alpha_2\beta_2$ тетрамер называется гемоглобином А (HbA) или гемоглобин взрослых. Как правило, альфа- и бета-глобиновые цепи синтезируются в приблизительном соотношении 1:1, и это соотношение представляется критическим с точки зрения стабилизации гемоглобина и эритроцитов. У развивающегося плода вырабатывается другая форма гемоглобина, фетальный гемоглобин (HbF), который обладает более высокой аффинностью связывания с кислородом, чем гемоглобин А, так что кислород может доставляться в систему ребенка через кровотоки матери. Есть два гена, которые кодируют эмбриональный глобин, которые очень похожи по последовательности и называются HBG1 (также называемый Ggamma) и HBG2 (Agamma), что основано на их порядке расположения в локусе гена бета-глобина.

Как и гемоглобин взрослых, фетальный гемоглобиновый белок содержит две α -глобиновые цепи, но вместо β -глобиновых цепей взрослых он имеет две фетальные гамма(γ)-глобиновые цепи (например, фетальный гемоглобин представляет собой $\alpha_2\gamma_2$). Примерно через 30 недель беременности синтез гамма-глобина у плода начинает снижаться, а выработка бета-глобина увеличивается. Примерно к 10-месячному возрасту гемоглобин новорожденного почти весь представляет собой $\alpha_2\beta_2$, хотя некоторый HbF сохраняется в зрелом возрасте (примерно 1-3% от общего количества гемоглобина). Регуляция перехода от продуцирования гамма- к бета-глобину довольно сложна и, главным образом, включает подавление транскрипции гамма-глобина с одновременным повышением транскрипции бета-глобина.

Генетические дефекты в последовательностях, кодирующих цепи гемоглобина, могут быть ответственны за группу заболеваний, известных как гемоглобинопатии, которые включают серповидноклеточную анемию и альфа- и бета- талассемии. У большинства пациентов с гемоглобинопатиями гены, кодирующие гамма-глобин, остаются, но экспрессия относительно низкая из-за нормальной репрессии генов, происходящей в период родов, как описано выше.

По оценкам 1 из 5000 человек в США страдает серповидноклеточной анемией (SCD), в основном она наблюдается у людей африканского происхождения из региона к югу от Сахары (Roseff (2009) *Immunohematology* 25(2):67). Есть польза для гетерозиготных носителей мутации серповидноклеточной анемии, которая состоит в защите от малярии, так что эта черта, возможно, была отобрана в результате позитивной селекции с течением времени, в итоге, как оценивают, в странах Африки к югу от Сахары мутацию серповидноклеточной анемии имеют до 28% населения (Elguero et al., (2015) *PNAS USA* 112(22):7051). Болезнь серповидноклеточной анемии вызвана мутацией в гене β -глобина вследствие замещения валина на глутаминовую кислоту в аминокислотной позиции № 6 (мутация GAG-GTG на уровне ДНК), причем полученный гемоглобин называется "гемоглобин S" или "HbS". В условиях пониженного содержания кислорода конформационный сдвиг в дезокси-форме HbS обнажает гидрофобную область на белке между спиралями E и F. Гидрофобные остатки валина в позиции 6 бета-цепи в гемоглобине способны связываться с гидрофобной областью, вызывая агрегацию молекул HbS и образование волокнистых преципитатов. Эти агрегаты, в свою очередь, вызывают аномалию или "серповидность" эритроцитов, что приводит к потере гибкости клеток. Серповидные эритроциты больше не могут втиснуться в капиллярные ложа и могут привести к вазо-окклюзионному кризису у пациентов с серповидноклеточной анемией. Кроме того, серповидные эритроциты являются более хрупкими, чем нормальные эритроциты, и имеют тенденцию к гемолизу, что в конечном итоге приводит к анемии у пациента.

Лечение и забота о пациентах с серповидноклеточной анемией продолжается всю жизнь и включает обработку антибиотиками, обезболивание и переливание крови во время острых эпизодов. Одним из подходов является использование гидроксимочевина, которая оказывает влияние частично за счет увеличения выработки гамма-глобина. Однако долгосрочные побочные эффекты продолжительной терапии гидроксимочевиной до сих пор неизвестны, и обработка дает нежелательные побочные эффекты, которые приводят к низкой степени соблюдения режима лечения, и имеет различную эффективность от пациента к пациенту (Brandow and Panepinto (2011) *Am J Hematol* 86(9):804-806). Несмотря на повышение эффективности способов лечения серповидноклеточной анемии, ожидаемая продолжительность жизни пациентов все еще только в середине и конце 50 лет, и связанные с этим заболеванием патологические состояния оказывают сильное влияние на качество жизни пациентов.

Талассемии также являются заболеваниями, связанными с гемоглобином, и обычно включают снижение экспрессии цепей глобина. Это может происходить посредством мутаций в регуляторных областях генов или из-за мутации в последовательности, кодирующей глобин, которая приводит к снижению экспрессии или снижению уровня функционального белка глобина. Альфа-талассемии, вызванные мутациями в локусе альфа-глобина, в основном связаны с людьми из Западной Африки и южно-азиатского происхождения и могут вызывать устойчивость к малярии. Бета-талассемия, вызванная мутациями в локусе бета-глобина, в основном связана с людьми средиземноморского происхождения, как правило, из Греции и прибрежных районов Турции и Италии. При малой талассемии только один из аллелей β -глобина несет мутацию. Эти люди страдают от микроцитарной анемии, и ее обнаружение обычно связано с меньшим, чем обычно, средним корпускулярным объемом (<80 фл). Аллели субъектов с малой талассемией представляют собой β^+/β или β^0/β_3 (где " β^+ " относится к аллелям, которые позволяют образовывать некоторое количество β -цепи, " β " относится к аллелям β -глобина дикого типа и " β^0 " относится

к мутациям β -глобина, связанным с полным отсутствием экспрессии бета-глобина). Субъекты с промежуточной талассемией часто могут вести нормальный образ жизни, но могут нуждаться время от времени в переливании крови, особенно во время болезни или беременности, в зависимости от тяжести их анемии. Аллели этих пациентов могут быть β^+/β^+ или β^0/β^+ . Большая талассемия возникает, когда оба аллеля имеют мутации талассемии. Это сильно микроцитарная и гипохромная анемия. Без лечения он вызывает анемию, спленомегалию и тяжелые деформации костей и приводит к смерти в возрасте до 20 лет. Лечение состоит из периодического переливания крови; спленэктомии для лечения спленомегалии и хелатирования перегрузки железом вызванной трансфузией. Трансплантация костного мозга также используется для лечения людей с тяжелой талассемией, если можно идентифицировать подходящего донора, но эта процедура может иметь значительные риски.

Один из подходов, который был предложен для лечения как SCD, так и бета-талассемии, заключается в повышении экспрессии гамма-глобина с целью функциональной замены aberrантного гемоглобина взрослых на HbF. Как упомянуто выше, считается, что обработка пациентов с SCD гидроксимочевой является успешной отчасти из-за ее влияния на увеличение экспрессии гамма-глобина. Первой группой соединений, обнаруженных в качестве влияющих на экспрессию гамма-глобина, были цитотоксические препараты. Способность вызывать синтез гамма-глобина de novo путем фармакологических манипуляций была впервые продемонстрирована с использованием 5-азациитидина у экспериментальных животных (DeSimone (1982) Proc Natl Acad Sci USA 79(14):4428-31). Последующие исследования подтвердили способность 5-азациитидина повышать HbF у пациентов с β -талассемией и серповидноклеточной анемией (Ley, et al., (1982) N. Engl. J. Medicine, 307:1469-1475, and Ley, et al., (1983) Blood 62: 370-380). Кроме того, в экспериментальных системах было показано, что короткоцепочечные жирные кислоты (например, бутират и производные) повышают уровень HbF (Constantoulakis et al., (1988) Blood 72(6):1961-1967). Кроме того, существует часть человеческой популяции с состоянием, известным как "наследственная персистенция фетального плода" (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 физиологическое состояние, но в отсутствие каких-либо связанных бета-глобин-аномалий, оно не связано с какими-либо значительными клиническими проявлениями, даже если 100% гемоглобина индивида является HbF. Когда у людей, страдающих бета-талассемией, также имеется сопутствующий HPFH, экспрессия HbF может снизить тяжесть заболевания (Potoka and Gladwin (2015) A MJ Physiol Lung Cell Mol Physiol. 308(4): L314-L324). Кроме того, тяжесть естественного течения серповидноклеточной анемии может значительно варьироваться от пациента к пациенту, и эта изменчивость, отчасти, может быть связана с тем фактом, что некоторые люди с более легкой формой заболевания экспрессируют более высокие уровни HbF.

Один из подходов к увеличению экспрессии HbF включает идентификацию генов, чьи продукты играют роль в регуляции экспрессии гамма-глобина. Одним из таких генов является BCL11A, впервые идентифицированный из-за его роли в развитии лимфоцитов. BCL11A кодирует белок цинкового пальца, который, как полагают, участвует в специфической регуляции экспрессии гамма-глобина на стадии развития. BCL11A экспрессируется в эритроидных клетках-предшественниках взрослых, и подавление его экспрессии приводит к увеличению экспрессии гамма-глобина. Кроме того, сплайсинг мРНК BCL11A регулируется в зависимости от стадии развития организма. В эмбриональных клетках в основном экспрессируются более короткие варианты мРНК BCL11A, известные как BCL11A-S и BCL11A-XS, тогда как в клетках взрослых преимущественно экспрессируются более длинные варианты мРНК BCL11A-L и BCL11A-XL. См. Sankaran et al. (2008) Science 322 p. 1839. Белок BCL11A, по-видимому, взаимодействует с локусом бета-глобина, изменяя его конформацию и, следовательно, его экспрессию на разных стадиях развития. Было предложено использование ингибирующей РНК, нацеленной на ген BCL11A (см., например, публикацию патента США № 2011/0182867), но эта технология имеет несколько потенциальных недостатков, которые состоят в том, что полный нокаут не может быть достигнут, доставка таких РНК может быть проблематичной, и РНК должны присутствовать постоянно, требуя многократных обработок в течении всей жизни.

Нацеливание на энхансерные последовательности BCL11A обеспечивает механизм увеличения количества HbF. См., например, публикацию патента США № 2015/0132269 и публикацию PCT № WO 2016/183298. Исследования геномных ассоциаций выявили ряд генетических вариаций в локусе гена BCL11A, которые связаны с повышенным уровнем HbF. Эти вариации представляют собой совокупность небольших (единичных) нуклеотидных полиморфизмов (SNP), обнаруженных в некодирующих областях BCL11A, которые функционируют как специфичная для стадии ограниченная по линии происхождения энхансерная область. Дальнейшие исследования показали, что этот энхансер BCL11A необходим в эритроидных клетках для экспрессии BCL11A, но не требуется для его экспрессии в В-клетках (см. Bauer et al. (2013) Science 343:253-257). Область энхансера была обнаружена в интроне 2 гена BCL11A, и были идентифицированы три области гиперчувствительности к ДНКзе I (часто свидетельствующие о состоянии хроматина, которое связано с регуляторным потенциалом) в интроне 2. Эти три области были определены как "+62", "+58" и "+55" в соответствии с расстоянием в килобазах от стартового сайта транс-

крипции BCL11A.

Эти энхансерные области примерно 350(+55); 550(+58) и 350(+62) нуклеотидов в длину (Bauer 2013, там же).

При разработке нуклеазы для использования в терапевтическом лечении людей важно, чтобы нуклеаза имела максимальные характеристики безопасности для пациентов. В частности, нуклеазы должны иметь очень низкий уровень разрезания вне мишени. Значительное количество двухцепочечных разрезов в местах, отличных от указанной пользователем мишени, может привести к репрессии нецелевых генов, а в редких случаях - к появлению хромосомных транслокаций (см. Hoban and Bauer (2016) *Blood*, 127(21):2525-2535 и Tsang et al. (2017) *Nature Methods*, в печати). Улучшения в специфичности могут быть достигнуты путем устранения неспецифических взаимодействий между сконструированной нуклеазой и геномной ДНК (см. предварительные заявки на патент США №№ 62/378978 и 62/443981).

Таким образом, остается потребность в дополнительных высокоспецифичных способах и композициях для изменения экспрессии гена BCL11A, например, для лечения гемоглобинопатий, таких как серповидноклеточная анемия и бета-талассемия.

Сущность изобретения

Данное изобретение описывает высокоспецифичные композиции и способы для применения в генной терапии и геномной инженерии. В частности, описанные способы и композиции относятся к инактивации (например, путем полной или частичной отмены его экспрессии) гена BCL11A, например, гена, который действует как регулятор одного или более дополнительных генов. В частности, изобретение описывает способы и композиции для воздействия на функцию энхансера в гене BCL11A для уменьшения или подавления его активности в специфических клеточных линиях (например, эритроидных). Кроме того, изобретение предлагает способы и композиции для вмешательства в функции энхансера BCL11A, причем последовательности энхансера не локализованы в кодирующих последовательностях гена BCL11A, и причем предложенные реагенты проявляют высокоспецифичную активность. Результирующее подавление гена BCL11A в этих обстоятельствах приводит к увеличению экспрессии гамма-глобина, и количество явлений разрезания вне мишени снижается.

В некоторых аспектах изобретение включает не встречающийся в природе белок цинкового пальца, содержащий белок цинкового пальца (ZFP), содержащий 4, 5 или 6 пальцев, причем каждый палец содержит область спирали распознавания, которая распознает второстепенный сайт-мишень в ДНК, причем области спирали распознавания содержат последовательности в порядке, указанном в одной строке табл. 1. Внутри каждого цинкового пальца спираль распознавания область из 7 аминокислот пронумерована от -1 до + 6 в остове цинкового пальца (приблизительно из 30 остатков, включая координирующие остатки для цинка). В определенных вариантах осуществления 1, 2, 3 или более компонентных цинковых пальцев белков цинковых пальцев, описанных в данном документе, дополнительно содержат мутации в одном или более остатках вне области спирали распознавания, включая, но не ограничиваясь этим, мутации аминокислот в позиции -5 позиции -14 или мутации в обеих позициях -5 и -14 (нумерация продолжается от нумерации от -1 до + 6, используемой для области распознающей спирали). См., например, мутации Qm4 и Qm14, описанные в предварительных патентных заявках США 62/378978 и 62/443981. Компонентные цинковые пальцы белка цинковых пальцев могут быть связаны любыми линкерами, например, как описано в патенте США № 8772453. В некоторых вариантах осуществления ZFP (белок цинкового пальца) содержит распознающие спирали, показанные в табл. 1 для белков, обозначенных следующим образом: 63014 (который связывается с сайтом-мишенью, показанным в SEQ ID NO: 1) и 65722 (который связывается с сайтом-мишенью, показанным в SEQ ID NO: 2).

В определенных вариантах осуществления белки цинкового пальца, как описано в данном документе, слиты с функциональным доменом (например, доменом активации транскрипции, доменом репрессии транскрипции, доменом разрезания (для образования нуклеазы цинкового пальца) и т.д.). Любой линкер может быть использован для осуществления функциональной связи домена разрезания и белка цинкового пальца, включая, но не ограничиваясь этим, линкеры, описанные в патентах США №№ 9394531 и 9567609. Кроме того, когда используется домен разрезания FokI, могут присутствовать дальнейшие мутации в каталитическом домене, домене димеризации в фосфат-контактных остатках (не присутствующих в димеризационном или каталитическом домене), и комбинации мутаций в любом одном из каталитического домена, домена димеризации в фосфат-контактных остатках, включая, но не ограничиваясь этим, мутации ELD или KKR в домене димеризации, мутации в остатках 525 (K-S) домена FokI и комбинации мутаций ELD или KKR в домене димеризации и мутации в остатках 525 (от K до S) домена FokI, пронумерованные относительно дикого типа. См. патенты США № 7888121; 7914796; 8034598; 8623618 и публикации патента США № 2011/0201055 и предварительные заявки на патент США №№ 62/378978 и 62/443981.

В некоторых вариантах осуществления нуклеазы цинковых пальцев (ZFN) можно использовать в димеризующихся парах для разрезания в одном или обоих целевых сайтах для ZFN пары, например, "левый партнер" в табл. 1 (например, 63014) может образовывать димер с "правым партнером" из табл. 1 (например, 65722) для разрезания энхансерных последовательностей BCL11A. В определенных вариантах осуществления пара ZFN содержит следующие аминокислотные последовательности:

63014:

MDYKDHGDYKDHDIYKDDDDKMAPKKRKRKVGIVPAAMAERPFQCRICMQNFSQSNLRAH
 IRTHTEGKPFACDICGRKFARNFSLTMTKIHTGSQKPFQCRICMQNFSSTGNLTNHIRTHTEG
 KPFACDICGRKFATSGSLTRHTKIHTHPRAPIPKPFQCRICMQNFSQSNLRAHIRTHTEGKPF
 ACDICGRKFAAQCLFHHTKIH- Линкер -
 ELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSR
 KPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPFIGQADEMERYVEENQTRDKHLNPNEWVKVYPS
 VTEFKFLFVSGHFKNYKAQLTRLNHIITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNG
 EINFRS (SEQ ID NO:3); а также
 MDYKDHGDYKDHDIYKDDDDKMAPKKRKRKVGIVPAAMAERPFQCRICMQK FARND
 HRTHTKIHTGKPFQCRICMQNFSQKALIRHIRTHTGKPFACDICGRKFAQKGTGENTKI
 HTGSQKPFQCRICMQNFSRGRDLRHRIRTHTGKPFACDICGRKFARRDNLHSHTKIH-
 Линкер-
 ELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSR
 KPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPFIGQADEMORYVKENQTRNKHINPNEWVKVYPS
 VTEFKFLFVSGHFSGNYKAQLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNG

EINF (SEQ ID NO: 4), в которых последовательность линкера может быть любой последовательностью линкера известной в данной области, например, как описано в патентах США №№ 9394531 и 9567609. В некоторых вариантах осуществления линкер для 63014 содержит или состоит из линкера L7c5 (LRGSISRARPLNPHP (SEQ ID NO: 5)), и линкер, используемый в 65722, содержит или состоит из линкера LO (LRGSQLVKS (SEQ ID NO: 6), см. патент США № 9567609). Последовательность домена разрезания FokI, которая является С-концевой относительно линкера последовательностей, показанных выше, также может содержать альтернативные домены FokI, функционально связанные с белком цинкового пальца. В некоторых вариантах осуществления домен разрезания FokI может включать альтернативные или дополнительные мутации в каталитическом домене, домене димеризации, фосфат-контактных остатках и комбинации мутаций в любом одном из каталитического домена, домена димеризации и фосфат-контактных остатках.

В другом аспекте изобретение включает доставку по меньшей мере одной нуклеазы (например, нуклеазы, которая связывается с последовательностью энхансера BCL11A) в стволовую клетку человека или клетку-предшественник (HSC/PC) с целью инженерии генома. В некоторых вариантах осуществления нуклеаза содержит белок цинкового пальца (ZFP), содержащий 4, 5 или 6 пальцев, каждый палец содержит область спирали распознавания, которая распознает целевой сайт, причем области спирали распознавания содержат последовательности в порядке, показанном в одном ряду табл. 1. В других вариантах осуществления нуклеаза ZFN содержит пару нуклеаз, обозначенных 63014/65722. Нуклеаза(ы), как описано в данном документе, может дополнительно содержать линкер (например, между ДНК-связывающим доменом и доменом разрезания), например линкер, как показано в 9567609, включая, но не ограничиваясь этим, LRGSISRARPLNPHP (SEQ ID NO: 5) или (LRGSQLVKS (SEQ ID NO: 6)).

В некоторых вариантах осуществления нуклеаза доставляется в виде пептида, тогда как в других она доставляется в виде нуклеиновой кислоты, кодирующей по меньшей мере одну нуклеазу. В некоторых вариантах осуществления используется более одной нуклеазы. В некоторых предпочтительных вариантах осуществления нуклеиновая кислота, кодирующая нуклеазу, представляет собой мРНК, а в некоторых случаях мРНК защищена. В некоторых аспектах мРНК может быть химически модифицирована (см., например, Kormann et al. (2011) Nature Biotechnology 29(2):154-157). В других аспектах мРНК может содержать ARCA кэп (см. патенты США №№ 7074596 и 8153773). В других вариантах осуществления мРНК может содержать смесь немодифицированных и модифицированных нуклеотидов (см. публикацию патента США № 2012/0195936). В предпочтительном варианте осуществления нуклеиновая кислота, кодирующая нуклеазу(ы), доставляется в HSC/PC посредством электропорации. В некоторых вариантах осуществления нуклеаза разрезает в сайте или вблизи сайта связывания фактора транскрипции. В некоторых аспектах фактором транскрипции является GATA-1.

В других аспектах изобретение включает клетку или клеточную линию, в которой эндогенная энхансерная последовательность BCL11A генетически модифицирована нуклеазой, как описано в данном документе (например, показанной в табл. 1), например, по сравнению с последовательностью дикого типа клетки. Генетическая модификация энхансера BCL11A приводит к модификации экспрессии гена глобина (бета и гамма). Модифицированные нуклеазой клетки или клеточные линии, как описано в данном документе, отличаются по структуре, функции и сочетаниям как структуры, так и функции от дикого типа. Генетически модифицированные клетки или клеточные линии могут быть гетерозиготными или гомозиготными по модификации. Модификации могут включать вставки (например, вставки трансгена) и комбинации вставок и делеций; такие вставки (инсерции), делеции и комбинации вставок и делеций обычно упоминаются как "инделы" (indel). В некоторых предпочтительных вариантах осуществления изобретения инделы приводят к разрушению сайта связывания фактора транскрипции. В некоторых ва-

риантах осуществления модификация происходит в или около сайта(ов) связывания нуклеазы(нуклеаз), сайта(ов) разрезания и комбинаций сайтов связывания и разрезания, например, в пределах 1-300 (или любого значения между ними) пар оснований выше- или ниже расположенных от сайта(ов) разрезания, более предпочтительно в пределах 1-100 пар оснований (или любого значения между ними) с любой стороны от сайта(ов) связывания, сайта(ов) разрезания и комбинаций сайтов связывания и разрезания как показано в табл. 1, еще более предпочтительно в пределах от 1 до 50 пар оснований (или любого значения между ними) на любой стороне сайта(ов) связывания, сайта(ов) разрезания и комбинаций сайтов связывания и разрезания. В определенных вариантах осуществления генетическая модификация последовательности энхансера BCL11A находится внутри и/или между последовательностями, показанными в табл. 1 (сайты-мишени). Модификация также может включать модификации одного или более нуклеотидов в сайтах разрезания. Модификация также может включать модификации одного или более нуклеотидов в сайтах связывания. Модификация может дополнительно включать модификации одного или более нуклеотидов в сайтах разрезания и в одном или более сайтах связывания. В определенных вариантах осуществления один или более сайтов-мишеней нуклеазы не модифицированы. В других вариантах осуществления, по меньшей мере, один из сайтов-мишеней для нуклеазы(нуклеаз) модифицирован. В некоторых вариантах осуществления модификация находится в или около области "+58" энхансера BCL11A, например, в или около сайта связывания нуклеазы, показанного в любом из SEQ ID NO: 1 и SEQ ID NO: 2. Любая клетка или клеточная линия может быть модифицирована нуклеазами, как описано в данном документе, например, стволовая клетка (гемопоэтическая стволовая клетка, такая как CD34+ гемопоэтическая стволовая клетка) или клетка-предшественник эритроцитов (RBC).

Также описаны клетки или клеточные линии, полученные после модификации нуклеазой, как описано в данном документе, например клетки или клеточные линии, происходящие от модифицированной нуклеазой клетки или клеточной линии, как описано в данном документе. Также предложены частично или полностью дифференцированные клетки, происходящие от модифицированных стволовых клеток, как описано в данном документе (например, эритроциты или клетки-предшественники эритроцитов). Клетки, происходящие от модифицированных нуклеазой(ами) клеток, могут размножаться, дифференцироваться и комбинациями как с размноженных, так и дифференцированных в культуре *in vitro* или могут дифференцироваться в живом субъекте, например, после введения *ex vivo* модифицированной нуклеазой(ами) стволовой клетки. Любые из генетически модифицированных клеток или клеточных линий, раскрытых в данном документе, могут демонстрировать повышенную экспрессию гамма-глобина. Композиции, такие как фармацевтические композиции, содержащие генетически модифицированные клетки, как описано в данном документе, также предлагаются.

В других аспектах изобретение включает доставку донорной нуклеиновой кислоты в клетку-мишень, чтобы обеспечить генетически модифицированную клетку, в которой указанный донор интегрирован в клетку. Донор может быть доставлен до, после или вместе с нуклеиновой кислотой, кодирующей нуклеазу(ы) из табл. 1. Донорная нуклеиновая кислота может содержать экзогенную последовательность (трансен) для интеграции в геном клетки, например, эндогенный локус. В некоторых вариантах осуществления донор может содержать полноразмерный ген или его фрагмент, фланкированные областями гомологии с целевым сайтом разрезания. В некоторых вариантах осуществления донор не имеет гомологичных областей и интегрируется в локус-мишень посредством механизма, независимого от гомологии (т.е. NHEJ). Донор может содержать любую последовательность нуклеиновой кислоты, например нуклеиновую кислоту, которая при использовании в качестве субстрата для гомологически направленной репарации нуклеазо-индуцированного двухцепочечного разрыва приводит к генерации указанной донором делеции на эндогенной локусе хромосомы (например, энхансерная область BCL11A) или, альтернативно (или в дополнение к), новые аллельные формы (например, точечные мутации, которые инактивируют сайт связывания фактора транскрипции) создаваемого эндогенного локуса. В некоторых аспектах донорная нуклеиновая кислота представляет собой олигонуклеотид, причем интеграция приводит к явлению коррекции гена или целевой делеции.

В других аспектах нуклеаза, донор и комбинации как нуклеазы, так и донора доставляются(доставляются) вирусными, невирусными и комбинациями вирусных и невирусных способов переноса генов. В предпочтительных вариантах осуществления донор доставляется в клетку посредством аденоассоциированного вируса (AAV). В некоторых случаях AAV содержит LTR, которые принадлежат гетерологичному серотипу по сравнению с серотипом капсида.

В некоторых аспектах содержащие делеции области внутри гиперчувствительных к ДНКазе I областей энхансера (например, области +58 энхансера BCL11A), получают с использованием одной или большего количества нуклеаз, как показано в табл. 1. Эти делеции могут содержать от около 1 нуклеотида до около 551 нуклеотида. Таким образом, делеции могут содержать 1, 5, 10, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550 нуклеотидов или любое значение между ними. В некоторых вариантах осуществления делеции содержат области связывания для одного или более факторов транскрипции. В некоторых предпочтительных вариантах осуществления делеции содержат сайт связывания GATA-1 или сайт связывания для GATA-1 в сочетании с другими факторами.

В некоторых вариантах осуществления ДНК-связывающие домены из табл. 1 слиты с функцио-

нальным доменом. Некоторые аспекты включают в себя слияние ДНК-связывающих доменов с доменами способными регулировать экспрессию гена. В некоторых вариантах осуществления слитые белки содержат ДНК-связывающий домен из табл. 1, слитый с доменом модуляции экспрессии гена, причем модулятор подавляет экспрессию гена.

В некоторых вариантах осуществления HSC/PC клетки приводят в контакт с нуклеазами, ДНК-связывающими белками и комбинациями нуклеаз и ДНК-связывающих белков по данному изобретению (т.е. ZFP, как показано в табл. 1). В некоторых вариантах осуществления нуклеазы, ДНК-связывающие белки и комбинации нуклеаз и ДНК-связывающих белков доставляются в виде нуклеиновых кислот, а в других вариантах осуществления они доставляются в виде белков. В некоторых вариантах осуществления нуклеиновые кислоты представляют собой мРНК, кодирующие ДНК-связывающие белки и комбинации нуклеаз и ДНК-связывающих белков, и в дополнительных вариантах осуществления мРНК могут быть защищены. В некоторых вариантах осуществления мРНК может быть химически модифицирована и может включать ARCA кэп, смесь немодифицированных и модифицированных нуклеотидов и комбинации ARCA кэпа и смеси немодифицированных и модифицированных нуклеотидов. Также предусмотрены клетки или клеточные линии, происходящие от этих клеток, включая частично или полностью дифференцированные клетки.

В некоторых аспектах HSC/PC приводятся в контакт с нуклеазами, ДНК-связывающими белками и комбинациями нуклеаз и ДНК-связывающих белков по данному изобретению *ex vivo*, после афереза HSC/PC от субъекта или очистки из полученного костного мозга. В некоторых вариантах осуществления нуклеазы, описанные в данном документе, вызывают модификации в областях энхансера BCL11A, например, приводя к генетически модифицированной клетке, которая структурно, функционально, и комбинированно структурно и функционально, отлична от клеток дикого типа, других модифицированных (например, модифицированных нуклеазой(ами)) клеток и комбинации дикого типа и других модифицированных клеток. В других вариантах осуществления HSC/PC, содержащие модификации энхансерной области BCL11A, вводят обратно субъекту. В некоторых случаях HSC/PC, содержащие модификации энхансерной области BCL11A, размножают перед введением. В других аспектах генетически модифицированные HSC/PC вводят субъекту в составе трансплантата костного мозга, причем HSC/PC приживаются, дифференцируются и созревают *in vivo*. В некоторых вариантах осуществления HSC/PC выделяют у субъекта после мобилизации, индуцированной G-CSF, мобилизации, индуцированной плериксафором (plerixafor), и комбинаций мобилизации, индуцированной G-CSF, и плериксафором, а в других клетки выделяют из костного мозга человека или пуповин человека. В некоторых аспектах субъекта подвергают мягкой миелоабляционной (myeloablative) процедуре до введения трансплантата, содержащего модифицированный HSC/PC, в то время как в других аспектах субъекта подвергают интенсивным условиям миелоабляционного режима. В некоторых вариантах осуществления способы и композиции по данному изобретению используются для лечения или профилактики гемоглобинопатии. В некоторых аспектах гемоглобинопатия представляет собой талассемию. В некоторых аспектах гемоглобинопатия представляет собой бета-талассемию, в то время как в других аспектах гемоглобинопатия представляет собой серповидноклеточную анемию.

В некоторых вариантах осуществления HSC/PC дополнительно приводят в контакт с донорной молекулой. В некоторых вариантах осуществления донорная молекула доставляется вирусным вектором. Донорная молекула может содержать одну или более последовательностей, кодирующих функциональный полипептид (например, кДНК или его фрагмент), с промотором или без него. Дополнительные последовательности (кодирующие или некодирующие последовательности) могут быть включены, когда донорная молекула используется для инактивации, в том числе, но не ограничиваясь ими, последовательности, кодирующие пептид 2A, сайт SA, IRES и др.

В одном аспекте способы и композиции по данному изобретению включают способы приведения в контакт с HSC/PC *in vivo*. Нуклеазы, ДНК-связывающие белки или комбинация нуклеаз и ДНК-связывающих белков доставляются в HSC/PC *in situ* способами, известными в данной области техники. В некоторых вариантах осуществления нуклеазы и/или ДНК-связывающие белки по данному изобретению содержат вирусную частицу, которую вводят нуждающемуся субъекту, тогда как в других нуклеазы, ДНК-связывающие белки или комбинация нуклеаз и ДНК-связывающих белков содержат наночастицу (например, липосому). В некоторых вариантах осуществления вирусные частицы, наночастицы или комбинация вирусных частиц и наночастиц доставляются в орган (например, костный мозг), в котором находятся HSC/PC.

В другом аспекте описанном в данном документе описаны способы интеграции донорной нуклеиновой кислоты в геном клетки посредством механизмов независимых от гомологии. Указанные способы включают создание двуцепочечного разрыва (DSB) в геноме клетки и разрезание донорной молекулы с использованием нуклеазы, как описано в данном документе, так что донорная нуклеиновая кислота интегрируется в сайт DSB. В некоторых вариантах осуществления донорная нуклеиновая кислота интегрируется посредством способов независимых от гомологии (например, NHEJ). Как отмечено выше, после разрезания *in vivo* донорные последовательности могут быть целенаправленно интегрированы в геном клетки в месте расположения DSB. Донорная последовательность может включать один или более оди-

наковых сайтов-мишеней для одной или большего количества нуклеаз, используемых для создания DSB. Таким образом, донорная последовательность может быть разрезана одной или более одинаковыми нуклеазами, используемыми для разрезания эндогенного гена, в который желательна интеграция. В некоторых вариантах осуществления донорная последовательность включает сайты-мишени нуклеазы, отличных от нуклеаз, используемых для индукции DSB. DSB в геноме клетки-мишени могут быть созданы по любому механизму. В некоторых вариантах осуществления DSB создается одной или более нуклеазами цинковых пальцев (ZFN), слитыми белками, содержащими домен связывания цинковых пальцев, который сконструирован для связывания последовательности в пределах области представляющей интерес, и доменом разрезания или полудоменом разрезания.

В одном аспекте донор может кодировать представляющий интерес регуляторный белок (например, TF ZFP, TF TALE или TF CRISPR/Cas), который связывается с, модулирует экспрессию или как связывается, так и модулирует экспрессию представляющего интерес гена. В одном варианте осуществления указанные регуляторные белки связываются с последовательностью ДНК и предотвращают связывание других регуляторных факторов. В другом варианте осуществления связывание регуляторного белка может модулировать (т.е. индуцировать или репрессировать) экспрессию ДНК-мишени.

В некоторых вариантах осуществления трансгенная HSC/PC клетка, трансгенное животное или комбинация трансгенных HSC/PC клетки и животного включает трансген, который кодирует ген человека. В некоторых случаях трансгенное животное содержит нокаут в эндогенном локусе и замену эндогенного гена его человеческим аналогом, что позволяет разработать систему *in vivo*, где человеческий белок можно изучать изолированно. Такие трансгенные модели могут быть использованы для целей скрининга, чтобы идентифицировать малые молекулы или большие биомолекулы или другие объекты, которые могут взаимодействовать с или модифицировать человеческий белок представляющий интерес. В некоторых аспектах трансген интегрируется в выбранный локус (например, безопасную геномную "гавань" (safe-harbor)) в стволовую клетку (например, эмбриональную стволовую клетку, индуцированную плюрипотентную стволовую клетку, гемопоэтическую стволовую клетку и т.д.) или эмбрион животного, полученный путем любого из методов, описанных в данном документе, и затем эмбрион имплантируется таким образом, что рождается живое животное. Затем животное доводится до половой зрелости и ему дают возможность произвести потомство, причем, по меньшей мере, некоторые из потомков содержат отредактированную последовательность эндогенного гена или интегрированный трансген.

В другом аспекте в данном документе представлен способ изменения экспрессии гена (например, BCL11A, гена глобина и комбинаций BCL11A и гена глобина) в клетке, включающий: введение в клетку одной или большего количества нуклеаз описанных в данном документе (показаны в табл. 1), в условиях, когда один или более белков экспрессируются и экспрессия гена изменяется. В некоторых вариантах осуществления экспрессия гена глобина (например, гамма-глобина или бета-глобина) изменяется (например, увеличивается). Любой из способов, описанных в данном документе, может дополнительно включать в себя интеграцию донорной последовательности (например, трансгена или его фрагмента под контролем экзогенного или эндогенного промотора) в геном клетки, например, интегрирование донора в или вблизи сайта разрезания нуклеазой в гене BCL11A. Донорную последовательность вводят в клетку с использованием вирусного вектора, как олигонуклеотид, на плазмиде, и комбинации одного или более методов, выбранных из использования вирусного вектора, как олигонуклеотид, или на плазмиде. Клетка, в которой экспрессия генов изменена может быть, например, предшественником красных кровяных клеток (эритроцитов, RBC), гемопоэтических стволовых клеток (например, CD34+ клеток), а также комбинации клеток -предшественников эритроцитов и гемопоэтических стволовых клеток.

В других вариантах осуществления в данном документе представлен способ получения генетически модифицированной клетки, содержащей геномную модификацию в эндогенной последовательности энхансера BCL11A (модификацию нуклеотидной последовательности энхансера BCL11A), причем способ включает стадии: а) приведения в контакт клетки с полинуклеотидом (например, ДНК или мРНК), кодирующим нуклеазу цинкового пальца, содержащую 4, 5 или 6 доменов цинкового пальца, в которой каждый из доменов цинкового пальца содержит область спирали распознавания в порядке, показанном в одной строке табл. 1; б) подвергание клетки условиям, способствующим экспрессии белка цинкового пальца из полинуклеотида; и с) модифицирование эндогенной энхансерной последовательности BCL11A экспрессированным белком цинкового пальца, достаточное для получения генетически модифицированной клетки. В некоторых вариантах осуществления клетки стимулируют по меньшей мере одним цитокином (например, перед стадией (а)). Полинуклеотид может приводиться в контакт с клеткой с использованием любого подходящего способа, включая, но не ограничиваясь этим, посредством трансфекции, с использованием невирусного вектора, с использованием вирусного вектора, с использованием химических средств или путем воздействия электрического поля (например, электропорацией).

Также предлагаются клетки, содержащие одну или комбинацию геномных модификаций, описанных в данном документе, включая клетки, происходящие от клеток, полученных способами, описанными в данном документе.

Также предложен способ лечения пациента, нуждающегося в повышении экспрессии гена глобина, включающий введение пациенту фармацевтического препарата, причем фармацевтический препарат со-

держит генетически модифицированные клетки, белки, полинуклеотиды и комбинации одного или большего количества выбранного из генетически модифицированных клеток, белков и полинуклеотидов, как описано в данном документе, в количестве, достаточном для увеличения экспрессии гена глобина у пациента. В определенных вариантах осуществления известно, что пациент имеет диагноз, подозревается в наличии или подвержен риску развития талассемии или серповидноклеточной анемии.

Также предлагается набор, содержащий нуклеиновые кислоты, белки, генетически модифицированные клетки и комбинации одного или большего количества выбранного из нуклеиновых кислот, белков и генетически модифицированных клеток по данному изобретению. Набор может включать нуклеиновые кислоты кодирующие нуклеазы (например, молекулы РНК или гены, кодирующие ZFN, TALEN или систему CRISPR/Cas, содержащиеся в подходящем векторе экспрессии), алиquotы нуклеазных белков, донорные молекулы, подходящие модификаторы самообновления стволовых клеток ("стволовости"), клетки, буферы, инструкции (например, для осуществления способов по данному изобретению) и тому подобное, включая различные комбинации этих компонентов набора. Изобретение включает, но не ограничивается ими, генетически модифицированную клетку (например, стволовую клетку, такую как гемопоэтическая (CD34+) стволовая клетка или клетка-предшественник эритроцитов), содержащую по меньшей мере одну геномную модификацию, сделанную нуклеазой (например, как показано в одном ряду табл. 1), причем геномная модификация находится в эндогенной энхансерной последовательности BCL11A, и причем, кроме того, геномная модификация выбрана из группы, состоящей из вставок, делеций и их комбинаций, и содержит модификацию в, вблизи или между любой из SEQ ID NO: 1 и SEQ ID NO: 2. В некоторых вариантах осуществления клетка представляет собой генетически модифицированную дифференцированную клетку, происходящую из стволовой клетки, как описано в данном документе (например, RBC, происходящую из гемопоэтической стволовой клетки или клетки-предшественника RBC).

Нуклеаза может содержать по меньшей мере одну нуклеазу цинкового пальца (ZFN) (например, как показано в табл. 1), по меньшей мере один TALEN и комбинации по меньшей мере одного ZFN и по меньшей мере одного TALEN. Указанная нуклеаза(ы) может быть введена в клетку в виде белка, как полинуклеотид, кодирующий нуклеазу(ы), или в виде комбинации белковой формы и полинуклеотида кодирующего нуклеазу(ы). В определенных вариантах осуществления геномная модификация включает инсерцию (вставку), которая включает в себя интеграцию донорного полинуклеотида кодирующего трансген. Также предложены фармацевтические композиции, содержащие одну или более генетически модифицированных клеток, как описано в данном документе.

Также предложен ДНК-связывающий белок, содержащий белок цинкового пальца, содержащий 4, 5 или 6 доменов цинкового пальца, содержащих область спирали распознавания, причем белок цинкового пальца содержит области спирали распознавания в порядке, показанном в одном ряду табл. 1. Также предложен белок TALE, содержащий множество повторов, которые связываются с последовательностью, содержащей часть (например, по меньшей мере, 4, 5, 6 или более) пар оснований сайтов-мишеней, показанных в табл. 1. Слитый белок, содержащий белок цинкового пальца или белок TALE, как описано в данном документе, и домен разрезания дикого типа или сконструированный или полудомен разрезания, также предлагаются, как и полинуклеотиды, кодирующие белки (ZFP, TALE, ZFN, TALEN), описанные в данном документе. Также предусмотрены клетки (например, изолированные стволовые клетки, такие как гематопоэтические (CD34+) стволовые клетки), содержащие один или более полинуклеотидов, белков и комбинаций полинуклеотидов и белков, как описано в данном документе. Также предложены наборы, содержащие один или более белков, полинуклеотидов, клеток или их комбинаций, как описано в данном документе.

Также описан способ изменения экспрессии гена глобина в клетке (например, клетке-предшественнике эритроцитов, гемопоэтической стволовой клетке и комбинаций клетки-предшественника эритроцитов и гемопоэтической стволовой клетки), включающий: введение в клетку одного или более полинуклеотидов, кодирующих одну или более нуклеаз, как описано в данном документе, в условиях, при которых один или более белков экспрессируются, а экспрессия гена глобина (например, гамма-глобина, бета-глобина и комбинаций гамма-и бета-глобина) изменяется (например, увеличивается). В некоторых вариантах осуществления способы дополнительно включают в себя интеграцию донорной последовательности в геном клетки, например, с использованием вирусного вектора, в качестве олигонуклеотида или на плазмиде. Указанная донорная последовательность может содержать трансген под контролем эндогенного или экзогенного промотора.

Также предложен способ получения генетически модифицированной клетки, содержащей геномную модификацию в эндогенной энхансерной последовательности BCL11A (например, сайте-мишени, как показано в табл. 1), причем способ включает стадии: (a) приведение в контакт клетки с полинуклеотидом кодирующим слитый белок, содержащий нуклеазу цинкового пальца, содержащую 4, 5 или 6 доменов цинкового пальца, в которой каждый из доменов цинкового пальца содержит область спирали распознавания в порядке, показанном в одном ряду табл. 1; (b) подвергание клетки условиям, способствующим экспрессии слитого белка из полинуклеотида; и (c) модифицирование эндогенной энхансерной последовательности BCL11A экспрессированным слитым белком, достаточное для получения генетиче-

ски модифицированной клетки. В определенных вариантах осуществления способ дополнительно включает стимуляцию клеток по меньшей мере одним цитокином. Полинуклеотид(ы) может быть доставлен внутрь клетки, например, с использованием невирусной системы доставки, вирусной системы доставки, доставляющего носителя и комбинаций, выбранных из невирусной системы доставки, вирусной системы доставки и доставляющего носителя и может включать воздействие на клетки электрическим полем или использование разрушения клеточной мембраны в качестве механизма доставки (так называемая "технология сжатия", см., например, Sharei et al. (2015) PLOS ONE doi: 10.1371/journal.pone.0118803).

Способы лечения пациента, нуждающегося в повышении экспрессии гена глобина (например, известно, что пациент имеет диагноз, подозреваемый в наличии или подвержен риску развития гемоглобинопатии, такой как талассемия (например, β -талассемия) или серповидноклеточная анемия также предусмотрены, при этом способ включает введение пациенту фармацевтической композиции, как описано в данном документе (например, белков, полинуклеотидов, клеток или комбинации, выбранной из белков, полинуклеотидов и клеток) в количестве, достаточном для увеличения экспрессии гена глобина у пациента.

Эти и другие аспекты будут очевидны для опытного специалиста в свете раскрытия в целом.

Краткое описание графических материалов

Фиг. 1А и 1В представляют собой схемы, изображающие обзор анализа сайта интеграции олигонуклеотидного дуплекса. На фиг. 1 изображена первая стадия анализа, в которой клетки обрабатывают ZFN в присутствии дополнительной олигонуклеотидной дуплексной ДНК, которая захватывается во фракцию результирующих явлений разрезания. На фиг. 1В продемонстрирована следующая стадия, на которой клетки культивируют в течение семи дней, выделяют геномную ДНК и сегменты генома, фланкирующие сайты интеграции донора, амплифицируют с помощью ПЦР, опосредованной адаптером, с использованием праймера для интегрированного олигонуклеотидного дуплекса. Затем ампликоны секвенируют для выявления потенциальных сайтов разрезания.

На фиг. 2 изображено расположение 455 потенциальных локусов вне мишени, идентифицированных с помощью анализа сайта интеграции олигонуклеотидного дуплекса с использованием исходной 51857/51949 или родительской пары ZFN. Лучшие 63 локуса выделены серым цветом и были проанализированы в последующих анализах. Каждый локус может быть идентифицирован по хромосоме и номеру основания, указывающему наиболее вероятное местоположение разрезания, а также указывает количество интегрантов, обнаруженных в этом локусе.

На фиг. 3А-3С представлены иллюстративные фосфат-контактные остатки, внутри каркаса цинкового пальца. На фиг. 3А изображено взаимодействие цинкового пальца с молекулой ДНК и изображено расположение боковой цепи аргинина дикого типа и как она взаимодействует с фосфатным остовом молекулы ДНК. На фиг. 3В (SEQ ID NO: 7-17) представлены типичные последовательности ZFN, показывающие, где этот аргинин расположен в первичной последовательности каждого цинкового пальца (вертикальный ряд остатков 'R', обозначенный жирной стрелкой), а также выделены те аргинины, которые были замещены глутамином в остове ZFP, чтобы устранить соответствующий фосфатный контакт (индивидуально помеченные 'R' остатки). Последовательности также показывают области спирали распознавания (где остатки от -1 до +3, +5 и +6 обведены прямоугольником и затенены), а также часть линкера между С-терминальным доменом цинкового пальца и доменом разрезания (домен разрезания не изображен). На фиг. 3С дополнительно изображено пространственное расположение другой боковой цепи лизина, потенциально контактирующей с остовом, присутствующей в домене разрезания FokI, которая во время оптимизации специфичности может быть заменена серином.

Фиг. 4 представляет собой график, сравнивающий уровни модификации посредством обработки CD34 клеток либо исходной парой 51857/51949, либо оптимизированной парой 63014/65722 в различных тестируемых локусах. Внизу или по оси x показаны локусы, идентифицированные как потенциальные мишени разрезания, с процентным значением инделов для каждой области (сайта), показанной на вертикальной оси или оси Y. Обратите внимание, что ось Y показана в логарифмическом масштабе. Темно-серые столбцы показывают локусы, расщепленные парой 51857/51949, и количество обнаруженного разрезания, в то время как светлосерые столбцы - представляют локусы, расщепленные оптимизированной парой, где почти все измеренное разрезание находится в целевой последовательности BCL11a.

Фиг. 5А и 5В представляют собой графики, изображающие относительные соотношения мРНК глобина произведенной в hCD34+ клетках после обработки специфичными для BCL11A ZFN и эритроидной дифференцировки. CD34+ клетки, полученные от двух здоровых доноров-людей (PB-MR-003 или PB-MR-004), обрабатывали или не обрабатывали парой ZFN, а затем анализировали экспрессию α , β и γ глобина. Наилучшим методом определения количества мРНК γ -глобина, обнаруживаемого после обработки ZFN, является выражение изменения экспрессии либо в виде отношения γ -глобина к β -глобину (фиг. 5А), либо мРНК γ -глобина к α -глобину (фиг. 5В).

Фиг. 6 представляет собой график, изображающий относительные количества белка γ -глобина, продуцируемого в обработанных CD34+ клетках. Как и выше, использовали две серии CD34+ клеток, полученных от здоровых доноров-людей (PB-MR-003 и PB-MR-004). В этом эксперименте приблизительно 3-

4-кратное повышение процентного содержания белка глобина плода до уровней около 15-20% наблюдалось в эритроидном потомстве HSPC при 63014/65722-опосредованном разрушении энхансера BCL11A в обеих партиях доноров.

Фиг. 7 представляет собой график, показывающий относительный человеческий химеризм у мышей, которым трансплантированы партии обработанных 63014/65722 доноров ("+" ZFN"), как описано выше. Человеческий химеризм измеряли путем детектирования клеток, несущих маркер hCD45 на их поверхности с помощью FACS. Показано процентное содержание hCD4 5+ клеток человека в периферической крови, собранной через 8 или 12 недель после трансплантации. Данные показали хорошие уровни приживления в этом исследовании со сравнимым человеческим химеризмом после приживления нетрансфицированного контроля ("(-)") и HSPC, трансфицированных ZFN ("+" ZFN"). Незаполненные круги и треугольники представляют отдельных животных.

На фиг. 8 изображен процент химерности, обнаруженный в костном мозге привитых мышей, у которых клетки человека были идентифицированы по присутствию hCD45 на их клеточной поверхности. Образцы анализировали через 12 недель после приживления.

Фиг. 9A-9D представляют собой графики, изображающие восстановление различных гематопозитических линий клеток, протестированных FACS-анализом клеток костного мозга у привитых мышей, полученных на 12 неделе, антителами, распознающими маркеры клеточной поверхности, специфичные для определенных популяций. Данные показали сравнимое присутствие всех проанализированных гематопозитических линий человека в костном мозге на 12-й неделе после инъекции в потомстве CD34+ клеток ("14/22") обработанных BCL11A-специфической мРНК ZFN и нетрансфицированных клетках ("(-)"). Показаны данные по лимфоидным, миелоидным, эритроидным клеткам и HSPC (фиг. 9A-9D, соответственно) для клеток, полученных от обоих доноров ("003" и "004"). Данные показали сравнимое представление всех проанализированных гематопозитических линий человека в костном мозге на 12-й неделе после инъекции в потомстве клеток CD34+, обработанных мРНК ZFN BCL11A, и потомстве нетрансфицированных клеток.

Фиг. 10 представляет собой график, показывающий уровень модификации гена на мишени BCL11A в ДНК, выделенной из периферической крови модифицированных мышей, проанализированной методом глубокого секвенирования. Данные приведены для введенных клеток (через 2 дня после трансфекции ZFN ("+")), а затем для клеток крови через 8 или 12 недель после приживления и продемонстрировали хорошее сохранение модификации гена. Нетрансфицированные клетки обозначены "(-)" в строке "Обработка".

Фиг. 11 представляет собой график, показывающий количество модификации гена на мишени BCL11A для образцов клеток костного мозга после приживления ZFN-обработанных клеток ("+"). Необработанные клетки обозначены "(-)" в строке "Обработка". Сопоставимая модификация наблюдалась как в BCL11A-зависимых линиях (B-клетки, экспрессирующие маркер CD19; примитивные предшественники, экспрессирующие CD45 и высокие уровни CD38), так и в BCL11A-независимых (миелоидных) линиях. Хотя уровни генной модификации при вводе были выше в образце донора PB-MR-003, чем в образце донора PB-MR-004, клетки полученные из PB-MR-004 повторяемо демонстрировали более высокие уровни модификации, т.е. лучшее сохранение модификации у мышей, чем полученные из PB-MR-003.

Фиг. 12 представляет собой график, показывающий количество модификации генов в эритроидных клетках, полученных из клеток костного мозга на 12 неделе после дифференцировки *in vitro* в течение 14 дней. Данные были получены от мышей, изначально привитых двумя различными донорами, описанными выше, и продемонстрировали, что модификация BCL11A, опосредованная обработкой ZFN ("+" ZFN"), заметно не изменяется во время дифференцировки эритроида. Клетки, которые не были обработаны ZFN, обозначены как "(-) ZFN".

Фиг. 13A и 13B представляют собой графики, изображающие относительное количество мРНК, кодирующей γ -глобин, где концентрация мРНК γ -глобина изображается либо как отношение мРНК γ -глобина/ β -глобина (фиг. 9A), либо как отношение мРНК γ -глобина/ α -глобина (фиг. 9B). Как в нетрансфицированных ("(-) ZFN"), так и в образцах, обработанных ZFN ("+" ZFN"), отношения мРНК γ -глобина к β глобину или γ -глобина к α -глобину сильно различались между потомками эритроидных клеток отдельных мышей из одной и той же группы. Однако, несмотря на эту изменчивость и изменчивость, вызванную использованием двух разных доноров-людей, средние значения для обработанных 63014/65722 образцов демонстрируют увеличение уровней мРНК γ -глобина в $\sim 1,5$ -2 раза по сравнению с их соответствующими нетрансфицированными аналогами.

Фиг. 14 представляет собой график, изображающий разницу в количестве белка γ -глобина (выраженную в виде отношения γ -глобин/ α -глобин или γ -глобин/общий β -подобный белок) в клетках, полученных из костного мозга от модифицированных мышей, причем клетки костного мозга подвергались протоколу дифференцировки *in vitro*. Уровни белка измеряли через 16 дней после начала дифференцировки. Были определены отношения гамма (γ) глобина (сумма пиков A γ и G γ) к альфа (α) глобину, а также гамма глобин (сумма пиков A γ и G γ)/относительно соотношений бета-подобного глобина (сумма пиков A β , G β , бета и дельта-глобинов). Согласуясь со слабой

эритроидной дифференцировкой образцов, полученных из PB-MR-003, уровни гамма-глобина в не-трансфицированных клетках, полученных от этого донора, были очень высокими (~30%), и поэтому обработка ZFN ("+" ZFN") приводила к увеличению уровня гамма-глобина только в 1,2 раза по сравнению с необработанными клетками ("(-) ZFN"). PB-MR-004 демонстрировал более типичные нетрансфицированные уровни (~ 9%) и показывал ~ 2-кратное увеличение уровней белка гамма-глобина после 12-недельной циркуляции через тело мыши.

Подробное описание сущности изобретения

Здесь раскрыты композиции и способы для геномной инженерии для модуляции экспрессии BCL11A, гамма-глобина и комбинаций BLC11A и гамма-глобина, а также для лечения, профилактики или лечения и профилактики гемоглобинопатии. В частности, посредством воздействия с использованием нуклеаз, содержащих ZFP, имеющих области спирали узнавания, как показано в одном ряду табл. 1, разрушение энхансера BCL11A эффективно достигается в HSC/PC и приводит к изменению относительной экспрессии гамма-глобина во время последующего эритропоэза. Эта модуляция экспрессии BCL11A и гамма-глобина особенно полезна для лечения гемоглобинопатий (например, бета-талассемии, серповидноклеточной анемии), при которых наблюдается недостаточная экспрессия бета-глобина или экспрессия мутированной формы бета-глобина. При использовании способов и композиций по данному изобретению осложнения и последствия, связанные с заболеванием, вызываемые aberrантным количеством бета-глобина, можно преодолеть путем изменения экспрессии гамма-глобина в клетках-предшественниках эритроцитов.

Общее описание

Практикование способов, а также приготовление и использование композиций, раскрытых в данном документе, используют, если не указано иное, обычные методы молекулярной биологии, биохимии, структуры и анализа хроматина, вычислительной химии, клеточных культур, рекомбинантных ДНК и смежных областей, которые находятся в пределах навыков в данной области техники. Эти методы полностью объяснены в литературе, см. например, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Second edition, Cold Spring Harbor Laboratory Press, 1989 и третье издание, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 и периодические переиздания; серию METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, третье издание, 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 дезоксирибонуклеотидному или рибонуклеотидному полимеру в линейной или кольцевой конформации и в одноцепочечной или двухцепочечной форме. Для целей данного раскрытия эти термины не должны рассматриваться как ограничивающие по отношению к длине полимера. Указанные термины могут включать известные аналоги природных нуклеотидов, а также нуклеотиды, которые модифицированы по фрагменту основания, сахарному, фосфатному фрагментам (например, фосфоротиоатным остовам) и комбинациям, выбранным из фрагмента основания, сахарного и фосфатного фрагментов. Как правило, аналог конкретного нуклеотида имеет такую же специфичность спаривания с другим основанием; т.е. аналог А будет парой основания Т.

Термины "полипептид", "пептид" и "белок" используются взаимозаменяемо для обозначения полимера из аминокислотных остатков. Указанный термин также применяется к аминокислотным полимерам, в которых одна или более аминокислот являются химическими аналогами или модифицированными производными соответствующих встречающихся в природе аминокислот.

"Связывание" относится к сиквенс-специфичной нековалентной взаимосвязи между макромолекулами (например, между белком и нуклеиновой кислотой). Не все компоненты связывающего взаимодействия должны быть сиквенс-специфичными, (например, контакты с фосфатными остатками в остовах ДНК), если взаимодействие в целом является сиквенс-специфичным. Такие взаимодействия обычно характеризуются константой диссоциации (K_d), равной 10^{-6} M^{-1} или ниже. "Сродство" относится к силе связывания: повышенная аффинность связывания коррелирует с более низким K_d .

"Связывающий белок" - это белок, который способен связываться с другой молекулой. Связывающий белок может связываться, например, с молекулой ДНК (ДНК-связывающий белок), молекулой РНК (РНК-связывающий белок), белковой молекулой (белок-связывающий белок) или может связываться с комбинацией молекул, выбранных из молекулы ДНК, молекулы РНК или белка. В случае белок-связывающего белка, он может связываться с самим собой (с образованием гомодимеров, гомотримеров и т.д.), он может связываться с одной или более молекулами другого белка или белков или он может связываться как с самим собой, так и с одной или более молекул другого белка или белков. Связывающий белок может иметь более одного типа связывающей активности. Например, белки цинкового пальца обладают ДНК-связывающей, РНК-связывающей и белок-связывающей активностью.

"ДНК-связывающий белок цинкового пальца" (или связывающий домен) представляет собой белок или домен в более крупном белке, который связывает ДНК сиквенс-специфичным образом через один или более цинковых пальцев, которые являются областями аминокислотной последовательности в пределах связывающего домена, структура которого стабилизируется за счет координации иона цинка. Термин ДНК-связывающий белок цинкового пальца часто сокращается как белок цинкового пальца или ZFP.

"ДНК-связывающий домен TALE" или "TALE" представляет собой полипептид, содержащий один или более доменов/единиц повтора TALE. Домены повторов участвуют в связывании TALE с узнаваемой им последовательностью ДНК-мишени. Одна "единица повтора" (также называемая "повтор") обычно имеет длину 33-35 аминокислот и проявляет, по меньшей мере, некоторую гомологию последовательности с другими последовательностями повторов TALE в пределах встречающегося в природе белка TALE.

Связывающие домены цинкового пальца и TALE могут быть "сконструированы" для связывания с заранее определенной нуклеотидной последовательностью, например, путем конструирования (изменения одной или большего количества количества аминокислот) области спирали распознавания встречающегося в природе цинкового пальца или белка TALE. Следовательно, сконструированные ДНК-связывающие белки (цинковые пальцы или TALE) - это белки, которые не встречаются в природе. Неограничивающими примерами способов конструирования ДНК-связывающих белков являются дизайн и отбор. Разработанный ДНК-связывающий белок - это белок, не встречающийся в природе, дизайн/состав которого главным образом вытекает из рациональных критериев. Рациональные критерии для проектирования включают применение правил замещения и компьютеризированных алгоритмов для обработки информации в базе данных, хранящей информацию о существующих конструкциях ZFP и/или TALE и данные по связыванию. См., например, патенты США № 6140081; 6453242; 6534261 и 8585526; см. также публикации PCT № WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 и WO 03/016496.

"Отобранный" белок цинкового пальца или TALE - это белок, не встречающийся в природе, производство которого происходит в следствие эмпирического процесса, такого как фаговый дисплей, ловушка взаимодействия или гибридный отбор. См., например, патенты США № 5789538; 5925523; 6007988; 6013453; 6200759; 8586526; публикации PCT № WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970, WO 01/88197, WO 02/099084.

"TtAgo" - прокариотический белок Аргонавт (Argonaute), который, как полагают, участвует в выключении генов. TtAgo происходит из бактерии *Thermus thermophilus*. См., например, Swarts et al. там же, G. Sheng et al. (2013) Proc. Natl. Acad. Sci. USA 111, 652). А "система TtAgo" представляет собой все необходимые компоненты, в том числе, например, направляющие ДНК для разрезания с помощью фермента TtAgo.

"Рекомбинация" относится к процессу обмена генетической информацией между двумя полинуклеотидами, включая, но не ограничиваясь этим, захват донора негомологичным присоединением концов (NHEJ - non-homologous end joining) и гомологичную рекомбинацию. В целях данного раскрытия термин "гомологичная рекомбинация (HR)" относится к специализированной форме такого обмена, который имеет место, например, во время репарации двухцепочечных разрывов в клетках с помощью механизмов репарации, направляемых гомолгией. Этот процесс требует гомологии нуклеотидных последовательностей, использует "донорную" молекулу для матричного репарации "целевой" молекулы (т.е. той, которая подверглась двухцепочечному разрыву) и по-разному известен как "некрссоверная конверсия гена" (non-crossover gene conversion) или "короткоцепочечная конверсия гена" (short tract gene conversion), потому что это приводит к передаче генетической информации от донора к мишени. Не желая быть связанными какой-либо конкретной теорией, такой перенос может включать коррекцию несоответствия гетеродуплексной ДНК, которая образуется между поврежденной мишенью и донором, "синтез-зависимый отжиг цепи", при котором донор используется для повторного синтеза генетической информации, которая станет частью мишени, связанные процессы или комбинацию вышеуказанного. Такая специализированная HR часто приводит к изменению последовательности молекулы-мишени таким образом, что часть или вся последовательность донорного полинуклеотида включается в полинуклеотид-мишень.

В способах по настоящему изобретению одна или более нацеленных нуклеаз, как описано в данном документе, создают двухцепочечный разрыв (DSB) в последовательности-мишени (например, клеточном хроматине) в заранее определенном сайте. DSB может приводить к инделам с помощью механизмов гомологичной репарации, или негомологичной репарации. Делении могут включать любое количество пар оснований. Аналогично, вставки могут включать любое количество пар оснований, включая, например, интеграцию "донорного" полинуклеотида, необязательно имеющего гомологию с нуклеотидной последовательностью в области разрыва. Донорная последовательность может быть физически интегрирована или, альтернативно, донорный полинуклеотид используется в качестве матрицы для репарации разрыва посредством гомологичной рекомбинации, что приводит к введению всей или части нуклеотидной последовательности, как у донора, в клеточный хроматин. Таким образом, первая последовательность в клеточном хроматине может быть изменена и, в определенных вариантах осуществления, может быть преобразована в последовательность, присутствующую в донорном полинуклеотиде. Таким образом, использование терминов "заменить" или "замена" может пониматься как замена одной нуклеотидной

последовательности другой (т.е. замена последовательности в информационном смысле) и необязательно требует физической или химической замены одного полинуклеотида другим.

В любом из способов, описанных в данном документе, дополнительные пары белков цинкового пальца или TALEN могут использоваться для дополнительного двухцепочечного разрезания дополнительных сайтов-мишеней в клетке.

Любой из способов, описанных в данном документе, можно использовать для введения донора любого размера или для частичной или полной инактивации одной или большего количества последовательностей-мишеней в клетке путем направленной интеграции донорной последовательности, которая нарушает экспрессию гена(ов) представляющего интерес. Также предложены клеточные линии с частично или полностью инактивированными генами.

В любом из способов, описанных в данном документе, экзогенная нуклеотидная последовательность ("донорная последовательность" или "трансген") может содержать последовательности, которые гомологичны, но не идентичны геномным последовательностям в области представляющей интерес, тем самым стимулируя гомологичную рекомбинацию для вставки неидентичной последовательности в область представляющую интерес. Таким образом, в определенных вариантах осуществления части донорной последовательности, которые гомологичны последовательностям в области представляющей интерес, демонстрируют от около 80 до 99% (или любое целое число между ними) идентичности последовательности с заменяемой геномной последовательностью. В других вариантах гомология между донорной и геномной последовательностями является выше 99%, например, если только 1 нуклеотид отличается между донорной и геномной последовательностями на 100 смежных пар оснований. В некоторых случаях негомологичная часть донорной последовательности может содержать последовательности, отсутствующие в области представляющей интерес, так что новые последовательности вводятся в область представляющую интерес. В этих случаях негомологичная последовательность обычно фланкируется последовательностями из 50-1000 пар оснований (или любым интегральным значением между ними) или любым числом пар оснований, превышающим 1000, которые гомологичны или идентичны последовательностям в области представляющей интерес. В других вариантах осуществления донорная последовательность является негомологичной первой последовательности и встраивается в геном с помощью механизмов негомологичной рекомбинации.

"Разрезание" относится к разрушению ковалентного остова молекулы ДНК. Разрезание может быть инициировано различными способами, в том числе, но не ограничиваясь ими, ферментативным или химическим гидролизом фосфодиэфирной связи. Возможны как одноцепочечное разрезание, так и двухцепочечное разрезание, и двухцепочечное разрезание может происходить в результате двух различных одноцепочечных явлений разрезания. Разрезание ДНК может привести к образованию либо тупых концов, либо липких концов. В определенных вариантах осуществления слитые полипептиды используются для целевого разрезания двухцепочечной ДНК.

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

"Сконструированный полудомен разрезания" представляет собой полудомен разрезания, который был модифицирован таким образом, чтобы образовывать облигатные гетеродимеры с другим полудоменом разрезания (например, другим сконструированным полудоменом разрезания). См. также патенты США №№ 7888121; 7914796; 8034598; 8623618 и патентную публикацию США № 2011/0201055, полностью включенные в данный документ посредством ссылки.

Термин "последовательность" относится к нуклеотидной последовательности любой длины, которая может представлять собой ДНК или РНК; может быть линейной, циркулярной или разветвленной и может быть как одноцепочечной, так и двухцепочечной. Термин "донорная последовательность" относится к нуклеотидной последовательности, которая вставлена в геном. Донорная последовательность может иметь любую длину, например, от 2 до 10000000 нуклеотидов в длину (или любое целочисленное значение между ними или выше), предпочтительно от примерно 100 до 100000 нуклеотидов в длину (или любое целое число между ними), более предпочтительно между примерно 2000 и 20000 нуклеотидов в длину (или любое значение между ними) и даже более предпочтительно между примерно 5 и 15 т. п.н. (или любое значение между ними).

"Хроматин" - это нуклеопротеиновая структура, составляющая клеточный геном. Клеточный хроматин содержит нуклеиновую кислоту, главным образом ДНК, и белок, включая гистоны и негистоновые хромосомные белки. Большая часть эукариотического клеточного хроматина существует в форме нуклеосом, где ядро нуклеосомы содержит приблизительно 150 пар оснований ДНК, связанных с октамером, содержащим два из каждого из гистонов H2A, H2B, H3 и H4; и линкерная ДНК (переменной длины в зависимости от организма) проходит между ядрами нуклеосом. Молекула гистона H1 обычно связана с линкерной ДНК. Для целей данного раскрытия термин "хроматин" предназначен для охвата всех типов клеточных нуклеопротеинов, как прокариотических, так и эукариотических. Клеточный хроматин вклю-

чае как хромосомный, так и эписомальный хроматин.

"Хромосома" представляет собой комплекс хроматина, содержащий весь или часть генома клетки. Геном клетки часто характеризуется ее кариотипом, который представляет собой совокупность всех хромосом, составляющих геном клетки. Геном клетки может содержать одну или более хромосом.

"Эписомой" является реплицирующаяся нуклеиновая кислота, комплекс нуклеопротеинов или другая структура, включающая нуклеиновую кислоту, которая не является частью хромосомного кариотипа клетки. Примеры эписом включают плазмиды и некоторые вирусные геномы.

"Доступная область" представляет собой сайт в клеточном хроматине, в котором целевой сайт, присутствующий в нуклеиновой кислоте, может быть связан экзогенной молекулой, которая распознает указанный целевой сайт. Не желая быть связанными какой-либо конкретной теорией, полагают, что доступной областью является та, которая не упакована в нуклеосомную структуру. Отличительную структуру доступной области часто можно обнаружить по ее чувствительности к химическим и ферментативным зондам, например, к нуклеазам.

"Сайт-мишень" или "последовательность-мишень" представляет собой последовательность нуклеиновой кислоты, которая определяет часть нуклеиновой кислоты, с которой связывается молекула, при условии наличия достаточных условий для связывания.

"Экзогенная" молекула представляет собой молекулу, которая обычно не присутствует в клетке, но может быть введена в клетку одним или более генетическими, биохимическими или другими методами. "Присутствие в клетке в норме" определяется с учетом конкретной стадии развития и условий окружающей среды клетки. Так, например, молекула, которая присутствует только во время эмбрионального развития мышц, является экзогенной молекулой по отношению к мышечной клетке взрослого человека. Аналогично, молекула, индуцированная тепловым шоком, является экзогенной молекулой по отношению к клетке, не подвергшейся тепловому шоку. Экзогенная молекула может содержать, например, функционирующую версию неправильно функционирующей эндогенной молекулы или неправильно функционирующую версию нормально функционирующей эндогенной молекулы.

Экзогенной молекулой может быть, среди прочего, малая молекула, такая как генерируемая комбинаторным химическим процессом, или макромолекула, такая как белок, нуклеиновая кислота, углевод, липид, гликопротеин, липопротеин, полисахарид, любое модифицированное производное вышеуказанных молекул или любой комплекс, включающий одну или более вышеуказанных молекул. Нуклеиновые кислоты включают ДНК и РНК, могут быть одноцепочечными или двухцепочечными; могут быть линейными, разветвленными или циркулярными; и могут быть любой длины. Нуклеиновые кислоты включают те, которые способны образовывать дуплексы, а также триплексообразующие нуклеиновые кислоты. См., например, патенты США №№ 5176996 и 5422251. Белки включают, но не ограничиваются ими, ДНК-связывающие белки, факторы транскрипции, факторы ремоделирования хроматина, белки связывающие метилированную ДНК, полимеразы, метилазы, деметилазы, ацетилазы, деацетилазы, киназы, фосфатазы, интегразы, рекомбиназы, лигазы, топоизомеразы, гиразы и геликазы.

Экзогенная молекула может представлять собой молекулу того же типа, что и эндогенная молекула, например, экзогенный белок или нуклеиновую кислоту. Например, экзогенная нуклеиновая кислота может содержать заражающий вирусный геном, плазмиду или эписому, введенную в клетку, или хромосому, которая в норме не присутствует в клетке. Способы введения экзогенных молекул в клетки известны специалистам в данной области и включают, но не ограничиваются ими, перенос опосредованный липидами (т.е. липосомы, включая нейтральные и катионные липиды), электропорацию, прямую инъекцию, слияние клеток, бомбардировку частицами, копреципитацию с фосфатом кальция, DEAE-декстран-опосредованный перенос и перенос опосредованный вирусным вектором. Экзогенная молекула также может быть тем же самым типом молекулы, что и эндогенная молекула, но быть получена из вида отличного от вида, к которому принадлежит клетка. Например, последовательность нуклеиновой кислоты человека может быть введена в клеточную линию, первоначально полученную от мыши или хомяка.

Напротив, "эндогенная" молекула - это молекула, которая обычно присутствует в конкретной клетке на определенной стадии развития в определенных условиях окружающей среды.

Например, эндогенная нуклеиновая кислота может содержать хромосому, геном митохондрии, хлоропласта или другой органеллы или природную эписомальную нуклеиновую кислоту. Дополнительные эндогенные молекулы могут включать белки, например, факторы транскрипции и ферменты.

Используемый в данном документе термин "продукт экзогенной нуклеиновой кислоты" включает как полинуклеотидные, так и полипептидные продукты, например, продукты транскрипции (полинуклеотиды, такие как РНК) и продукты трансляции (полипептиды).

"Слитая" молекула представляет собой молекулу, в которой две или более молекул субъединиц связаны, предпочтительно ковалентно. Молекулы субъединиц могут представлять собой молекулы одного и того же химического типа или могут представлять собой молекулы разных химических типов. Примеры первого типа слитой молекулы включают, но не ограничиваются ими, слитые белки (например, слияние между ДНК-связывающим доменом ZFP или TALE и одним или более активационными доменами) и слитые нуклеиновые кислоты (например, нуклеиновые кислоты, кодирующие белок слияния, описанный выше). Примеры слитой молекулы второго типа включают, но не ограничиваются ими, слияние между

нуклеиновой кислотой, образующей триплекс, и полипептидом, и слияние связывающего малой бороздки и нуклеиновой кислоты.

Экспрессия слитого белка в клетке может быть результатом доставки слитого белка в клетку или доставки полинуклеотида, кодирующего слитый белок, в клетку, причем полинуклеотид транскрибируется, а транскрипт транслируется для генерации слитого белка. Трансплайсинг, разрезание полипептида и лигирование полипептида также могут участвовать в экспрессии белка в клетке. Способы доставки полинуклеотидов и полипептидов в клетки представлены в других частях данного описания.

"Ген" для целей данного раскрытия включает область ДНК, кодирующую продукт гена (см. ниже), а также все области ДНК, которые регулируют продукцию продукта гена, независимо от того, являются ли такие регуляторные последовательности близлежащими к кодирующим последовательностям, транскрибируемым последовательностям и комбинации кодирующих и транскрибируемых последовательностей. Соответственно, ген включает, но не ограничивается ими, промоторные последовательности, терминаторы, трансляционные регуляторные последовательности, такие как сайты связывания рибосомы и внутренние сайты входа в рибосому, энхансеры, сайленсеры, инсуляторы, граничные элементы, точки начала репликации, сайты присоединения к матриксу и области контроля локуса.

"Экспрессия гена" относится к преобразованию информации, содержащейся в гене, в продукт гена. Продукт гена может быть прямым продуктом транскрипции гена (например, мРНК, тРНК, рРНК, анти-смысловой РНК, микроРНК, рибозимом, структурной РНК или любым другим типом РНК) или белком, полученным путем трансляции мРНК. Продукты генов также включают РНК, которые модифицированы такими процессами, как экзпирование, полиаденилирование, метилирование и редактирование, и белки, модифицированные, например, метилированием, ацетилизацией, фосфорилированием, юбиквитинированием, ДЦФ-рибозилированием, миристилированием и гликозилированием.

"Модуляция" экспрессии гена относится к изменению активности гена. Модуляция экспрессии может включать, но не ограничивается этим, активацию гена и репрессию гена. Редактирование генома (например, разрезание, изменение, инактивация, случайная мутация) может использоваться для модуляции экспрессии. Инактивация гена относится к любому снижению экспрессии гена по сравнению с клеткой, которая не включает ZFP, TALE или систему CRISPR/Cas, как описано в данном документе. Таким образом, инактивация генов может быть частичной или полной.

"Защищенная" мРНК представляет собой такую, в которой мРНК каким-то образом изменена для повышения стабильности или трансляции мРНК. Примеры защит включают использование замены до 25% остатков цитидина и уридина 2-тиоуридином (s2U) и 5-метилцитидином (m5C). Результирующая мРНК проявляет меньшую иммуногенность и большую стабильность по сравнению с немодифицированным аналогом (см. Karikó et al. ((2012), *Molecular Therapy*, Vol. 16, No. 11, pages 1833-1844). Другие изменения включают добавление так называемого ARCA эпа, который увеличивает трансляционность мРНК, продуцируемой *in vitro* (см. патент США № 7074596).

"Область представляющая интерес" представляет собой любую область клеточного хроматина, такую как, например, ген или некодирующую последовательность внутри или рядом с геном, в которой желательно связать экзогенную молекулу. Связывание может быть осуществлено с целью целевого разрезания ДНК, целевой рекомбинации и комбинаций целевого разрезания ДНК и целевой рекомбинации. Область представляющая интерес может присутствовать, например, в хромосоме, эписоме, геноме органеллы (например, митохондриальном, хлоропластном) или в заражающем вирусном геноме. Область представляющая интерес может находиться внутри кодирующей области гена, внутри транскрибируемых некодирующих областей, таких как, например, лидерные последовательности, концевые последовательности или интроны, или внутри нетранскрибированных областей, либо вышерасположенных, либо нижерасположенных по отношению к кодирующей области. Область, представляющая интерес, может быть такой же маленькой, как одна нуклеотидная пара, или длиной до 2000 нуклеотидных пар, или любым целым значением нуклеотидных пар.

"Эукариотические" клетки включают, но не ограничиваются ими, клетки грибов (таких как дрожжи), клетки растений, клетки животных, клетки млекопитающих и клетки человека (например, Т-клетки).

Термины "функциональная связь" и "функционально связаны" (или "оперативно связаны") используются взаимозаменяемо применительно к сопоставлению двух или более компонентов (таких как элементы последовательности), в которых компоненты расположены так, что оба компонента функционируют нормально и допускают возможность того, что по меньшей мере один из компонентов может опосредовать функцию, которая воздействует по меньшей мере на один из других компонентов. В качестве иллюстрации транскрипционная регуляторная последовательность, такая как промотор, функционально связана с кодирующей последовательностью, если транскрипционная регуляторная последовательность контролирует уровень транскрипции кодирующей последовательности в ответ на наличие или отсутствие одного или более транскрипционных регуляторных факторов. Транскрипционная регуляторная последовательность обычно функционально связана в *cis*-положении с кодирующей последовательностью, но не обязательно должна быть непосредственно смежной с ней. Например, энхансер представляет собой транскрипционную регуляторную последовательность, 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 ту же функцию в связи с другим компонентом, как и если бы они не были связаны таким образом. Например, в отношении слитого полипептида, в котором ДНК-связывающий домен ZFP, TALE или Cas слит с доменом активации, ДНК-связывающий домен ZFP, TALE или Cas и домен активации находятся в функциональной связи, если в слитом полипептиде ДНК-связывающий домен ZFP, TALE или Cas способен связывать свой сайт-мишень, свой сайт связывания и комбинаций сайта-мишени и сайта связывания, и в то же время домен активации способен усиливать экспрессию генов. Когда слитый полипептид, в котором ДНК-связывающий домен ZFP, TALE или Cas слит с доменом разрезания, ДНК-связывающий домен ZFP, TALE или Cas и домен разрезания находятся в функциональной связи, если в слитом полипептиде ДНК-связывающий домен ZFP, TALE или Cas способен связывать свой сайт-мишень, свой сайт связывания и комбинации сайта-мишени и сайта связывания, и в то же время домен разрезания способен расщеплять ДНК в непосредственной близости от сайта-мишени.

"Функциональный фрагмент" белка, полипептида или нуклеиновой кислоты представляет собой блок, полипептид или нуклеиновую кислоту, чья последовательность не идентична полноразмерному белку, полипептиду или нуклеиновой кислоте, но сохраняет ту же функцию, что и полноразмерный белок, полипептид или нуклеиновая кислота. Функциональный фрагмент может содержать больше, меньше или такое же количество остатков, что и соответствующая нативная молекула, может содержать одну или более аминокислотных или нуклеотидных замен и может представлять собой комбинации, содержащие больше, меньше или такое же количество остатков, как и соответствующая нативная молекула и содержащая одну или более аминокислотных или нуклеотидных замен. Методы определения функции нуклеиновой кислоты (например, кодирующей функции, способности гибридизоваться с другой нуклеиновой кислотой) хорошо известны в данной области техники. Точно так же способы определения функции белка хорошо известны. Например, ДНК-связывающая функция полипептида может быть определена, например, с помощью анализов связывания с фильтром, электрофоретической подвижности или иммунопреципитации.

Разрезание ДНК может быть проанализировано с помощью гель-электрофореза. See Ausubel et al., выше. Способность белка взаимодействовать с другим белком может быть определена, например, с помощью коиммунопреципитации, двугибридных анализов или комплементации, как генетической, так и биохимической.

"Вектор" способен переносить последовательности генов в клетки-мишени. Как правило, "векторная конструкция", "вектор экспрессии" и "вектор переноса гена" означают любую конструкцию нуклеиновой кислоты, способную направлять экспрессию представляющего интерес гена и которая может переносить последовательности гена в клетки-мишени. Таким образом, этот термин включает носители клонирования и экспрессии, а также интегрирующие векторы.

Термины "субъект" и "пациент" используются взаимозаменяемо и относятся к млекопитающим, таким как пациенты-люди и приматы, не относящиеся к человеку, а также к экспериментальным животным, таким как свиньи, коровы, кролики, собаки, кошки, крысы, мыши и другие животные. Соответственно, термин "субъект" или "пациент", используемый в данном документе, означает любого пациента или субъекта-млекопитающего, которому могут быть введены стволовые клетки по данному изобретению.

Субъекты данного изобретения включают тех, кто подвергся воздействию одного или более химических токсинов, включая, например, нервный токсин.

"Стволовость" относится к относительной способности любой клетки действовать подобным стволовым клеткам образом, т.е. к степени тотипотенции, плюри-, мульти- или олигопотенции и усиленной или бесконечной способностью к самообновлению, которые какая-либо конкретная стволовая клетка может иметь.

Нуклеазы

Здесь описаны композиции, в частности нуклеазы, которые полезны для разрезания *in vivo* донорной молекулы, несущей трансген, и нуклеазы для разрезания генома клетки таким образом, что трансген целенаправленно интегрируется в геном. В определенных вариантах осуществления одна или более нуклеаз встречаются в природе. В других вариантах осуществления одна или более нуклеаз не встречаются в природе, т.е. модифицированы в ДНК-связывающем домене, домене разрезания и комбинации ДНК-связывающего домена и домена разрезания. Например, ДНК-связывающий домен встречающейся в природе нуклеазы может быть изменен для связывания с выбранным сайтом-мишенью (например, мегануклеаза, которая была сконструирована для связывания с сайтом, отличным от ее природного сайта связывания). В других вариантах осуществления нуклеаза содержит гетерологичные ДНК-связывающие и разрезающие домены (например, нуклеазы цинкового пальца; ДНК-связывающие белки TAL-эффлекторного домена; ДНК-связывающие домены мегануклеазы с гетерологичными доменами разрезания).

А. ДНК-связывающие домены

В некоторых вариантах осуществления ДНК-связывающий домен одной или большего количества нуклеаз, используемых для разрезания *in vivo*, целевого разрезания генома клетки и комбинаций разрезания *in vivo* и целевого разрезания генома клетки, содержит белок цинкового пальца. Один белок цинко-

вого пальца состоит из нескольких/многих доменов цинкового пальца (например, 3, 4, 5, 6 или более доменов цинкового пальца). Каждый домен цинкового пальца имеет длину около 30 аминокислот и содержит бета-виток (содержащий два координационных остатка цинка) и альфа-спираль (содержащую два инвариантных координационных остатка цинка), которые удерживаются в определенной конформации, позволяющей белку связывать целевую последовательность. Можно использовать канонические (C2H2) домены цинкового пальца, имеющие два координирующих цинк цистеиновых остатка (Cys) в бета-витке и два координирующих цинк остатка гистидина (His) в альфа-спирали или неканонический (CH3) (см., например, патент США № 9234187). Спираль распознавания из 7 аминокислот содержится между координирующими цинк остатками бета-витка и координирующими цинк остатками альфа-спирали. Область спирали распознавания пронумерована от -1 до +6 в пределах домена цинкового пальца, а аминокислоты вне этой области распознавания (и исключая координирующие цинк остатки) называются остатками остова.

Предпочтительно белок цинкового пальца является не встречающимся в природе в отношении того, что спираль распознавания сконструирована так, чтобы связываться с выбранным сайтом-мишенью. 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, которые все включены в данный документ посредством ссылки во всей их полноте.

Сконструированный домен связывания цинковых пальцев может иметь новую специфичность связывания по сравнению с природным белком цинковых пальцев. Методы конструирования включают, но не ограничиваются ими, рациональный дизайн и различные типы отбора. Рациональный дизайн включает в себя, например, использование баз данных, содержащих триплетные (или квадруплетные) нуклеотидные последовательности и отдельные аминокислотные последовательности цинковых пальцев, в которых каждая триплетная или квадруплетная нуклеотидная последовательность ассоциирована с одним или более аминокислотными последовательностями цинковых пальцев, которые связывают последовательность конкретного триплета или квадруплета. См., например, совместные патенты США №№ 6453242 и 6534261, включенные в данный документ в качестве ссылки во всей их полноте.

Кроме того, в определенных вариантах осуществления ДНК-связывающие домены ZFP дополнительно содержат одну или более модификаций остова одного или более компонентных доменов цинкового пальца. Специфичность ZFP к последовательности ДНК-мишени зависит от сиквенс-специфических контактов между доменами цинкового пальца и специфическими основаниями ДНК, в частности, между областью спирали распознавания и сайтом-мишенью (обычно каждая спираль распознавания связывается с дочерним сайтом-мишенью из 3-х нуклеотидов). Кроме того, домены цинкового пальца также содержат аминокислотные остатки, которые принимают участие в неспецифических взаимодействиях с фосфатами остова ДНК. Elrod-Erickson et al. ((1996) *Structure* 4:1171), продемонстрировали путем кокристаллизации белка цинкового пальца и узнаваемой им мишени ДНК, что существуют специфические аминокислоты, способные взаимодействовать с фосфатами в остове ДНК посредством образования водородных связей. Белки цинкового пальца, которые используют хорошо известный Zif268 остов обычно имеют аргинин в качестве конечного аминокислотного остатка в их второй цепи бета - листа, который также является второй карбокси терминальной позицией по отношению ко второму инвариантному цистеину. Эта позиция может упоминаться как (-5) в каждом домене цинкового пальца, поскольку это 5-й остаток, предшествующий началу α -спирали (и позиция -5 относительно спирали распознавания, пронумерованной от -1 до +6). Аргинин в этой позиции может взаимодействовать с фосфатом на остове ДНК посредством образования заряженной водородной связи с его группой гуанидина в боковой цепи. Белки цинкового пальца в остове Zif268 также часто содержат лизин в позиции, которая сдвинута на 4 остатка в аминотерминальном направлении по отношению к первому инвариантному цистеину. Эту позицию можно обозначить как (-14) внутри каждого пальца, так как это 14-й остаток, предшествующий началу α -спирали для цинковых пальцев с двумя остатками между координирующими цинк остатками цистеина (и положением -14 относительно области спирали распознавания пронумерованной от -1 до +6). Лизин может взаимодействовать с фосфатом в остове ДНК через образование опосредованной водой заряженной водородной связи с его боковой аминогруппой. Поскольку фосфатные группы обнаруживаются по всей длине остова ДНК, этот тип взаимодействия между цинковым пальцем и молекулой ДНК обычно считается не сиквенс-специфичным (J. Miller, Massachusetts Institute of Technology Ph. D. Thesis, 2002).

Недавние исследования выдвинули гипотезу, что неспецифические фосфат-контактные боковые цепи в некоторых нуклеазах также могут объяснять некоторую неспецифичность этих нуклеаз (Kleistiver et al. (2016) *Nature* 529(7587):490-5; Guilinger et al. (2014) *Nat Meth.* 429-435). Исследователи предположили, что эти нуклеазы могут обладать "избыточной энергией связывания ДНК", что означает, что указанные нуклеазы могут обладать большей аффинностью к своей ДНК-мишени, чем это требуется для существенного связывания и разрезания сайта-мишени. Таким образом, были предприняты попытки

уменьшить катионные заряды в ДНК-связывающем домене TALE (Guilinger, там же) или в ДНК-связывающем домене Cas9 (Kleinstiver, там же), чтобы снизить энергию связывания ДНК этих нуклеаз, что привело к увеличению специфичности разрезания *in vitro*. Однако дополнительные исследования (Sternberg et al. (2015) Nature 527 (7576): 110-113) также указывают на роль в правильном сворачивании и активации нуклеазного домена Cas9 для некоторых катионных аминокислот, которые были мутированы в исследовании Kleinstiver ДНК-связывающего домена Cas9. Таким образом, точная роль этих аминокислот в активности Cas9 не известна.

Способы и композиции по данному изобретению, таким образом, включают мутации аминокислот в пределах ДНК-связывающего домена ZFP ("остова ZFP"), которые могут неспецифично взаимодействовать с фосфатами в остове ДНК, но они не содержат изменений в спиральных распознавания ДНК. Таким образом, изобретение включает мутации катионных аминокислотных остатков в остове ZFP, которые не требуются для специфичности к нуклеотидной мишени. В некоторых вариантах осуществления эти мутации в остове ZFP включают мутацию катионного аминокислотного остатка на нейтральный или анионный аминокислотный остаток. В некоторых вариантах осуществления эти мутации в остове ZFP включают мутацию полярного аминокислотного остатка на нейтральный или неполярный аминокислотный остаток. В предпочтительных вариантах осуществления мутации делаются в позиции (-5), позиции (-9), позиции (-14) и комбинации мутаций, выбранных из мутаций в позиции (-5), позиции (-9) и позиции (-14) относительно спирали связывания ДНК. В некоторых вариантах осуществления цинковый палец может содержать одну или более мутаций в (-5), (-9), (-14) и комбинации мутаций, выбранных из мутаций в (-5), (-9) и (-14). В других вариантах осуществления один или более цинковых пальцев в белке цинковых пальцев с несколькими пальцами, могут содержать мутации в (-5), (-9), (-14) и комбинации, выбранные из (-5), (-9) и (-14). В некоторых вариантах осуществления аминокислоты в (-5), (-9), (-14) и комбинации, выбранные из (-5), (-9) и (-14) (например, аргинин (R) или лизин (K)) мутированы в аланин (A), лейцин (L), Сер (S), Asp (D), Glu (E), Тир (Y) и/или глутамин (Q).

В любом из этих слитых полипептидов, описанных в данном документе, партнеры ZFP могут дополнительно содержать мутации в ДНК-связывающем домене цинкового пальца в позициях (-5), (-9), (-14) и комбинации мутаций, выбранных из мутаций в (-5), (-9) и (-14). В некоторых вариантах осуществления Arg (R) в позиции -5 меняется на Тир (Y), Asp (D), Glu (E), Leu (L), Gln (Q) или Ala (A). В других вариантах осуществления Arg (R) в позиции (-9) меняется на Ser (S), Asp (D) или Glu (E). В других вариантах осуществления Arg (R) в позиции (-14) заменяется Ser (S) или Gln (Q). В других вариантах осуществления слитые полипептиды могут содержать мутации в ДНК-связывающем домене цинкового пальца, где аминокислоты в позициях (-5), (-9), (-14), и комбинации мутаций, выбранных из мутаций в (-5), (-9) и (-14) заменены на любую из перечисленных выше аминокислот в любой комбинации.

Иллюстративные способы 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, усиление специфичности связывания для доменов связывания цинковых пальцев было описано, например, в совместной заявке WO 02/011221.

Подбор целевых сайтов; ZFP и способов проектирования и конструирования и слитых белков (и полинуклеотидов, кодирующих их) известны специалистам в данной области и подробно описаны в патентах США №№ 6140081; 5789538; 6453242; 6534261; 5925523; 6007988; 6013453; 6200759; публикациях РСТ № 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.

Почти любой линкер (спейсер) может использоваться между одним или более компонентами ДНК-связывающего домена (например, цинкового пальца), между одним или более ДНК-связывающими доменами, между ДНК-связывающим доменом и функциональным доменом (например, нуклеазным) и между одним или более ДНК-связывающими доменами и между ДНК-связывающим доменом и функциональным доменом. Неограничивающие примеры подходящих линкерных последовательностей включают патенты США №№ 8772453; 7888121; 6479626; 6903185 и 7153949; и патентные публикации США №№ 2009/0305419; 2015/0064789 и 2015/0132269. Таким образом, белки, описанные в данном документе, могут включать любую комбинацию подходящих линкеров между отдельными ДНК-связывающими компонентами, между ДНК-связывающим доменом и функциональным доменом или между одним или более ДНК-связывающими доменами и между ДНК-связывающим доменом и функциональным доменом композиций, описанных в данном документе.

В. Домены разрезания

Любой подходящий домен разрезания может быть функционально связан с ДНК-связывающими доменами, как описано в данном документе, с образованием нуклеазы. Домен разрезания может быть гетерологичным по отношению к ДНК-связывающему домену, например ДНК-связывающий домен из цинкового пальца и домен разрезания из нуклеазы. Гетерологичные домены разрезания могут быть получены из любой эндонуклеазы или экзонуклеазы. Иллюстративные эндонуклеазы, из которых может быть получен домен разрезания, включают, но не ограничиваются ими, эндонуклеазы рестрикции и эндонуклеазы хоуминга. См., например, Каталог 2002-2003, New England Biolabs, Беверли, Массачусетс; и

Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388. Известны дополнительные ферменты, которые расщепляют ДНК (например, нуклеаза S1; нуклеаза бобов мунг; ДНКазы I поджелудочной железы; микрорНКазы; эндонуклеаза HO дрожжей; 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 разрезания и полудоменов разрезания.

Аналогично, полудомен разрезания может быть получен из любой нуклеазы или ее части, как указано выше, которая требует димеризации для активности разрезания. Как правило, два слитых белка необходимы для разрезания, если слитые белки содержат полудомены разрезания. Альтернативно, можно использовать один белок, содержащий два полудомена разрезания. Два полудомена разрезания могут быть получены из одной и той же эндонуклеазы (или ее функциональных фрагментов), или каждый полудомен разрезания может быть получен из другой эндонуклеазы (или ее функциональных фрагментов). Кроме того, сайты-мишени для двух слитых белков предпочтительно расположены относительно друг друга так, что связывание двух слитых белков с их соответствующими сайтами-мишенями позиционирует полудомены разрезания в пространственной ориентации друг к другу таким образом, что это позволяет полудоменам разрезания формировать функциональный домен разрезания, например, путем димеризации. Таким образом, в определенных вариантах осуществления ближние края сайтов-мишеней разделены 5-8 нуклеотидами или 15-18 нуклеотидами. Однако любое целое число нуклеотидов или нуклеотидных пар может находиться между двумя сайтами-мишенями (например, от 2 до 50 нуклеотидных пар или более). В общем случае, сайт разрезания лежит между целевыми сайтами.

Эндонуклеазы рестрикции (ферменты рестрикции) присутствуют во многих видах и способны специфически связываться с последовательностью ДНК (в сайте узнавания) и расщеплять ДНК в сайте связывания или рядом с ним. Некоторые ферменты рестрикции (например, типа IIS) расщепляют ДНК в сайтах, удаленных от сайта узнавания, и имеют пригодные для разделения домены связывания и разрезания. Например, фермент типа IIS FokI катализирует двухцепочечное разрезание ДНК на расстоянии в 9 нуклеотидов от ее сайта узнавания на одной цепи и на расстоянии в 13 нуклеотидов от ее сайта узнавания на другой цепи. 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 осуществления слитые белки содержат домен разрезания (или полудомен разрезания) по меньшей мере от одного фермента рестрикции типа IIS и один или более доменов связывания цинковых пальцев, которые могут быть или не быть изменены.

Для оптимальной специфичности разрезания сиквенс-селективной (искусственной) нуклеазой желательнее создать условия, чтобы связывание и активность на мишени не были насыщающими. В условиях насыщения по определению - используется избыток нуклеазы по сравнению с необходимым количеством для достижения полной целевой активности. Это превышение не дает преимуществ при воздействии на цель, но все же может привести к увеличению разрезания в нецелевых областях. Для мономерных нуклеаз можно легко избежать условий насыщения, выполнив простое исследование реакции на дозу, чтобы идентифицировать и избежать насыщающего плато на кривой титрования. Однако для димерной нуклеазы, такой как ZFN, TALEN или dCas-FokI, идентификация и избегание условий насыщения могут быть более сложными, если аффинности связывания отдельных мономеров различны. В таких случаях исследование доза-ответ с использованием простого соотношения нуклеаз 1:1 покажет только точку насыщения более слабо связывающего мономера. При таком сценарии, если, например, сродство к мономеру отличается в 10 раз, то в точке насыщения, определенной в исследовании титрования 1:1, мономер с более высоким сродством будет присутствовать в концентрации, которая в 10 раз выше, чем он должен присутствовать. Получающийся в результате избыток мономера с более высокой аффинностью может, в свою очередь, приводить к повышенной активности вне мишени, не обеспечивая какого-либо полезного увеличения разрезания намеченной мишени, что потенциально ведет к снижению общей специфичности для любой данной пары нуклеаз.

Типичным ферментом рестрикции IIS типа, домен разрезания которого отделим от домена связывания, является FokI. Этот конкретный фермент активен в качестве димера. Bitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:10, 570-10,575. Соответственно, для целей данного раскрытия часть фермента FokI, используемая в описанных слитых белках, считается полудоменом разрезания. Таким образом, для целевого двухцепочечного разрезания, целенаправленная замена клеточных последовательностей с использованием слияний цинкового пальца -FokI и комбинаций целевого двухцепочечного разрезания и целенаправленной замены клеточных последовательностей с использованием слияний цинкового пальца - FokI, два слитых белка, каждый из которых содержит полудомен разрезания FokI, могут быть использованы для репарации каталитически активного домена разрезания. Альтернативно, одна полипептидная молекула, содержащая домен связывания цинкового пальца и два полудомена разрезания FokI, тоже может использоваться. Параметры для целевого разрезания и целевого изменения последовательности с использованием слияний цинкового пальца - FokI представлены в других частях данного раскрытия.

Домен разрезания или полудомен разрезания может представлять собой любую часть белка, которая сохраняет активность разрезания или которая сохраняет способность к мультимеризации (например,

димеризации) с образованием функционального домена разрезания.

Типичные рестрикционные ферменты типа IIS описаны в патенте США № 7888121, полностью включенном в данный документ. Дополнительные рестрикционные ферменты также содержат разделяемые связывающие и разрезающие домены, и они предусмотрены настоящим раскрытием (См., например, Roberts et al. (2003) *Nucleic Acids Res.* 31:418-420).

В некоторых вариантах осуществления домен разрезания содержит один или более сконструированных полудоменов разрезания (также называемых мутантами домена димеризации), которые минимизируют или предотвращают гомодимеризацию, как описано, например, в см., например, патенты США №№ 7914796; 8034598 и 8623618, раскрытия которых включены в данное описание в качестве ссылки во всей их полноте. Аминокислотные остатки в позициях 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 и 538 FokI все являются мишенями для воздействия на димеризацию полудоменов разрезания FokI, причем нумерация относится к кристаллическим структурам 1FOK.pdb и 2FOK.pdb (см. Wah et al. (1997) *Nature* 388:97-100), имеющим последовательность, продемонстрированную ниже.

Полудомен разрезания FokI дикого типа (SEQ ID NO: 18)

```
QLVKSELEEKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKH
LGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNNEWK
VYPSSVTEFKFLFVSGHFKGNKAKQLTRLNHNITNCNGAVLSVEELLIIGGEMIKAGTLTLEEVRR
KFNNGEINF
```

Иллюстративные сконструированные полудомены разрезания FokI, которые образуют облигатные гетеродимеры, включают пару, в которой первый полудомен разрезания включает мутации в аминокислотных остатках в позициях 490 и 538 FokI, а второй полудомен разрезания включает мутации в аминокислотных остатках 486 и 499.

Таким образом, в одном варианте осуществления мутация в 490 заменяет Glu (E) на Lys (K); мутация в 538 заменяет Ile (I) на Lys (K); мутация в 486 заменила Gln (Q) на Glu (E); и мутация в позиции 499 заменяет Ile (I) на Lys (K). В частности, сконструированные полудомены разрезания, описанные в данном документе, были получены путем мутации позиций 490 (E → K) и 538 (I → K) в одном полудомене разрезания для получения сконструированного полудомена разрезания, обозначенного "E490K:I538K", и путем мутации позиций 486 (Q → E) и 499 (I → L) в другом полудомене разрезания, чтобы получить сконструированный полудомен разрезания, обозначенный "Q486E:I499L".

Сконструированные полудомены разрезания, описанные в данном документе, являются облигатными гетеродимерными мутантами, у которых абберрантное разрезание с помощью гомодимеров ZFN минимизировано или устранено. См., например, патентную публикацию США № 2008/0131962, раскрытие которой включено в качестве ссылки во всей своей полноте для всех целей. В определенных вариантах осуществления сконструированный полудомен разрезания содержит мутации в позициях 486, 499 и 496 (пронумерованные относительно FokI дикого типа), например мутации, которые заменяют остаток Gln (Q) дикого типа в позиции 486 на Glu (E), остаток Ile (I) дикого типа в позиции 499 остатком Leu (L) и остаток Asn (N) дикого типа в позиции 496 остатком Asp (D) или Glu (E) (также называемые как домены "ELD" и "ELE" соответственно). В других вариантах осуществления сконструированный полудомен разрезания содержит мутации в позициях 490, 538 и 537 (пронумерованные относительно FokI дикого типа), например мутации, которые заменяют остаток Glu (E) дикого типа в позиции 490 на Lys (K), остаток Ile (I) дикого типа в позиции 538 остатком Lys (K) и остаток His (H) дикого типа в позиции 537 остатком Lys (K) или остатком Arg (R) (также называемые доменами "KKK" и "KKR" соответственно). В других вариантах осуществления сконструированный полудомен разрезания содержит мутации в позициях 490 и 537 (пронумерованные относительно FokI дикого типа), например мутации, которые заменяют остаток Glu (E) дикого типа в позиции 490 остатком Lys (K) и остаток His (H) дикого типа в позиции 537 остатком Lys (K) или остатком Arg (R) (также называемые доменами "KIK" и "KIR" соответственно). См., например, патенты США №№ 7914796; 8034598 и 8623618. В других вариантах осуществления сконструированный полудомен разрезания содержит "Sharkey", мутации "Sharkey" и комбинации "Sharkey" и мутаций "Sharkey" (см. Guo et al. (2010) *J. Mol. Biol.* 400(1):96-107).

Таким образом, полудомены разрезания, полученные из FokI, могут содержать мутацию в одном или более аминокислотных остатках, как показано в SEQ ID NO: 18, включая мутации в домене димеризации, как описано выше; мутации в каталитическом домене, мутации в других аминокислотных остатках, таких как фосфат-контактные остатки, и любую комбинацию мутаций, выбранных из мутаций в домене димеризации, мутаций в каталитическом домене и мутаций в других аминокислотных остатках, таких как фосфат-контактные остатки. Мутации включают замены (аминокислотного остатка дикого типа другим остатком), вставки (одного или более аминокислотных остатков), делеции (одного или более аминокислотных остатков) и любую комбинацию мутаций, выбранных из замен, вставок и делеций. В некоторых вариантах осуществления один или более остатков 414-426, 443-450, 467-488, 501-502, 521-531 (пронумерованных относительно SEQ ID NO: 18) и любая комбинация таких остатков мутированы, поскольку эти остатки расположены близко к остову ДНК в молекулярной модели ZFN, связанной с его

сайтом-мишенью, описанной в Miller et al. ((2007) Nat Biotechnol 25:778-784). В определенных вариантах осуществления один или более остатков в позициях 416, 422, 447, 448 и 525 являются мутированными. В некоторых вариантах осуществления мутация включает замену остатка дикого типа другим остатком, например остатком серина (S). В определенных вариантах осуществления домен разрезания FokI нуклеаз, описанных в данном документе, содержит мутацию домена димеризации ELD, мутацию домена димеризации KKR, мутацию K525S и любую комбинацию, выбранную из мутации домена димеризации ELD или мутации домена димеризации KKR и мутации K525S.

Сконструированные домены разрезания, описанные в данном документе, могут быть получены с использованием любого подходящего способа, например, путем сайт-направленного мутагенеза полудоменов разрезания дикого типа (FokI), как описано в патентах США №№ 7888121; 7914796; 8034598 и 8623618. Кроме того, описанные в данном документе домены разрезания могут быть слиты с ДНК-связывающим доменом (например, ZFP) с использованием любого подходящего линкера, включая, но не ограничиваясь этим, линкеры, описанные в патентах США №№ 9394531 и 9567609.

Альтернативно, нуклеазы могут быть собраны *in vivo* на сайте-мишени нуклеиновой кислоты с использованием так называемой технологии "расщепленного фермента" (см., например, публикацию патента США № 2009/0068164). Компоненты таких расщепленных ферментов могут быть экспрессированы либо на отдельных конструкциях экспрессии, либо могут быть связаны в одной открытой рамке считывания, где отдельные компоненты разделены, например, с помощью саморазрезающегося пептида 2A или последовательности IRES. Компоненты могут быть отдельными доменами связывания цинковых пальцев, или доменами домена мегануклеазы, связывающего нуклеиновую кислоту.

Нуклеазы могут быть подвергнуты скринингу на активность перед использованием, например, в хромосомной системе на основе дрожжей, как описано в WO 2009/042163 и 20090068164. Экспрессия нуклеазы может находиться под контролем конститутивного промотора или индуцибельного промотора, например промотора галактокиназы, который активируется (дерепрессируется) в присутствии рафинозы, галактозы и комбинации рафинозы и галактозы и репрессируется в присутствии глюкозы.

Нуклеаза(ы), как описано в данном документе, может производить один или более двухцепочечных, один или более одноцепочечных разрезов и комбинации одного или более двухцепочечных и одного или более одноцепочечных разрезов в целевом сайте. В некоторых вариантах осуществления нуклеаза содержит каталитически неактивный домен разрезания (например, FokI, белок Cas и комбинации FokI и белка Cas). См., например, патенты США №№ 9200266; 8703489 и Guillinger et al. (2014) Nature Biotech. 32(6):577-582. Каталитически неактивный домен разрезания может, в сочетании с каталитически активным доменом, действовать как нуклеаза, образуя одноцепочечный разрез. Следовательно, две нуклеазы могут использоваться в комбинации, чтобы сделать двухцепочечный разрез в определенной области.

Дополнительные нуклеазы также известны в данной области техники, например, McCaffery et al. (2016) Nucleic Acids Res. 44(2):e11. doi: 10.1093/nar/gkv878. Epub 2015 Oct 19.

Целевые сайты

Как подробно описано выше, ДНК-домены могут быть сконструированы так, чтобы связываться с любой выбранной последовательностью. Сконструированный ДНК-связывающий домен может иметь новую специфичность связывания по сравнению с природным ДНК-связывающим доменом. В некоторых вариантах осуществления ДНК-связывающие домены связываются с последовательностью в энхансерной последовательности BCL11A, например, сайт-мишень (обычно 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 или даже больше пар оснований) находится между экзоном 2 и экзоном 3 BCL11A, включая ДНК-связывающие домены, которые связываются с последовательностью в сайте гиперчувствительном к ДНКазе I в последовательности энхансера BCL11A (например, +58), как показано в табл. 1. Методы конструирования включают, но не ограничиваются ими, рациональный дизайн и различные типы отбора. Рациональный дизайн включает в себя, например, использование баз данных, содержащих триплетные (или квадруплетные) нуклеотидные последовательности и отдельные аминокислотные последовательности цинковых пальцев, в которых каждая триплетная или квадруплетная нуклеотидная последовательность ассоциирована с одним или более аминокислотными последовательностями цинковых пальцев, которые связывают последовательность конкретного триплета или квадруплета. См., например, совместные патенты США №№ 6453242 и 6534261, включенные в данный документ в качестве ссылки во всей их полноте. Рациональное проектирование TAL-эффektorных доменов также может быть осуществлено. См., например, патентную публикацию США № 2011/0301073.

Иллюстративные способы отбора, применимые к ДНК-связывающим доменам, включая фаговый дисплей и двухгибридные системы, раскрыты в патентах США № 5789538; 5925523; 6007988; 6013453; 6410248; 6140466; 6200759 и 6242568; а также публикации PCT № WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 и патент Великобритании № GB 2338237. Кроме того, усиление специфичности связывания для доменов связывания цинковых пальцев было описано, например, в совместной заявке WO 02/077227.

Подбор целевых сайтов; нуклеазы и способы проектирования и конструирования слитых белков (и полинуклеотидов, кодирующих их) известны специалистам в данной области и подробно описаны в патентных публикациях США №№ 2005/0064474 и 2006/0188987, полностью включенных в данное описа-

ние посредством ссылки.

Кроме того, как раскрыто в этих и других ссылках, ДНК-связывающие домены (например, многопальцевые белки цинкового пальца) и слияния ДНК-связывающего(их) домена(ов) и функционально(ых) домена(ов) могут быть связаны вместе с использованием любых подходящих линкерных последовательностей, включая, например, линкеры из 5 или более аминокислот. Патенты США №№ 8772453; 7888121 (например, линкер "ZC"); 6479626; 6903185 и 7153949; публикация США № 2009/0305419) и 2015/0064789. Описанные в данном документе белки могут включать любую комбинацию подходящих линкеров между отдельными ДНК-связывающими доменами белка (см. также патент США № 8586526).

Доноры

В определенных вариантах осуществления данное раскрытие относится к нуклеазо-опосредованной целевой интеграции экзогенной последовательности в геном клетки с использованием молекул, связывающих энхансерную область BCL11A, описанные в данном документе. Как отмечалось выше, вставка экзогенной последовательности (также называемой "донорной последовательностью" или "донором" или "трансгеном"), например, для делеции указанной области, для коррекции мутантного гена, для комбинации делеции указанной области и коррекции мутантного гена или для повышенной экспрессии гена дикого типа. Очевидно, что донорная последовательность обычно не идентична геномной последовательности, в которую она вставляется. Донорная последовательность может содержать негомологичную последовательность, фланкированную двумя областями гомологии, чтобы обеспечить эффективный HDR в месторасположении представляющем интерес, или может быть интегрирована с помощью механизмов негомологично направляемой репарации, таких как NHEJ. Кроме того, донорные последовательности могут содержать векторную молекулу, содержащую последовательности, которые не гомологичны области представляющей интерес в клеточном хроматине. Донорная молекула может содержать несколько прерывающихся областей гомологии с клеточным хроматином и, например, приводить к делеции энхансерной области BCL11A (или ее фрагмента) при использовании в качестве субстрата для репарации DSB, индуцированного одной из нуклеаз описанных в данном документе. Кроме того, для направленной вставки последовательностей, которые обычно не присутствуют в области представляющей интерес, указанные последовательности могут присутствовать в молекуле донорной нуклеиновой кислоты и фланкироваться областями гомологии с последовательностью в области представляющей интерес.

Полинуклеотиды для вставки могут также называться "экзогенными" полинуклеотидами, "донорными" полинуклеотидами или молекулами или "трансгенами". Донорный полинуклеотид может быть ДНК или РНК, одноцепочечным или двухцепочечным, и может быть введен в клетку в линейной или циркулярной форме. См., например, публикации патентных заявок США № 2010/0047805 и 2011/0207221. Донорная последовательность(и) предпочтительно содержится в ДНК-МС, которая может быть введена в клетку в циркулярной или линейной форме. При введении в линейной форме концы донорной последовательности могут быть защищены (например, от экзонуклеолитической дегградации) способами, известными специалистам в данной области техники. Например, один или более дидезокси-нуклеотидных остатков добавляются к 3'-концу линейной молекулы, а самокомплементарные олигонуклеотиды необязательно лигированы с одним или обоими концами. См., например, Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls et al. (1996) Science 272:886-889. Дополнительные способы защиты экзогенных полинуклеотидов от дегградации включают, но не ограничиваются ими, добавление концевой аминоксигруппы(групп) и использование модифицированных межнуклеотидных соединений, таких как, например, фосфоротиоаты, фосфорамидаты и остатки О-метил рибозы или дезоксирибозы. При введении в двухцепочечной форме донор может включать один или более сайтов-мишеней нуклеазы, например, сайты-мишени нуклеазы, фланкирующие трансген, который должен быть интегрирован в геном клетки. См., например, публикацию патента США № 2013/0326645.

Полинуклеотид может быть введен в клетку как часть молекулы вектора, имеющей дополнительные последовательности, такие как, например, точки начала репликации, промоторы и гены, кодирующие устойчивость к антибиотикам. Кроме того, донорные полинуклеотиды могут быть введены в виде "голой" нуклеиновой кислоты, в виде нуклеиновой кислоты в комплексе с агентом, таким как липосома или полксамер, или могут быть доставлены вирусами (например, аденовирусом, AAV, герпесвирусом, ретровирусом, лентивирусом и дефектным по интегразе лентивирусом (IDLV)).

В некоторых вариантах осуществления двухцепочечный донор включает в себя последовательности (например, кодирующие последовательности, также называемые трансгенами) длиной более 1 т. п.о., например, между 2 и 200 т. п.о., между 2 и 10 т. п.о. (или любым значением между ними). Двухцепочечный донор также включает, например, по меньшей мере один сайт-мишень нуклеазы. В определенных вариантах осуществления донор включает по меньшей мере 2 сайта-мишени, например, для пары ZFN или TALEN. Как правило, сайты-мишени нуклеазы находятся вне трансгенных последовательностей, например, 5' и/или 3' относительно трансгенных последовательностей для разрезания трансгена. Сайт(ы) разрезания нуклеазой может быть для любой нуклеазы(нуклеаз). В некоторых вариантах осуществления сайт(ы)-мишень(и) нуклеазы(нуклеаз), содержащиеся в двухцепочечном доноре, предназначены для той же самой нуклеазы(нуклеаз), которая используется для разрезания эндогенной мишени, в которую разрезанный донор интегрируется гомологически независимыми способами.

Донор обычно встраивается таким образом, что его экспрессия управляется эндогенным промотором в сайте интеграции, а именно промотором, который управляет экспрессией эндогенного гена, в который встраивается донор (например, глубина, AAVS1 и т.д.). Однако будет очевидно, что донор может содержать промотор, энхансер и комбинации как промотора, так и энхансера, например, конститутивно-го промотора или индуцибельного или тканеспецифичного промотора.

Донорная молекула может быть встроена в эндогенный ген таким образом, что весь эндогенный ген полностью или частично не экспрессируется. В других вариантах осуществления трансген (например, с последовательностями, кодирующими глобин или без них) интегрируется в любой эндогенный локус, например локус "безопасной гавани" (см., например, патентные публикации США №№ 2008/0299580; 2008/0159996 и 2010/00218264).

Кроме того, хотя это и не требуется для экспрессии, экзогенные последовательности могут также включать транскрипционные или трансляционные регуляторные последовательности, например, промоторы, энхансеры, инсуляторы, внутренние сайты входа в рибосомы, последовательности, кодирующие пептиды 2A, сигналы полиаденилирования и их комбинации.

Трансгены, переносимые на донорных последовательностях, описанных в данном документе, могут быть выделены из плазмид, клеток или других источников с использованием стандартных методик, известных в данной области техники, таких как ПЦР. Доноры для использования могут включать в себя различные типы топологии, включая циркулярную суперспиральную, циркулярную релаксированную, линейную и тому подобное. Альтернативно, они могут быть химически синтезированы с использованием стандартных методов синтеза олигонуклеотидов. Кроме того, доноры могут быть метилированы или не иметь метилирования. Доноры могут быть в форме бактериальных или дрожжевых искусственных хромосом (BAC или YAC).

Описанные в данном документе двухцепочечные донорные полинуклеотиды могут включать одно или более неприродных оснований, один или более остовов и комбинации одного или более неприродных оснований и одного или более остовов. В частности, вставка донорной молекулы с метилированными цитозинами может быть осуществлена с использованием способов, описанных в данном документе, для достижения состояния транскрипционного покоя в области представляющей интерес.

Экзогенный (донорный) полинуклеотид может содержать любую представляющую интерес последовательность (экзогенную последовательность). Типичные экзогенные последовательности включают, но не ограничиваются ими, любую кодирующую полипептид последовательность (например, кДНК), промоторные последовательности, энхансерные последовательности, эпитопные тэги, маркерные гены, сайты узнавания фермента разрезания и различные типы конструкций экспрессии. Маркерные гены включают, но не ограничиваются ими, последовательности, кодирующие белки, которые обеспечивают устойчивость к антибиотикам (например, устойчивость к ампициллину, устойчивость к канамицину, устойчивость к неомицину, устойчивость к G418, устойчивость к пурамицину, устойчивость к гигромицину, устойчивость к бластицину), последовательности, кодирующие окрашенные или флуоресцентные или люминесцентные белки (например, зеленый флуоресцентный белок, усиленный зеленый флуоресцентный белок, красный флуоресцентный белок, люциферазу) и белки, которые обеспечивают усиленный рост клеток и амплификацию генов (например, дигидрофолатредуктазу) и комбинации усиленного роста клеток и амплификации генов. Эпитопные тэги включают, например, одну или более копий FLAG, His, Мус, тэг тандемной аффинной очистки (ТАР), HA, биотинилируемый пептид или любую пригодную для обнаружения аминокислотную последовательность.

В предпочтительном варианте осуществления экзогенная последовательность (трансген) содержит полинуклеотид, кодирующий любой полипептид, экспрессия которого в клетке желательна, включая, но не ограничиваясь этим, антитела, антигены, ферменты, рецепторы (клеточные или ядерные), гормоны, лимфокины, цитокины, репортерные полипептиды, факторы роста и функциональные фрагменты любого из вышеперечисленного. Кодирующими последовательностями могут быть, например, кДНК.

Например, экзогенная последовательность может содержать последовательность, кодирующую полипептид, который отсутствует или не функционирует у субъекта, имеющего генетическое заболевание, включая, но не ограничиваясь этим, любое из следующих генетических заболеваний: ахондроплазия, ахроматопсия, дефицит кислой мальтазы, дефицит аденозиндеаминазы (OMIM № 102700), адренолейкодистрофия, синдром Экарди, дефицит альфа-1-антитрипсина, альфа-талассемия, синдром нечувствительности к андрогенам, синдром Аперта, аритмогенный правожелудочковый, дисплазия, атаксия телеангиктазия, синдром Барта, бета-талассемия, синдром синего пузырчатого невуза, болезнь Канавана, хронические гранулематозные заболевания (CGD), синдром кошачьего крика, муковисцидоз, Болезнь Деркума, эктодермальная дисплазия, анемия Фанкони, прогрессирующая фибродисплазия, Синдром Мартина, галактоземия, болезнь Гоше, генерализованный ганглиозидоз, например, GM1), мутация гемоглобина С в 6^м кодоне бета-глобина (HbC), гемофилия, болезнь Хантингтона, Синдром Гурлера, гипофосфатазия, синдром Клайнфелтера, болезнь Краббе, синдром Лангера-Гидиона, дефицит адгезии лейкоцитов (LAD, OMIM № 116920), лейкоцидистрофия, синдром удлиненного интервала QT, синдром Марфана, синдром Мебиуса, мукополисахаридоз (MPS), синдром patella ногтя, нефрогенный несахарный диабет, нейрофиброматоз, болезнь Неймана-Пика, синдром ломких костей, порфирия, синдром Прадера-Вилли, прогерия,

синдром Протея, ретинобластома, синдром Ретта, синдром Рубинштейна-Тайби, синдром Санфилиппо, синдром тяжелого комбинированного иммунодефицита (SCID), синдром Швахмана, серповидноклеточная болезнь (серповидноклеточная анемия), синдром Смита-Магениса, синдром Стиклера, болезнь Тея-Сакса, синдром тромбоцитопении с отсутствием лучевой кости (TAR), синдром Тришера-Коллинза, трисомия, туберозный склероз, синдром Тернера, нарушение цикла мочевины, болезнь Гиппеля - Линдау, синдром Варденбурга, синдром Вильямса, болезнь Вильсона, синдром Вискотга-Олдрича, X-сцепленный лимфопролиферативный синдром (XLP, OMIM № 308240).

Дополнительные иллюстративные заболевания, которые можно лечить путем целенаправленной интеграции, включают приобретенные иммунодефициты, лизосомные болезни накопления (например, болезнь Гоше, GM1, болезнь Фабри и болезнь Тея-Сакса), мукополисахаридоз (например, болезнь Хантера, болезнь Гурлера), гемоглобинопатии (например, серповидноклеточная анемия, HbS, α -талассемия, β -талассемия) и гемофилии.

В некоторых вариантах осуществления экзогенные последовательности могут содержать маркерный ген (описанный выше), позволяющий выбирать клетки, которые подверглись целевой интеграции, и связанную последовательность, кодирующую дополнительную функциональность. Неограничивающие примеры маркерных генов включают GFP, маркер(ы) для селекции антибиотиками и тому подобное.

Дополнительные генные последовательности, которые могут быть вставлены, могут включать, например, гены дикого типа для замены мутированных последовательностей. Например, последовательность гена фактора IX дикого типа может быть вставлена в геном стволовой клетки, в которой мутирована эндогенная копия гена. Копия дикого типа может быть вставлена в эндогенный локус или, в качестве альтернативы, может быть нацелена на локус "безопасной гавани".

Конструирование таких экспрессионных кассет, следуя указаниям данного описания, использует методологии, хорошо известные в области молекулярной биологии (см., например, Ausubel или Maniatis). Перед использованием экспрессионной кассеты для создания трансгенного животного можно проверить чувствительность экспрессионной кассеты к индуктору стресса, связанному с выбранными контрольными элементами, путем введения экспрессионной кассеты в подходящую клеточную линию (например, первичные клетки, трансформированные клетки, или иммортализованные клеточные линии).

Кроме того, хотя это и не требуется для экспрессии, экзогенные последовательности могут также содержать транскрипционные или трансляционные регуляторные последовательности, например, промоторы, энхансеры, инсуляторы, сайты внутренней посадки рибосомы, последовательности, кодирующие пептиды 2A, сигналы полиаденилирования и комбинации полипептидов 2A и сигналов полиаденилирования. Кроме того, контрольные элементы представляющих интерес генов могут быть функционально связаны с репортерными генами для создания химерных генов (например, репортерных кассет экспрессии).

Целевая вставка некодирующей последовательности нуклеиновой кислоты также может быть осуществлена.

Последовательности, кодирующие антисмысловые РНК, РНКи, кшРНК (shRNA) и микроРНК (miRNA), также могут быть использованы для целевых вставок.

В дополнительных вариантах осуществления донорная нуклеиновая кислота может содержать некодирующие последовательности, которые являются специфическими сайтами-мишенями для дополнительных конструкций нуклеаз. Впоследствии дополнительные нуклеазы могут быть экспрессированы в клетках так, что исходная донорная молекула расщепляется и модифицируется путем вставки другой представляющей интерес донорной молекулы. Таким образом, могут генерироваться повторяющиеся интеграции донорных молекул, позволяющие укладывать признаки в конкретном интересующем локусе или в локусе безопасной гавани.

Доставка

Нуклеазы, как описано в данном документе (табл. 1), полинуклеотиды, кодирующие эти нуклеазы, донорные полинуклеотиды и композиции, содержащие белки, полинуклеотиды и комбинации белков и полинуклеотидов, описанные в данном документе, могут доставляться *in vivo* или *ex vivo* любым подходящим способом в любой тип клеток.

Подходящие клетки включают эукариотические (например, животные) и прокариотические клетки и эукариотические и прокариотические клеточные линии. Неограничивающие примеры таких клеток или клеточных линий, полученных из таких клеток, включают COS, CHO (например, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHO-K1SV), VERO, MDCK, WI38, V79, B14AF28-G3, ВНК, НаК, NSO, SP2/0-Ag14, HeLa, HEK293 (например, HEK293-F, HEK293-H, HEK293-T) и клетки perC6, а также клетки насекомых, такие как *Spodoptera fugiperda* (Sf) или грибковые клетки, такие как клетки *Saccharomyces*, *Pichia* и *Schizosaccharomyces*. В некоторых вариантах осуществления клеточная линия представляет собой клеточную линию CHO, MDCK или HEK293. Подходящие клетки также включают стволовые клетки, такие как, например, эмбриональные стволовые клетки, индуцированные плюрипотентные стволовые клетки, гемопоэтические стволовые клетки, нейрональные стволовые клетки и мезенхимальные стволовые клетки.

Способы доставки нуклеаз, как описано в данном документе, описаны, например, в патентах США №№ 6453242; 6503717; 6534261; 6599692; 6607882; 6689558; 6824978; 6933113; 6979539; 7013219 и 7163824, раскрытия всех из которых включены в данный документ посредством ссылки во всей их полноте.

Нуклеазы, донорные конструкции и комбинации нуклеаз и донорных конструкций, как описано в данном документе, также могут быть доставлены с использованием векторов, содержащих последовательности, кодирующие одну или более ZFN, описанных в данном документе. Могут быть использованы любые векторные системы, включая, но не ограничиваясь этим, плазмидные векторы, ретровирусные векторы, лентивирусные векторы, аденовирусные векторы, поксвирусные векторы; герпесвирусные векторы и аденоассоциированные вирусные векторы и др. См. также патенты США № 6534261; 6607882; 6824978; 6933113; 6979539; 7013219 и 7163824, полностью включенные в данное описание посредством ссылки. Кроме того, будет очевидно, что любой из этих векторов может содержать одну или более последовательностей, необходимых для лечения. Таким образом, когда одну или более нуклеаз и донорную конструкцию вводят в клетку, нуклеазы, донорный полинуклеотид и комбинации нуклеаз и донорного полинуклеотида могут переноситься на одном и том же векторе или на разных векторах (ДНК МС). Когда используются несколько векторов, каждый вектор может содержать последовательность, кодирующую одну или более нуклеаз, одну или более донорных конструкций и комбинации одной или большего количества нуклеаз и одной или большего количества донорных конструкций. Обычные способы переноса генов на вирусной и невирусной основе могут быть использованы для введения нуклеиновых кислот, кодирующих нуклеазы, донорные конструкции и комбинаций нуклеаз и донорных конструкций в клетки (например, клетки млекопитающих) и ткани-мишени. Системы доставки не основанные на вирусных векторах включают ДНК или РНК плазмиды, ДНК-МС, ""голую"" нуклеиновую кислоту и нуклеиновую кислоту в комплексе с носителем для доставки, таким как липосома или поллоксамер. Подходящие невирусные векторы включают нанотаксис векторы, включая векторы, коммерчески доступные от InCellArt (Франция). Системы доставки вирусных векторов включают ДНК- и РНК-вирусы, которые имеют или эписомальные или интегрированные геномы после доставки в клетку. Для ознакомления с доставкой *in vivo* сконструированных ДНК-связывающих белков и гибридных белков, содержащих эти связывающие белки, см., например, Rebar (2004) *Expert Opinion Invest. Drugs* 13(7):829-839; Rossi et al. (2007) *Nature Biotech.* 25(12):1444-1454 также как и общие статьи по доставке генов Anderson, *Science* 256:SOS-SIS (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 и 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 Böhm (eds.)* (1995); and Yu et al., *Gene Therapy* 1:13-26 (1994).

Методы невирусной доставки нуклеиновых кислот включают электропорацию, липофекцию, микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, конъюгаты поликаатион или липид:нуклеиновая кислота, ""голую"" ДНК, искусственные вирионы, деформацию мембраны и усиленное агентом поглощение ДНК. Сонопорация с использованием, например, системы Sonitron 2000 (Rich-Mar) также может быть использована для доставки нуклеиновых кислот.

Дополнительные иллюстративные системы доставки нуклеиновых кислот включают системы, предоставляемые Amaha Biosystems (Кельн, Германия), Maxcyte, Inc. (Роквил, Мэриленд), VTX Molecular Delivery Systems (Хьюстон, Массачусетс) и Copernicus Therapeutics Inc. (см., например, US 6008336). Липофекция описана, например, в патентах США №№ 5049386; 4946787 и 4897355 и реагенты для липофекции продаются коммерчески (например, Transfectam™ и Lipofectin™). Катионные и нейтральные липиды, которые подходят для эффективной липофекции полинуклеотидов опосредованной рецепторами, включают липиды Feigner, WO 91/17424, WO 91/16024.

Получение комплексов липид:нуклеиновая кислота, включая целевые липосомы, такие как иммунолипидные комплексы, хорошо известно специалисту в данной области 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). Другие комплексы липид:нуклеиновая кислота включают комплексы, включающие новые катионные липиды, новые пегилированные липиды и комбинации новых катионных липидов и новых пегилированных липидов (см. например, предварительные заявки на патент США №№ 62/432042 и 62/458373).

Дополнительные способы доставки включают использование упаковки нуклеиновых кислот для доставки в средства доставки EnGeneIC (EDV). Эти EDV специфически доставляются в ткани-мишени с использованием биспецифичных антител, причем одно плечо антитела обладает специфичностью к ткани-мишени, а другое специфично для EDV. Антитело доставляет EDV к поверхности клетки-мишени, а затем EDV вводится в клетку путем эндоцитоза. Попадая в клетку, содержимое высвобождается (см. MacDiarmid et al. (2009) *Nature Biotechnology* 27(7):643).

Использование РНК или ДНК систем на вирусной основе для доставки нуклеиновых кислот, коди-

рующих сконструированные системы ZFP, TALE и CRISPR/Cas пользуется преимуществом хорошо разработанных процессов для нацеливания вируса на конкретные клетки организма и передачи груза, который несет вирус в ядро. Вирусные векторы можно вводить непосредственно пациентам (*in vivo*) или их можно использовать для обработки клеток *in vitro* и модифицированные клетки вводят пациентам (*ex vivo*). Обычные вирусные системы для доставки ZFP включают, но не ограничиваются ими, ретровирусные, лентивирусные, аденовирусные, адено-ассоциированные вирусные векторы, векторы вируса коварной оспы и простого герпеса для переноса генов. Интеграция в геном хозяина возможна при использовании методов переноса генов использующих ретровирусы, лентивирусы и аденоассоциированные вирусы, что часто приводит к длительной экспрессии вставленного трансгена. Кроме того, высокая эффективность трансдукции наблюдалась во многих различных типах клеток и тканях-мишенях.

Тропизм ретровируса может быть изменен путем включения чужеродных белков оболочки, что расширяет потенциальную целевую популяцию клеток-мишеней. Лентивирусные векторы представляют собой ретровирусные векторы, которые способны трансдуцировать или инфицировать неделящиеся клетки и обычно продуцируют высокие вирусные титры. Выбор ретровирусной системы переноса генов зависит от ткани-мишени. Ретровирусные векторы состоят из цис- действующих длинных концевых повторов с упаковочной способностью до 6-10 т. п.н. чужеродной последовательности. Минимальные цис-действующие LTR достаточны для репликации и упаковки векторов, которые затем используются для интеграции терапевтического гена в клетку-мишень для обеспечения постоянной экспрессии трансгена. Широко используемые ретровирусные векторы включают векторы, основанные на вирусе мышинного лейкоза (MuLV), вирусе лейкемии гиббоновых обезьян (GaLV), вирусе иммунодефицита обезьян (SIV), вирусе иммунодефицита человека (ВИЧ) и их комбинациях (см., например, Buchscher et al., *J. Virol.* 66:2731-2739 (1992); Johann et al., *J. Virol.* 66:1635-1640 (1992); Sommerfelt et al., *Virology* 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 приложениях, в которых временная экспрессия является предпочтительной, могут использоваться системы на основе аденовируса. Векторы на основе аденовируса способны к очень высокой эффективности трансдукции во многих типах клеток и не требуют деления клеток. С такими векторами получали высокий титр и высокий уровень экспрессии. Этот вектор может быть получен в больших количествах в относительно простой системе. Векторы аденоассоциированного вируса ("AAV") также используются для трансдукции клеток нуклеиновыми кислотами-мишенями, например, в производстве нуклеиновых кислот и пептидов *in vitro*, и для процедур генной терапии *in vivo* и *ex vivo* (см., например, West et al., *Virology* 160: 38-47 (1987); патент США № 4797368; WO 93/24641; Kotin Human Therapy 5: 793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Конструирование рекомбинантных AAV-векторов описано в ряде публикаций, в том числе в патенте 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); and Samulski et al., *J. Virol.* 63:03822-3828 (1989).

По крайней мере, шесть подходов к вирусным векторам в настоящее время доступны для переноса генов в клинических испытаниях, в которых используются подходы, которые включают комплектацию дефектных векторов генами, вставленными в линии вспомогательных клеток для генерации трансдуцирующего агента.

pLASN и MFG-S являются примерами ретровирусных векторов, которые использовались 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, использованным в испытании генной терапии. (Blaese et al., *Science* 270:475-480 (1995)). Эффективность трансдукции 50% или более наблюдалась 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) представляют собой многообещающие альтернативные системы доставки генов, основанные на дефектном и непатогенном аденоассоциированном вирусе типа 2, относящемся к семейству парвовирусов. Все векторы получены из плазмиды, которая содержит только инвертированные концевые повторы AAV длиной 145 п.н., фланкирующие каскету экспрессии трансгена. Эффективный перенос генов и стабильная доставка трансгенов благодаря интеграции в геномы трансдуцированной клетки являются ключевыми характеристиками этой векторной системы. (Wagner et al., *Lancet* 351:9117 1702-3 (1998), Kearns et al., *Gene Ther.* 9:748-55 (1996)). Другие серотипы AAV, включая AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 и AAVrh. 10 и любой новый серотип AAV, также могут быть использованы в соответствии с данным изобретением.

Неспособные к репликации рекомбинантные аденовирусные векторы (Ad) могут продуцироваться с высоким титром и легко инфицировать ряд различных типов клеток. Большинство аденовирусных векторов сконструированы таким образом, что трансген заменяет гены Ad E1a, E1b или E3; впоследствии вектор с дефектом репликации размножается в 293 клетках человека, которые обеспечивают функцию удаленных генов в транс положении. Ad векторы могут трансдуцировать различные типы тканей *in vivo*, включая неделящиеся, дифференцированные клетки, такие как клетки печени, почек и мышц. Обычные Ad векторы имеют большую несущую способность. Пример использования Ad вектора в клинических испытаниях включал полинуклеотидную терапию для противоопухолевой иммунизации с помощью

внутримышечной инъекции (Sterman et al., Hum. Gene Ther. 7:1083-9 (1998)). Дополнительные примеры использования аденовирусных векторов для переноса генов в клинических испытаниях включают Rose-necker et al., Infection 24:1 5-10 (1996); Sterman 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); Sterman et al., Hum. Gene Ther. 7:1083-1089 (1998).

Упаковочные клетки используются для формирования вирусных частиц, которые способны инфицировать клетку-хозяин. Такие клетки включают клетки НЕК293 и Sf9, которые можно использовать для упаковки AAV и аденовируса, и клетки ψ 2 или клетки PA317, которые упаковывают ретровирус. Вирусные векторы, используемые в генной терапии, обычно генерируются линией клеток-производителей, которая упаковывает вектор нуклеиновой кислоты в вирусную частицу. Векторы обычно содержат минимальные вирусные последовательности, необходимые для упаковки и последующей интеграции в хозяина (если применимо), при этом другие вирусные последовательности заменяются экспрессионной кассетой, кодирующей экспрессируемый белок. Отсутствующие вирусные функции обеспечиваются в транс положении упаковочной клеточной линией. Например, векторы AAV, используемые в генной терапии, обычно содержат только последовательности инвертированных концевых повторов (ITR) из генома AAV, которые необходимы для упаковки и интеграции в геном хозяина. Вирусная ДНК упакована в клеточную линию, которая содержит хелперную плазмиду, кодирующую другие гены AAV, а именно *her* и *cap*, но без последовательностей ITR. Клеточная линия также инфицирована аденовирусом в качестве хелпера. Хелперный вирус способствует репликации вектора AAV и экспрессии генов AAV с хелперной плазмиды. Хелперная плазида не упакована в значительных количествах из-за отсутствия последовательностей ITR.

Загрязнение аденовирусом может быть уменьшено, например, тепловой обработкой, к которой аденовирус более чувствителен, чем AAV. В некоторых вариантах осуществления AAV получают с использованием бакуловирусной экспрессионной системы (см., например, патенты США №№ 6723551 и 7271002).

Очистка частиц AAV из 293 или бакуловирусной системы обычно включает рост клеток, которые продуцируют вирус, с последующим сбором вирусных частиц из клеточного супернатанта или лизирование клеток и сбор вируса из неочищенного лизата. AAV затем очищают с помощью методов, известных в данной области техники, включая ионообменную хроматографию (например, смотри патент США №№ 7419817 и 6989264), ионообменной хроматографии и центрифугирование в градиенте плотности CsCl (например, в публикации PCT WO 2011094198A10), иммуноаффинной хроматографии (например, WO 2016128408) или очистки с использованием AVB сефарозы (например, GE Healthcare Life Sciences).

Во многих применениях генной терапии желательно, чтобы вектор генной терапии доставлялся с высокой степенью специфичности к конкретному типу ткани. Соответственно, вирусный вектор может быть модифицирован для придания специфичности для данного типа клеток путем экспрессии лиганда в виде белка слияния с белком оболочки вируса на внешней поверхности вируса. Выбирается лиганд, который имеет сходство с рецептору, о котором известно, что он присутствует в типе клеток представляющих интерес. Например, Han et al., Proc. Natl. Acad. Sci. USA 92: 9747-9751 (1995), сообщили, что вирус лейкоза мышей Молони может быть модифицирован для экспрессии человеческого герегулина слитого с gp70, и указанный рекомбинантный вирус инфицирует некоторые клетки рака молочной железы человека, экспрессирующие рецептор эпидермального фактора роста человека. Этот принцип может быть распространен на другие пары вирус-клетка-мишень, в которых клетка-мишень экспрессирует рецептор, а вирус экспрессирует гибридный белок, содержащий лиганд для рецептора на поверхности клетки. Например, нитевидный фаг может быть сконструирован для выставления на поверхности фрагментов антител (например, FAb или Fv), имеющих специфическую аффинность связывания практически с любым выбранным клеточным рецептором. Хотя приведенное выше описание относится в первую очередь к вирусным векторам, те же принципы могут быть применены к невирусным векторам. Такие векторы могут быть сконструированы так, чтобы они содержали определенные последовательности поглощения, которые способствуют поглощению конкретными клетками-мишенями.

Векторы генной терапии могут доставляться *in vivo* применением к отдельному пациенту, обычно путем системного применения (например, внутривенного, внутрибрюшинного, внутримышечного, подкожного или внутрочерепной инфузии) или местного применения, как описано ниже. Альтернативно, векторы можно доставлять в клетки *ex vivo*, такие как клетки, эксплантированные из отдельного пациента (например, лимфоциты, аспираты костного мозга, биопсия ткани) или универсальные донорские гематопозитические стволовые клетки, с последующей реимплантацией клеток пациенту, обычно после отбора для клеток, которые включили вектор.

Векторы (например, ретровирусные, аденовирусные, липосомные и т.д.), содержащие нуклеазы, донорные конструкции и комбинации нуклеаз и донорных конструкций, также можно вводить непосредственно в организм для трансдукции клеток *in vivo*. Альтернативно, можно вводить "голую" ДНК. Введение осуществляется любым путем, обычно используемым для введения молекулы в окончательный контакт с клетками крови или ткани, включая, без ограничения, инъекцию, инфузию, местное применение

ние и электропорацию. Подходящие способы введения таких нуклеиновых кислот доступны и хорошо известны специалистам в данной области техники, и, хотя для введения конкретной композиции можно использовать более одного пути, конкретный путь часто может обеспечить более немедленную и более эффективную реакцию, чем другой путь.

Векторы, подходящие для введения полинуклеотидов (например, кодирующих нуклеазы, двухцепочечных доноров и комбинаций нуклеазо-кодирующих и двухцепочечных доноров), описанные в данном документе, включают неинтегрирующиеся 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; патентная публикация США № 2009/0117617.

Фармацевтически приемлемые носители частично определяются конкретной вводимой композицией, а также конкретным способом, применяемым для введения композиции. Соответственно, существует большое разнообразие доступных подходящих составов фармацевтических композиций, как описано ниже (см., например, Remington's Pharmaceutical Sciences, 17th ed., 1989).

Будет очевидно, что кодирующие нуклеазы последовательности и донорные конструкции могут быть доставлены с использованием одной и той же или разных систем. Например, нуклеазы и доноры могут переноситься одной и той же ДНК МС. Альтернативно, донорный полинуклеотид можно перенести с помощью МС, тогда как одну или более нуклеаз может переноситься с помощью стандартной плазмиды или вектора AAV. Кроме того, разные векторы можно вводить одним и тем же или разными путями (внутримышечная инъекция, инъекция в хвостовую вену, другая внутривенная инъекция, внутрибрюшинное введение или внутримышечная инъекция). Векторы могут быть доставлены одновременно или в любом последовательном порядке.

Таким образом, данное раскрытие включает лечение *in vivo* или *ex vivo* заболеваний и состояний, которые поддаются вставке трансгенов, кодирующих терапевтический белок. Композиции вводят пациенту-человеку в количестве, эффективном для получения желаемой концентрации терапевтического полипептида в сыворотке или целевом органе или клетках. Введение может осуществляться любыми способами, в которых полинуклеотиды доставляются в желаемые клетки-мишени. Например, рассматриваются как *in vivo*, так и *ex vivo* способы. Внутривенная инъекция в воротную вену является предпочтительным способом введения. Другие способы введения *in vivo* включают, например, прямую инъекцию в доли печени или желчного протока и внутривенную инъекцию, дистальную по отношению к печени, в том числе через печеночную артерию, прямую инъекцию в паренхиму печени, инъекцию через печеночную артерию, и ретроградную инъекцию через желчные протоки. Способы введения *ex vivo* включают трансдукцию *in vitro* вырезанных гепатоцитов или других клеток печени с последующей инфузией трансдуцированных вырезанных гепатоцитов обратно в портальную сосудистую сеть, паренхиму печени или билиарное дерево пациента-человека, см., например, Grossman et al. (1994) Nature Genetics, 6: 335-341.

Эффективное количество вводимых нуклеазы(нуклеаз) и донора будет варьироваться от пациента к пациенту и в зависимости от терапевтического полипептида, представляющего интерес.

Соответственно, эффективные количества лучше всего определяются врачом, вводящим композицию, и подходящие дозы могут быть легко определены специалистом в данной области техники. После предложения достаточного времени для интеграции и экспрессии (обычно, например, 4-15 дней), анализ уровней терапевтического полипептида в сыворотке или других тканях и сравнение с начальным уровнем до введения будет определять, является ли вводимое количество слишком низким, в правильном диапазоне или слишком высоким. Подходящие режимы для начального и последующего введений также являются переменными, но они типизируются по первоначальному введению, за которым следуют последующие введения, если необходимо. Последующие введения могут осуществляться через различные интервалы времени, от ежедневного до ежегодного до каждых нескольких лет. Специалисту в данной области техники будет понятно, что соответствующие иммуносупрессивные методы могут быть рекомендованы, чтобы избежать ингибирования или блокирование трансдукции с иммуносупрессией векторов доставки, см., например, Vilquin et al., (1995) Human Gene Ther., 6:1391-1401.

Композиции для введения как *ex vivo*, так и *in vivo* включают суспензии в жидкостях или эмульгированных жидкостях. Активные ингредиенты часто смешивают с наполнителями, которые являются фармацевтически приемлемыми и совместимыми с активным ингредиентом. Подходящие наполнители включают, например, воду, физиологический раствор, декстрозу, глицерин, этанол или тому подобное и их комбинации. Кроме того, композиция может содержать незначительные количества вспомогательных веществ, таких как смачивающие или эмульгирующие агенты, pH-буферизирующие агенты, стабилизирующие агенты или другие реагенты, которые повышают эффективность фармацевтической композиции.

Клетки

Также в данном документе описаны клетки и клеточные линии, в которых эндогенная энхансерная последовательность BCL11A модифицирована нуклеазами, описанными в данном документе (табл. 1). Модификация может быть, например, по сравнению с последовательностью дикого типа клетки. Клетка или клеточные линии могут быть гетерозиготными или гомозиготными для модификации. Модификации последовательности BCL11A могут включать в себя инделы.

Модификация предпочтительно происходит в или около сайта(ов) связывания нуклеазы(нуклеаз), сайта(ов) разрезания и комбинаций сайта(ов) связывания и сайта(ов) разрезания, например, в пределах 1-300 (или любого значения между ними) пар оснований вышерасположенных или нижерасположенных по отношению к сайту(ам) разрезания, более предпочтительно в пределах 1-100 пар оснований (или любого значения между ними) с любой стороны от сайта(ов) связывания, сайта(ов) разрезания или сайта(ов) связывания и сайт(ов) разрезания, еще более предпочтительно в пределах от 1 до 50 пар оснований (или любого другого значения между ними) на любой стороне сайта(ов) связывания, сайта(ов) разрезания или сайта(ов) связывания и сайта(ов) разрезания. В некоторых вариантах осуществления модификация находится в или около области "+58" энхансера BCL11A, например, в или рядом с сайтом связывания нуклеазы, показанном в любой ячейке первого столбца табл. 1.

Любая клетка или клеточная линия может быть модифицирована, например, стволовая клетка, например, эмбриональная стволовая клетка, индуцированная плюрипотентная стволовая клетка, гемопоэтическая стволовая клетка, нейрональная стволовая клетка и мезенхимальная стволовая клетка. Другие неограничивающие примеры клеточных линий, как описано в данном документе, включают Т-клетки (например, CD4+, CD3+, CD8+ и т.д.); дендритные клетки; В-клетки. Также предлагается потомок стволовой клетки, включая частично или полностью дифференцированную клетку (например, клетку-предшественник RBC или RBC). Неограничивающие примеры других клеточных линий, включая модифицированную последовательность BCL11A, включают COS, CHO (например, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHO-K1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NSO, SP2/0-Ag14, HeLa, HEK293 (например, HEK293-F, HEK293-N, HEK293-T) и клетки репС6, а также клетки насекомых, таких как *Spodoptera fugiperda* (Sf), или грибковые клетки, такие как *Saccharomyces*, *Pichia* и *Schizosaccharomyces*.

Описанные в данном документе клетки полезны для лечения или профилактики расстройства, например, с помощью терапии *ex vivo*. Клетки, модифицированные нуклеазой, могут быть размножены и затем повторно введены пациенту с использованием стандартных методик (см, например, Tebas et al. (2014) *New Eng J Med* 370(10):901). В случае стволовых клеток, после инфузии субъекту также происходит дифференцировка *in vivo* этих предшественников в клетки, экспрессирующие функциональный трансген. Фармацевтические композиции, содержащие клетки, как описано в данном документе, также предлагаются. Кроме того, клетки могут быть криоконсервированы до введения пациенту.

Любые из модифицированных клеток или клеточных линий, раскрытых в данном документе, могут демонстрировать повышенную экспрессию гамма-глобина. Композиции, такие как фармацевтические композиции, содержащие генетически модифицированные клетки, как описано в данном документе, также предлагаются.

Применения

Способы и композиции, раскрытые в данном документе, предназначены для модификации экспрессии белка или для исправления аберрантной последовательности гена, которая кодирует белок, экспрессируемый при генетическом заболевании, таком как серповидноклеточная анемия или талассемия. Таким образом, способы и композиции обеспечивают лечение или профилактику таких генетических заболеваний. Редактирование генома, например стволовых клеток, может использоваться для коррекции аберрантного гена, вставки гена дикого типа или изменения экспрессии эндогенного гена. В качестве неограничивающего примера, ген дикого типа, например, кодирующий по меньшей мере один глобин (например, α -глобин, γ -глобин, β -глобин и их комбинации), могут быть вставлены в клетки (например, в эндогенную последовательность BCL11A энхансера с использованием одной или более нуклеаз, как описано в данном документе), чтобы обеспечить белки глобина присутствующие в пониженном количестве или отсутствующие в клетке и, таким образом, лечение генетического заболевания, например гемоглобинопатии, вызванной неправильной экспрессией глобина. Альтернативно или в дополнение, геномное редактирование с или без введения соответствующего донора может исправить дефектный эндогенный ген, например, исправить точечную мутацию в α - или β -гемоглобине, восстановить экспрессию гена или вылечить генетическое заболевание, например, серповидно-клеточную анемию, нок аут или изменение (сверхэкспрессию или подавление) какого-либо гена прямой или непрямой регуляции глобина (например, инактивация гена-регулятора γ -глобина BCL11A или регулятора BCL11A KLF1). В частности, способы и композиции по данному изобретению находят применение при лечении или профилактике гемоглобинопатии.

Нуклеазы по данному изобретению нацелены на энхансерную область BCL11A, которая, как известно, необходима для экспрессии BCL11A во время эритропоэза и, следовательно, для подавления экспрессии гамма-глобина. Модификация этой области энхансера может привести к появлению эритроцитов с повышенной экспрессией гамма-глобина и, таким образом, может быть полезна для лечения или профилактики серповидноклеточной анемии или бета-талассемии.

Следующие примеры относятся к иллюстративным вариантам осуществления данного раскрытия, в которых нуклеаза содержит нуклеазу цинкового пальца (ZFN). Понятно, что это сделано только для иллюстративных целей и что могут быть использованы другие нуклеазы, например, системы TtAgo и

CRISPR/Cas, хоуминг эндонуклеазы (мегануклеазы) со сконструированными ДНК-связывающими доменами, слияния встречающихся в природе ДНК-связывающих доменов сконструированных хоуминг эндонуклеаз (мегануклеаз), в том числе комбинации хоуминг эндонуклеаз (мегануклеаз) со сконструированными ДНК-связывающими доменами и слияния встречающихся в природе ДНК-связывающих доменов сконструированных хоуминг эндонуклеаз (мегануклеаз) и гетерологичных доменов разрезания, слияния мегануклеаз и белков TALE, включая комбинации гетерологичных доменов разрезания и слияния мегануклеаз и белков TALE.

Примеры

Пример 1. Сборка нуклеаз цинковых пальцев

ZFN собирали против человеческого гена BCL11A, и активность проверяли путем глубокого секвенирования ДНК, выделенной из трансфицированных клеток, как описано ниже. ZFN, специфичные для +58 области энхансерной области получали как описано. Пара ZFN 51857/51949 была описана ранее (см. WO 2016/183298).

Пример 2. Анализ вне мишени.

Для анализа нецелевого разрезания парами ZFN был проведен двухэтапный непредвзятый анализ специфичности. На первом этапе (фиг. 1) сайты-кандидаты, не являющиеся мишенями, для каждого ZFN были идентифицированы с помощью анализа сайта интеграции олигонуклеотидного дуплекса с использованием процедуры, аналогичной описанной Tsai et al. ((2015), *Nat Biotechnol* 33(2):187-197. doi:10.1038/nbt.3117).

Анализ сайта интеграции олигонуклеотидного дуплекса основан на наблюдении, что совместное введение нуклеазы и короткого сегмента дуплексной ДНК в клетку-мишень приводит к интеграции дуплекса во время репарации части явлений разрезания генома через путь репарации ДНК NHEJ (Орландо et al. (2010), *Nucleic Acids Res*, 38 (15) e152. doi: 10.1093/nar/gkq512; Gabriel et al. (2011), *Nat Biotechnol*. 2011 Aug 7; 29(9):816-23. doi: 10.1038/nbt.1948; Tsai et al., *ibid*). После интеграции дуплекс обеспечивает постоянный тэг этого конкретного явления разрезания. Сайты интеграции затем идентифицируют путем лигирования олигонуклеотидного адаптера к сдвинутой геномной ДНК с последующим 2 раундами 25 циклов гнездовой ПЦР и глубоким секвенированием полученных сочленений донора и генома. Этот анализ позволяет оценить все потенциальные сайты интеграции в геноме.

Анализ сайтов интеграции проводили в клетках K562 для максимизации доставки донора, экспрессии ZFN и интеграции донора. Более того, поскольку клетки K562 быстро делятся (время удвоения приблизительно 24 ч), ожидается, что они наложат минимальные эпигенетические ограничения на способность ZFN расщеплять клеточные мишени. Клетки (2×10^5) подвергали электропорации с 0,47 мкг олигонуклеотидного дуплексного донора и 400 нг каждой мРНК, кодирующей ZFN, с использованием челнока Атаха и настроек, оптимизированных для максимальной целевой активности ZFN. Четыре репликативных образца были приготовлены для каждой комбинации олиго и мРНК. На 7-е сутки после трансфекции выделяли геномную ДНК для каждого образца (набор Qiagen DNeasy для крови и тканей) и 400 нг (133000 гаплоидных геномов) было использовано в качестве исходного материала для протокола амплификации, изображенного на фиг. 1. Образцы затем обрабатывали по существу, как было описано раньше (Tsai et al. *Ibid*). Конечные продукты были объединены, количественно было определено и они были секвенированы на приборе MiSeq (Illumina) с использованием набора для секвенирования v2 на 300 циклов с парными считываниями (paired-end reads) по 150 п.н. и считываниями с двойными индексирующими последовательностями (dual index reads) по 8 п.н. /16 п.н. для обнаружения штрих-кодов образцов на каждом конце ампликона.

Чтобы сгенерировать список возможных вне целевых сайтов, данные секвенирования были отфильтрованы на предмет наличия правильной затравочной последовательности с последующим обрезанием последовательностей адаптера и картированием в геноме. Затем были картированы координаты соединений, и соединение дуплекс-геном, также как и позиция разрыва, вызванного сдвигом ДНК, были использованы для идентификации различных явлений интеграции. Затем явления интеграции были обработаны, чтобы идентифицировать кластеры интеграции в непосредственной близости в геноме (минимум 4 различных явления интеграции в пределах 100 п.н. друг от друга, суммированные по всем репликам). Кластеры, находящиеся на контигах, которые нельзя было картировать в сборке hg38 (т.е. chrUn в hg38), были удалены из дальнейшего анализа. Картирование кластеров в повторяющихся локусах (медиана трех или более попаданий в геном по всем последовательностям в кластере) также было удалено, так как предыдущий опыт показал, что это артефакты амплификации. Оставшиеся кластеры оценивали как сайты-кандидаты для разрезания ZFN, если они были получены по меньшей мере из 2 повторов образцов, обработанных ZFN (всего 4), и демонстрировали ≥ 5 -кратное превышение явлений интеграции в образцах, обработанных ZFN, по сравнению с контролем. Сайты-кандидаты для разрезания классифицировали по общему количеству уникальных интеграции в образцах, обработанных ZFN. Локусы-кандидаты, идентифицированные с помощью этого анализа, представлены на фиг. 2 для пары ZFN 51857/51949, классифицированные по количеству интегрантов.

Пример 3. Оптимизация ZFN

Чтобы уменьшить разрезание вне мишени, адаптировали стратегию оптимизации нуклеазы, в которой неспецифические фосфатные контакты выборочно удаляются, чтобы вызвать глобальное подавление разрезания вне мишени (Guilinger et al. (2014) *Nat Methods*. 11(4):429-35. doi:10.1038/nmeth.2845; Kleinstiver et al. (2016) *Nature* 529(7587):490-5. doi:10.1038/nature16526; Slaymaker et al. (2016) *Science* 351(6268): 84-8. doi:10.1126/science.aad5227) (см. предварительные заявки США №№ 62/443981 и 62/378978). Аминокислотные замены были сделаны в ключевой позиции внутри каркаса цинкового пальца, который взаимодействует с фосфатным остовом ДНК (Pavletich and Pabo, (1991) *Science* 252(5007):809-17; Elrod-Erickson et al. (1996) *Structure* 4(10):1171-80) (фиг. 3А-3В), а также в одной позиции в правом домене FokI ZFN которая также, как спрогнозировано, создает фосфатный контакт (фиг. 3С).

Специфичность была дополнительно улучшена благодаря возможности независимой экспрессии каждого ZFN из двух разделенных мРНК, что позволяет оптимизировать соотношения доставки. Эти усилия позволили получить оптимизированные пары ZFN, которые высоко родственны исходной паре, отличаясь заменами, которые снижают энергетику взаимодействия с фосфатным остовом ДНК, но которые минимально влияют или не влияют на сиквенс-специфическое распознавание оснований. В соответствии с этим, анализ сайта интеграции дал 455 локусов для потенциальных мишеней разрезания ZFN для исходной пары 51857/51949. Для оптимизированной пары с помощью этого анализа было определено гораздо меньшее количество локусов для дальнейшей проверки в качестве потенциальных мишеней разрезания ZFN (всего 72). Для обеих пар предполагаемой мишенью в энхансере BCL11A был локус высшего ранга. Более того, в энхансере BCL11A для оптимизированной пары была отмечена гораздо более высокая доля явлений интеграции, что согласуется с ее большей специфичностью.

Важно отметить, что при определении конвейера обработки данных последовательностей, ключевые параметры были выбраны консервативно, чтобы включить максимально возможное количество локусов вне мишени, а не отфильтровать их. Это было сделано для того, чтобы гарантировать, что каждый локус, который может представлять собой хороший сайт разрезания для оптимизированных ZFN, будет идентифицирован и протестирован в последующих исследованиях инделов, даже за счет принятия гораздо большего числа локусов, которые станут ложноположительными. Ожидалось, что на первом этапе анализа будет получен большой набор локусов-кандидатов для каждой пары ZFN, из которых подавляющее большинство (особенно для оптимизированных ZFN) не будут представлять истинные сайты разрезания вне мишени, а скорее фоновые явления, которые могут оказаться отрицательными для разрезания в последующих исследованиях инделов.

На втором этапе анализа локусы-кандидаты вне мишени, идентифицированные с помощью анализа сайта интеграции, были подвергнуты скринингу на наличие признаков модификации (например, наличия инделов) в CD34+ HSPC, обработанных ZFN.

В частности, CD34+ HSPC человека, полученные из мобилизованной периферической крови, обрабатывали исходными и оптимизированными парами ZFN с использованием клинической шкалы и клинических условий для трансфекции РНК (120 мкг/мл мРНК для исходной пары ZFN и 100 мкг/мл мРНК для оптимизированной пары). Геномную ДНК выделяли через 2 дня после трансфекции с последующей ПЦР-амплификацией локусов-кандидатов вне мишени и глубоким секвенированием для количественного определения уровней инделов. Как на исходной, так и на оптимизированной парах ZFN, на этом этапе был проведен скрининг одного и того же набора из 137 потенциальных локусов вне мишени, а также меньшее количество потенциальных сайтов вне мишени, которые были идентифицированы с помощью других методов в более ранних исследованиях с исходными ZFN.

Результаты показали, что оптимизированные ZFN заметно более специфичны, чем исходная пара. Это видно не только из числа локусов, которые были оценены как положительные на наличие доказательства разрезания ZFN (52 для исходной пары против 3 для оптимизированной пары), но также и из наблюдаемых уровней инделов, которые для оптимизированной пары были намного ниже. На фиг. 4 показаны графики значений инделов в каждом локусе, демонстрирующем признаки разрезания ZFN в этом исследовании (отметьте логарифмическую шкалу оси Y).

Суммирование инделов вне мишени по всем таким локусам указывает на снижение активности вне мишени в 300 раз (совокупные инделы вне мишени для исходной пары составили 46,5%, против инделов вне мишени для оптимизированной пары - 0,15%). Это снижение активности вне мишени было достигнуто без потери активности на предполагаемом сайте-мишени (72,5% инделов для исходной пары против 81,9% для оптимизированных ZFN). В этих исследованиях исходная пара (или родительская пара) была 51857/51949, тогда как оптимизированная пара ZFN была 63014/65722 (см. ниже). Конструкции нуклеаз показаны ниже в табл. 1.

Таблица 1. Пары ZFN, специфичные для области энхансера +58 BCL11A

SBS № (сайт-мишень, 5'-3')	Конструкция [Последовательность спирали, SEQ ID]						Линкер
	[Мутации в остове пальца]						Мутанты Fok
	F1	F2	F3	F4	F5	F6	
Левый партнер							
51857 aaAGCAActGTTA GCTTGCActagac ta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:19)	RNFSL TM (SEQ ID NO:20)	STG NLT N (SE Q ID NO: 21)	TSGSL TR (SEQ ID NO:22)	DQSNL RA (SEQ ID NO:19)	AQ CC LF H (S EQ ID NO :2 3)	L7c5
	нет	нет	нет	нет	нет	нет	ELD
63014 aaAGCAActGTTA	DQSNLRA (SEQ ID	RNFSL TM(SE	STG NLT	TSGSL TR(SE	DQSNL RA(SE	AQ CC	L7c5
GCTTGCActagac ta (SEQ ID NO:1)	NO:19)	Q ID NO:20)	N (SE Q ID NO: 21)	Q ID NO:22)	Q ID NO:19)	LF H(SE Q ID NO :2 3)	
	Qm5	нет	Qm5	нет	Qm5	нет	ELD
65459 aaAGCAActGTTA GCTTGCActagac ta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:19)	RNFSL TM(SE Q ID NO:20)	STG NLT N (SE Q ID NO: 21)	TSGSL TR(SE Q ID NO:22)	DQSNL RA(SE Q ID NO:19)	AQ CC LF H(SE Q ID NO :2 3)	L7c5
	Qm14Qm5	нет	Qm5	нет	Qm5	нет	ELD
Правый партнер							
51949 caCAGGCTCCAGG AAGGgtttgcct ct (SEQ ID NO: 2)	RNDHRTT (SEQ ID NO: 24)	QKAHL IR (SEQ ID NO: 25)	QKC TLG E (SE Q ID NO: 26)	RGRDL SR (SEQ ID NO: 26)	RRDNL HS (SEQ ID NO: 27)	H/ D	L0
	нет	нет	нет	нет	нет	H/ D	KKR
65722 caCAGGCTCCAGG	RNDHRTT (SEQ ID	QKAHL IR	QKC TLG	RGRDL SR	RRDNL HS	H/ D	L0

AAGGgtttgcct ct (SEQ ID NO: 2)	NO: 24)	(SEQ ID NO: 25)	E (SE Q ID NO: 26)	(SEQ ID NO: 26)	(SEQ ID NO: 27)		
	Qm5	Qm5	нет	Qm5	нет	H/ Д	KKR K525S
65526 caCAGGCTCCAGG AAGGgtttgcct ct (SEQ ID NO: 2)	RNDHRRTT (SEQ ID NO: 24)	QKAHL IR (SEQ ID NO: 25)	QKG TLG E (SE Q ID NO: 26)	RGRDL SR (SEQ ID NO: 26)	RRDNL HS (SEQ ID NO: 27)	H/ Д	L0
	Qm5	Qm5	нет	Qm5	Qm5	H/ Д	KKR R416S
65549 caCAGGCTCCAGG AAGGgtttgcct ct (SEQ ID NO: 2)	RNDHRRTT (SEQ ID NO: 24)	QKAHL IR (SEQ ID NO: 25)	QKG TLG E (SE Q ID NO: 26)	RGRDL SR (SEQ ID NO: 26)	RRDNL HS (SEQ ID NO: 27)	H/ Д	L0
	Qm5	Qm5	Qm5	Qm5	нет	H/ Д	KKR K525S
65550 caCAGGCTCCAGG AAGGgtttgcct ct (SEQ ID NO: 2)	RNDHRRTT (SEQ ID NO: 24)	QKAHL IR (SEQ ID NO: 25)	QKG TLG E SEQ ID NO: 26	RGRDL SR (SEQ ID NO: 26)	RRDNL HS (SEQ ID NO: 27)	H/ Д	L0
	Qm5	Qm5	нет	Qm5	Qm5	H/ Д	KKR K525S

Табл. 1 показывает характеризующую информацию, относящуюся к каждому ZFN. Начиная слева, SBS номер (например, 51857) отображается с ДНК-мишенью, с которой связывается ZFN, показанной под SBS номером. Далее показаны аминокислотные конструкции спирали распознавания для пальцев 1-6 или 1-5 (разделенный столбец 2 табл. 1). В табл. 1 под соответствующими конструкциями спирали также показаны мутации, сделанные в последовательностях остова ZFP указанного пальца, как описано в предварительных патентных заявках США №№ 62/378978 и 62/443981. В обозначениях, использованных в табл. 1, "Qm5" означает, что в позиции минус 5 (относительно спирали, которая пронумерована от -1 до +6) указанного пальца, аргинин в этой позиции был заменен глутамином (Q), в то время как "Qm14" означает, что аргинин (R), как правило, присутствующий в позиции минус 14 был заменен на глутамин (Q). "Нет" означает отсутствие изменений за пределами области спирали распознавания. Таким образом, например, SBS № 63014 включает мутацию Qm5 в пальцах 1, 3 и 5, в то время как пальцы 2, 4 и 6 не имеют мутаций в остове цинкового пальца (например, последовательность цинковых пальцев вне области спирали распознавания).

Наконец, в самом правом столбце табл. 1 показан линкер, используемый для связывания ДНК-связывающего домена с доменом разрезания FokI (например, "L7c5" (LRGSISRARPLNPHP (SEQ ID NO: 5), как описано, например, в патенте США № 9567609 отображается в верхней строке столбца, а сайты мутаций FokI в контакте с фосфатами и димеризационных мутаций показаны в рамке под обозначением линкера. В частности, в верхней строке рамки мутантов FokI указан тип мутации, обнаруженной в домене димеризации (например, ELD или KKR, как описано, например, в патенте США № 8962281). Ниже обозначений димеризационных мутантов показаны любые мутации, присутствующие в домене FokI, сделанные для удаления неспецифического фосфатного контакта, показанного внизу (например, K525S или R416S, где остатки серина в позициях аминокислот 525 или 416 были заменены либо лизином, либо аргинином, соответственно, как описано в предварительной заявке на патент США №№ 62/378978 и

62/443981). Так, например, в SBS № 63014, линкер представляет собой линкер L7c5, а домен разрезания FokI включает в себя мутанты димеризации ELD и не содержит фосфат-контактных мутаций. Кроме того, для SBS № 65722 линкер представляет собой линкер L0 (LRGSQLVKS (SEQ ID NO: 6), также называемый "стандартным" линкером, см. патент США № 9567609), и домен разрезания FokI включает мутации димеризации KKR, и фосфат-контактную мутацию K525S FokI.

Все ZFN были протестированы на функциональность (активность разрезания, как определено анализом на наличие инделов, как описано в примере 4 ниже) и были определены как активные.

Кроме того, чтобы определить, какие конструкции ZFN были наиболее специфичными, в обработанных ZFN CD34+ HSPC были проведены анализы инделов известных сайтов разрезания вне мишени исходной парой ZFN. Для этого человеческие CD34+ HSPC, полученные из мобилизованной периферической крови, обрабатывали исходными и оптимизированными парами ZFN с использованием клинических условий и концентраций мРНК (120 мкг/мл для исходной пары ZFN и 100 мкг/мл для оптимизированной пары). Геномную ДНК выделяли через 2 дня после трансфекции из этих клеток и необработанных контролей с последующей ПЦР-амплификацией каждого локуса-кандидата и глубоким секвенированием для количественного определения уровней инделов.

Уровни модификации в каждом локусе определяли путем глубокого секвенирования парных концов (paired-end deep sequencing) на Illumina MiSeq с использованием картриджа на 300 циклов. Парные последовательности были объединены, адапторы обрезаны через SeqPrep, отфильтрованы для оценки качества ≥ 15 по всем основаниям, а затем картированы в геном человека (сборка hg38). Последовательности, картированные в неправильный локус, были отброшены. Последовательности короче чем ампликон дикого типа, на >70 п.н. или $>70\%$ были удалены для минимизации продуктов праймер-димер. Выравнивание по Нидлману-Вуншу (Needleman and Wunsch, (1970), J Mol Biol 48(3):443-53) выполняли между целевым ампликоном и каждым прочтением MiSeq для картирования инделов. Инделов в выровненных последовательностях были определены, как описано в Gabriel et al. 2011 (там же), за исключением того, что были приняты индексы длиной 1 п.о., чтобы избежать недооценки реальных явлений. Обратите внимание, что часть локусов либо не амплифицировалась, либо не секвенировалась, либо была отклонена из анализа из-за высокого фона ($>1\%$ модификации в контрольных образцах) или недостаточной глубины секвенирования (<10000 считываний). Результаты этого анализа и сравнения с "родительской" парой 51857/51949 ZFN представлены ниже в табл. 2.

Таблица 2. Анализ разрезания вне мишени

Димер ZFN		мкг РНК Л:П	VCL11A % инделов	Вне мишени: количество локусов			количество индел-положительных локусов ВМ (вне мишени)		
Левый	Правый			мишень	ПЦР	Пригодно для анализа	P<0,05	ручная	В родителем? Захват/подтверждено
51857	51949	60:60	73,0	23	21	17	15	17	-
63014	65722	60:15	82,2	31	24	14	0	4	4/3
63014	65526	60:15	81,2	23	22	9	3	4	4/4
63014	65527	60:60	81,4	30	24	10	4	4	4/2
63014	65549	60:60	80,0	30	24	13	0	2	2/1
63014	65550	60:60	79,8	30	24	9	0	1	1/1
65459	65526	60:15	76,9	23	19	14	0	2	2/2

Пример 4. Активность ZFN в CD34+ клетках человека

Для тестирования *in vitro* нуклеазы тестировали в CD34+ клетках. ZFN поставляли в виде мРНК, где мРНК получали *in vitro* следующим образом: плазмиды, содержащие гены, кодирующие ZFN, линейризуют и используют для транскрипции мРНК *in vitro* с использованием набора mMessage mMachine® T7 Ultra (Ambion/Applied Biosystems). Затем мРНК очищали с использованием мини-набора RNeasy® (Qiagen). CD34+ клетки выделяли из мобилизованной периферической крови и поддерживали в среде X-VIVO 10, дополненной пенициллином, стрептомицином и глютамином, а также StemSpan CC110, и инкубировали при 37°C и 5% CO₂. Клетки трансфицировали через 48 ч после выделения или после оттаивания. Небольшую аликвоту смешивали в соотношении 1:1 с раствором трипанового синего 0,4% (вес/объем) в PBS (Corning), и количество клеток определяли на автоматическом счетчике клеток TC20 (Bio-Rad).

Для крупномасштабных трансфекций клетки промывали буфером электропорации MaxCyte (Maxcyte) и ресуспендировали при $3\text{-}5 \times 10^7$ клеток на мл в буфере электропорации в 100 мкл. Как правило, концентрации мРНК между 60 и 120 мкг/мл использовались для скрининга подходящих наборов ZFN. Затем клетки выращивали в среде для выращивания в 3×10^6 клеток на мл в течение 18 ч при температуре 30°C , а затем разводили до 1×10^6 клеток на мл в течение дополнительных 24 ч при 37°C . Для определения активности разрезания геномную ДНК выделяли через 2-3 дня после трансфекций, и уровень модификации гена в локусе энхансера BCL11A измеряли путем глубокого секвенирования на секвенаторе MiSeq (Illumina).

Пары ZFN из табл. 1 были протестированы в CD34+ клетках, и результаты активности показаны ниже в табл. 3.

Таблица 3. Активность пар ZFN против мишени BCL11A

Правый ZFN	Левый ZFN	П ZFN конц. (мкг)	Л ZFN конц. (мкг)	Инделлы (%)
51857	51949	60	60	72,98
63014	65722	15	60	80,62
63014	65722	60	15	82,19
63014	65526	60	15	81,22
63014	65527	60	60	81,43
63014	65549	60	60	79,82
63014	65550	60	60	79,96
GFP-контроль				0,07

В дополнение к анализу нуклеазной активности в CD34+ клетках до эритроидной дифференцировки, отредактированные клетки также дифференцировались *in vitro* в эритроидные клетки. Протокол был основан на публикации Giarratana et al. ((2011) Blood 120 (15):294 5-53). Вкратце следовали протоколу указанному ниже.

День 0 - день 7: 4×10^4 CD34+ клеток культивировали с плотностью 2×10^4 /мл в среде для дифференцировки (EDM) (модифицированная среда Дульбекко Искова [IMDM], 330 мкг/мл трансферрина, 10 мкг/мл человеческого инсулина, 2 Е/мл гепарина натрия, 5% человеческой АВ+плазмы) в присутствии 10^{-6} М гидрокортизона, 100 нг/мл фактора стволовых клеток (СКФ), 5 нг/мл ИЛ 3 и 3 МЕ/мл эритропоэтина (ЭПО).

День 4: клетки ресуспендируют в свежем EDM, содержащем SCF, ИЛ-3, ЭПО и гидрокортизон.

День 7 - День 11: Клетки ресуспендировали с плотностью $1,5 \times 10^5$ клеток на мл свежего EDM с добавлением SCF и ЭПО.

День 11 - день 21: В день 11 клетки пересаживали в концентрации 1×10^6 /мл в свежем EDM с добавлением ЭПО. Клетки впоследствии были пересажены в эту же среду в концентрации 5×10^6 /мл на 14 день. Рост выходит на плато в течение этого периода времени между 14 и 18 днями, когда жизнеспособность клеток начала падать до прекращения культур на 21 день.

Клетки подсчитывали перед посевом и на протяжении всей дифференциации путем измерения положительности по Акридиновому Оранжевому и исключению йодида пропидия (АОPI) с использованием клеткомера Nexcelom Bioscience K2 с режимом АОPI Erythroid Assay с флуоресцентным каналом 1 (АО), установленным на 700 мс, и флуоресцентным каналом 2 (PI) установленным на 5000 мс.

Процент энуклеированных клеток (клеток без ядра) определяли на 21 день дифференцировки с использованием следующего протокола. Уровень энуклеации был сопоставим среди нетрансфицированных контролей и ZFN-трансфицированных образцов с процентным содержанием 59-63% из этих двух групп.

1. Количество клеток.
2. 100000 клеток, центрифугирование при $450 \times g$, 5 мин, КТ.
3. Ресуспендирование в 50 мкл PBS-BSA+1 мкл GlyA-FITC (DAKO).
4. Окрашивание 15 мин в холодильнике.
5. Добавление 1 мл PBS-BSA, вортекс, центрифугирование.
6. Ресуспендирование в 250 мкл PBS-BSA-NucRed (2 капли NucRed на мл).
7. Получение данных на FACS Canto, используя канал APC для NucRed.
8. Эритроидные клетки с ядром будут находиться в GlyA-положительной NucRed-отрицательной/низкой фракции, а эритробласты будут в двойной GlyA-NucRed-положительной фракции.

Модификацию гена BCL11A измеряли с помощью глубокого секвенирования MiSeq в образцах ДНК, собранных: а) через 48 ч после электропорации, б) в день оттаивания клеток, в то время, когда начиналась дифференцировка *in vitro*, и с) на 14-й день эритроидной дифференциации *in vitro*. Хотя дифференцировка проводилась в течение 21 дня, была выбрана временная точка в 14-й день для анализа ДНК, поскольку она предшествует энуклеации большей части эритроидных клеток, которая приводит к потерям при выделении ДНК. Наблюдаемые проценты модификации на энхансере BCL11A приведены в табл. 4 вместе с подробностями условий трансфекции.

Mdykdhdgdykdhdi dykdddkM~~Ap~~kkkrkvGIHGVPAA~~MA~~ERPFQCRICM~~Q~~NFSDQSNLRAH
 IRHTHTGKPFACDICGRKFARNFSLTMTHTKIHTGSQKPFQCRICM~~Q~~NFSSSTGNLTNHIRTHTGE
 KFACDICGRKFATSGSLTRHTKIHTHPRAPIPKPFQCRICM~~Q~~NFSDQSNLRAHIRHTHTGKPF
 ACDICGRKFAAQCLFHHTKIHLRGSISRARPLNPHPELEEKSELRHKLKYVPHEYIELIEIA
 RNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPI
 GQADEMERIVEENQTRDKHLNPNNEWKVPSSVTEFKFLFVSGHFKGNYKAQLTRLNHIITNCNG
 AVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS (SEQ ID NO:29).

МРНК 65722 (1680 нуклеотидов):

5' gggagacaagcuugaauacaagcuugcuugcuuuuuugcagaagcucagaauaac
 gcucaacuuggcagaucgaaucgccuagagaucuggcggcggagagggcagaggaagcucuc
 uaaccucgugugacgugggagagaaucccgccuaggaaccuaggaacuacaagaccuagacg
 ugauuaaaagaucuagacaucgaaucagaaggaugacgaugacaagauggcccccagaagaag
 aggaagguccgcauucguggguaaccgcccgaugggcugagagggccuucagugucgaucuc
 gcaugcagaaguuugcccgcaaccgaccaccaccaccuacsaagaucacacggcgagaa
 gcccuuccagugucgaucgcaugcagaacucagucagaagggccaccugaucggccacauc
 cgacccacaccggcgagaagccuuugccugugacauuugggaggaauuugcccagaagg
 gcacccugggcgagcauacsaagaucacacgggaucicagaagccuuccagugucgaucug
 caugcagaacucagucggcgccgaccuguccgcccacaucggcaccacaccggcgagaag
 ccuuugccugugacauuugggaggaauuugcccggcgacaaccugcacucccauacca
 agauacaccugcgggaucaccagcuggugaagagcagcugggaggaagaaguccgagcugcg
 gcacaagcugaagucgucggccacgagucacucgagcugagcagcagcaggaacagcacc
 caggaccgcaucuggagagaaggggaguggaucucuaagaagggugacggcuacagggaa
 agcaccugggcggaagcagaagccugacggcgccaucuacagugggcagcccaucgaa
 cggcgugaucguggacacaagggccuacagcgggcgcuacaucugccuacggccagggcgac
 gagaugcagagaucgugaaggaagaccagaccggaaauagcacaucaccccaacgaguggu
 ggaaggguguaaccuagcagcugaccgaguucaaguuccugucgugagcggccacuacgcg
 caacuacaagggccagcugaccagggcugaaccgcaaaacsaucgaauggcggcugcugagc
 guggaggaugcugcugaucggcgcgagaucaaaagcggcaccucgacacuggaggauguc
 gggcgaaguucaacaacggcgagaucaucucgaaucucgagucuaagaagcucgcuuuuc
 cuguccaauuucauuuaagguuccuuuguccuaaguccaacuacuaaacugggggaauua
 ugaagggccuugagcucggaucucgcuauuaaaaaaacuuuuuuuauugcugcugcuga
 agcucgcuuuucguguccaauuucauuuaagguuccuuuguccuuaaguccaacuacuaa
 acugggggaauuaugaagggccuugagcucuggaucugccuauuaaaaaaacuuuuuuuc
 auugcugcgggacauucuu
 aaaaaaaaaaaaaaaaaaaacuag

(SEQ ID NO:30).

Аминокислотная последовательность 65722 (области спирали распознавания подчеркнуты; линкер изображен большими буквами курсивом; гидрофильный пептид маленькими буквами; сигнал локализации нуклеазы показан маленькими буквами курсивом; мутации в остатках остова пальцев 1, 2 и 4 показаны двойным подчеркиванием; мутации домена димеризации (ELD) показаны жирным шрифтом и курсивом; а фосфат-контактная мутация(ии) FokI показана волнистым подчеркиванием):

Mdykdhdgdykdhdi dykdddkM~~Ap~~kkkrkvGIHGVPAA~~MA~~ERPFQCRICM~~Q~~KFARN~~D~~
HRTHTHTKIHTGKPFQCRICM~~Q~~NFSDQSNLRAHIRTHTGKPFACDICGRKFA~~Q~~KGTLGENTKI
 HTGSQKPFQCRICM~~Q~~NFSSSTGNLTNHIRTHTGKPFACDICGRKFARN~~D~~NLHSHTKIHLRGSQ
 VKSELEEKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRK
 PDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRVYV~~K~~ENQTRNKHINPNNEWKVPSSVT
 EFKFLFVSGHFKGNYKAQLTRLN~~R~~KTN~~C~~NGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEI
 NF (SEQ ID NO:31).

Пример 5. Оценка уровня глобина в эритроидном потомстве

Уровни мРНК α -, β - и γ -глобина в клеточной мРНК, выделенной на 14-й день дифференцировки (до того, как общие уровни мРНК резко снижаются в ходе энуклеации и эритроидного созревания) определялись с помощью ОТ-кПЦР для двух препаратов клеток показанных выше в табл. 4. Значения мРНК γ -глобина показаны нормализованными относительно мРНК β -глобина (фиг. 5А) или значений мРНК α -глобина (фиг. 5В) из тех же образцов (с использованием произвольных единиц основанных на соотношении в нетрансфицированном ОТ-ПЦР стандарте определенном как 1).

В обращенно фазовой ВЭЖХ образцов белка, выделенных на 21 день, измеряемый параметр дифференцировки эритроида, использовали для определения того, повышает ли опосредованная ZFN моди-

фикация энхансера BCL11A эритроцитов гемоглобин плода на уровне белка. Определяли отношение гамма-глобина (сумма пиков Агамма и Ггамма) к альфа-глобину, а также отношения гамма-глобина (сумма пиков Агамма и Ггамма) к бета-подобным глобинам (сумме пиков Агамма, Ггамма, бета и дельта-глобина) и они показаны на фиг. 6.

В этом эксперименте приблизительно 3-4-кратное повышение процентного содержания белка глобина плода до уровней около 15-20% наблюдалось в эритроидном потомстве HSPC при 63014/65722-опосредованном разрушении энхансера BCL11A.

Пример 6. Приживление отредактированных клеток у мышей NSG

Отредактированные человеческие CD34+ клетки затем инъецировали мышам NSG для оценки приживления. Степень человеческой химерности (т.е. процент клеток CD45+ человека) измеряли с использованием проточной цитометрии (FACS fluorescence activated cell sorting) в периферической крови, отобранной через 8, и 12, 16 и 20 недель после трансплантации, и в костном мозге, отобранном через 12 недель и 20 недель. Кроме того, для проверки уровня приживления ZFN-модифицированных клеток уровень разрушения гена в локусе энхансера BCL11A оценивали путем прямого высокопроизводительного секвенирования локуса-мишени ZFN и сравнивали с уровнями модификации гена-мишени, измеренными в исходном материале.

HSPC от двух здоровых доноров (называемых PB-MR-003 и PB-MR-004) были мобилизованы с помощью G-SCF и Plerixafor и очищены, как описано в Yannaki et al. ((2012) Mol Ther 20(1):230-8. doi:10.1038/mt.2011). Деплецию тромбоцитов проводили на продукте лейкафереза с использованием устройства Fresenius-Kabi Lovo до его обогащения клетками CD34+ с использованием инструмента Miltenyi Biotec CliniMACS Plus. Затем очищенные клетки высевали в культуру для трансфекции.

Через два дня после очистки клеток CD34+ клетки подвергали электропорации с использованием инструмента Maxcyte в присутствии 120 мкг/мл одной мРНК, кодирующей родительскую пару ZFN, 63014/65722, или оптимизированных количеств двух отдельных мРНК, кодирующих оптимизированную пару ZFN 80мкг/мл 63014 и 20 мкг/мл 65722. Перед трансфекцией аликвоту клеток откладывали в качестве нетрансфицированного контроля. 95 миллионов клеток были трансфицированы из PB-MR-003, и 120 миллионов клеток были трансфицированы из PB-MR-004.

После электропорации проводили временное культивирование в течение ночи при 30°C, и затем клетки культивировали в течение дополнительных 24 ч при 37°C. Через два дня после электропорации отбирали аликвоты клеток для анализа ДНК, а оставшиеся клетки собирали, криоконсервировали и хранили в жидком азоте.

Кондиционирование:

Мышей обрабатывали 10 мг/кг/день воды с байтрилом (Baytril) за 1-2 дня до облучения и облучали сублетально 300 RAD за 16-24 ч до трансплантации. Трансплантация проводилась путем инъекции в хвостовую вену (см. ниже). Затем мыши получали свежую воду с байтрилом. Воду с байтрилом заменили через неделю, и добавление в воду байтрила было прекращено через 14 дней после трансплантации.

Трансплантация:

В день трансплантации предварительно разогрейте X-Vivo 10/1% PSG+коктейль из 3-х цитокинов (рекомбинантный фактор стволовых клеток человека (SCF), рекомбинантный тромбопоэтин человека (TPO) и рекомбинантный лиганд Flt-3 человека (Flt-3L)) при 37°C, приготовьте свежий PBS/0,1% BSA при температуре окружающей среды (стерильный/отфильтрованный). Криоконсервированные клетки оттаивали при 37°C, осаждали, ресуспендировали в предварительно нагретой среде X-Vivo, снова осаждали, ресуспендировали в PBS/0,1% BSA и подсчитывали. После другого осаждения осадок клеток ресуспендировали в 550 мкл на мышью PBS/0,1% BSA (2×10^6 клеток/мл в расчете на количество клеток). Затем клетки вводили при комнатной температуре в хвостовую вену мыши с помощью иглы 25 калибра. Группы исследования показаны ниже в табл. 5.

Таблица 5. Группы дозирования для приживления отредактированных клеток hCD34+

Виды: Мышь				Пол, Возраст, Количество: 60 самок мышей NSG			
Группа №	(Кол-во/группу)	Исследуемый препарат	% индекс лов (Сутки 2)	Жизнеспособность через 1d после оттаивания	Доза (клеток /мышь)	Кол-во/ Умерщвление	
						12-я неделя	20-я неделя

1	10 (M1-10)	Донорские клетки PB-MR-003, обработанные мРНК 63014 и 65722.	79%	83%	1 МИЛЛИОН	5	5
3	10 (M21-30)	PB-MR-003, нетрансфицированные	0,1%	95%	1 МИЛЛИОН	5	5
4	10 (M31-40)	Донорские клетки PB-MR-004 обработанные мРНК 63014 и 65722.	75%	77%	1 МИЛЛИОН	5	5
6	10 (M51-60)	PB-MR-004, нетрансфицированные	0,1%	92%	1 МИЛЛИОН	5	5

Животных наблюдали ежедневно для оценки общего состояния здоровья и взвешивали ежедневно в течение первых 2 недель, а затем взвешивали каждые две недели. Периферическую кровь собирали из поднижнечелюстной вены (100 мкл) через 8, 12, 16 и 20 недель после трансплантации или через пункцию сердца (1 мл) для умерщвленных животных через 12 и 20 недель после трансплантации. Половину животных в каждой группе (5 мышей в группе) подвергали эвтаназии через 12 недель после трансплантации, а костный мозг и терминальную кровь собирали для анализа. Оставшихся животных в каждой группе (5 мышей на группу) умерщвляли через 20 недель после трансплантации.

Забор крови, сбор и обработка клеток: периферическую кровь отбирали через поднижнечелюстную вену или пункцию сердца в пробирки с ЭДТА и центрифугировали при 500×g в течение 5 мин для удаления плазмы. После промывки фосфатно-солевым буфером (PBS) с бычьим сывороточным альбумином (BSA) и центрифугирования к осадку добавляли 10-кратный объем гемолитического буфера, и смесь инкубировали при 37°C в течение 15 мин, центрифугировали и снова промывали. Осажденную фракцию восстанавливали в 1 мл PBS BSA; аликвоту удаляли и центрифугировали при 1000×g в течение 5 мин, а полученный осадок сохраняли для генотипирования. Фракцию супернатанта использовали для FACS анализов.

Костный мозг, бедренную кость, большеберцовую кость и тазовые кости собирали в модифицированную среду Дульбекко Искова (IMDM), содержащую фетальную сыворотку теленка (FCS); весь костный мозг промывали в растворе PBS BSA и фильтровали с использованием нейлонового сита с размером отверстий 70 мкм. Объем доводили до 10 мл с помощью PBS BSA, и аликвоту использовали для подсчета клеток (Cellometer).

Активность ZFN анализировали с использованием глубокого секвенирования MiSeq. Вкратце, геномную ДНК от мышей, которым инъецировали либо нетрансфицированные контрольные CD34+ HSPC, либо CD34+ HSPC, трансфицированные энхансер-направленной мРНК ZFN, выделяли из образцов крови, полученных через 8 недель и 12 недель, или из костного мозга через 12 недель после инъекции. Область представляющую интерес (содержащую сайт связывания ZFN в локусе BCL11A) амплифицировали с помощью ПЦР и определяли уровень модификации путем глубокого секвенирования с парными концами на платформе Illumina (MiSeq).

Для создания библиотек, совместимых с платформой секвенирования Illumina MiSeq, адаптеры, штрих-коды и связующее для проточной ячейки (короткая последовательность ДНК) были прикреплены к целевым специфическим ампликонам с использованием двух наборов праймеров для слияния в последовательных ПЦР. Для MiSeq оценки модификации энхансера BCL11A человека в образцах крови и костного мозга мыши протокол должен быть скорректирован из-за низких количеств целевой ДНК в этих образцах.

Для адаптерного MiSeq ПЦР использовали следующие праймеры: PRJIYLFN-f2:

ACA CGA CGC TCT TCC GAT CTN NNN AGT CCT CTT CTA CCC CAC CCA

(SEQ ID NO:32) и

PRJIYLFN-r4:

GAC GTG TGC TCT TCC GAT CTC TAC TCT TAG ACA TAA CAC ACC AGG

G (SEQ ID NO:33).

Для анализа ДНК из образцов костного мозга мыши выделяли с помощью DNeasy и в каждой реакции ПЦР использовали приблизительно 100 нг ДНК. ДНК из образцов крови мыши выделяли с помощью Tissue XS, и в каждой реакции использовали 10 мкл 15 мкл выделенной ДНК. В дополнение к ДНК к каждой реакции MiSeq ПЦР добавляли следующее: 25 мкл смеси HotStar Taq (Qiagen), 0,5 мкл каждого из

перечисленных выше праймеров энхансера BCL11A (в концентрации 100 нМ) и воду до 50 мкл общего объема реакции. Типичными условиями для MiSeq ПЦР были: денатурация при 95°C в течение 15 мин и 30 циклов при 94°C в течение 30 с, 62°C в течение 30 с и 72°C в течение 40 с с последующим удлинением в течение 10 мин при 72°C. После MiSeq ПЦР продукт ПЦР разводили водой между 1:50 и 1:200 или оставляли неразбавленным для образцов с очень низким исходным количеством клеток. Штрих-код ПЦР была выполнена с 1 мкл продукта MiSeq ПЦР, разведенного, как описано выше, 25 мкл смеси HotStar Taq, 1 мкл прямого штрих-код праймера, 1 мкл обратного штрих-код праймера (оба в концентрации 10 нМ) и воды до 50 мкл общего объема реакции. Условия штрих-код ПЦР были следующими: денатурация при 95°C в течение 15 мин и 18 циклов при 94°C в течение 30 с, 60°C в течение 30 с и 72°C в течение 30 с с последующим удлинением на протяжении 10 мин при 72°C. Продукты штрих-код ПЦР объединяли и сиквенировали на сиквенаторе Illumina MiSeq. Результаты показаны в табл. 5 выше.

FACS анализ на химеризм и определение клеточной линии. Для оценки степени человеческого химеризма фракцию клеток периферической крови (через 8, 12, 16 и 20 недель после приживления) и костного мозга (через 12 и 20 недель после приживления) окрашивали hCD45-APC Cy7 (Biolegend) и антителами к hCD45-BV510 (BD Biosciences) соответственно и проводили FACS анализ. Кроме того, проводили анализ гематopoэтических линий путем окрашивания клеток костного мозга специфическими антителами, описанными в табл. 6 ниже:

Таблица 6. Источники антител для клеточных маркеров

Антитело	Ссылки
CD3-FITC: BD 561807 (клон UCNT1)	BD 561807
CD19-PE: BD 340364 (клон SJ25C1)	BD 340364
CD45-BV510	BD 563204
Lin-APC (CD3/UCNT1, CD14/hCD14, CD16/3G8, CD19/hIB19, CD20/2H7, CD56/hCD56)	BIOLEGEND 348803
CD33-PE-CF594	BD 562492
GlyA-FITC	DAKO 0870
CD38-PerCP Cy5. 5	BD 551400
CD14-PE	BD 555398
CD34-PE Cy7	BD 560710
CD71-APC Cy7/h7	BD 563671
CD15-BV650	BD 564232
CD8-PerCP Cy5. 5	BD 341051
CD4-PE Cy7	BD 344612
CD56-APC	BD 318310
IgM-APC Cy7	BD 314520
CD20-BV650	BD 563780

Кроме того, для очистки и сортировки популяций HSPC мы использовали стратегию обогащения/деплеции с использованием магнитного разделения клеток (MACS). Клетки костного мозга сначала окрашивали CD19-биотином, CD3-биотином, B220-биотином, TER119-биотином и m-ckit-биотином (BD Biosciences), а затем инкубировали с анти-биотин гранулами (Miltenyi Biotec).

Положительная фракция и обедненная фракция были разделены с использованием колонок LS (Miltenyi Biotec), помещенных в магнитное поле MACS. После разделения положительную фракцию окрашивали стрептавидин-APC, CD3-FITC, CD19-PE, CD45-BV510 (BD Biosciences) и деплетированную фракцию с помощью CD34-FITC (BD Biosciences), Gly-A-PE (DAKO), CD19-APC (BD), Lin-APC (Biolegend), стрептавидин-APC, CD45-BV510, CD33-PE-CF594 (BD) и CD38-PECy-7 (Biolegend).

Нетрансфицированные HSPC и 63014/65722 трансфицированные HSPC были привиты мышам NSG с использованием стандартных процедур, как описано выше. Степень человеческого химеризма у этих мышей после приживления оценивали путем измерения доли hCD45-положительных клеток с использованием FACS.

На фиг. 7 изображен процент клеток CD45+ человека в периферической крови, собранных через 8 и 12 недель после трансплантации, а на фиг. 8 изображены проценты клеток костного мозга, собранных на 12 неделе. Как показано, уровни приживления в этом исследовании были сопоставимы с человеческим химеризмом после приживления нетрансфицированного контроля и 63014/65722 трансфицированных HSPC. Только 3 мыши из 60, распределенных по группам, не имели клеток CD45+, что указывает на отсутствие приживления.

Восстановление различных гематopoэтических клеточных линий проверяли с помощью FACS-анализа клеток костного мозга, полученных на 12-й неделе, с помощью антител, распознающих маркеры клеточной поверхности, специфичные для линий, с использованием стандартных процедур. Как показано на фиг. 9, наблюдали сравнимое представление всех проанализированных гематopoэтических линий человека в костном мозге на 12-й неделе после инъекции между потомством CD34+ клеток обработанных

BCL11A-специфичной мРНК, кодирующей ZFN, и потомством нетрансфицированных клеток. Костный мозг мышей, умерщвленных на 12-й неделе после приживления, выделяли, и распределение различных гематопэтических линий было проанализировано с помощью FACS с использованием антител, распознающих указанные маркеры линий. Все числа приведены в виде отношения клеток, окрашенных положительно по указанному маркеру линии, к проценту человеческих CD45-положительных клеток, за исключением клеток, экспрессирующих эритроидный маркер Cd71+ (Ter119) на фиг. 9С, которые представлены как процент положительно окрашенных клеток во всей популяции, поскольку эритроидные клетки не являются CD45-положительными.

Уровни модификации гена в эритроидном энхансере BCL11A (% аллелей со вставками и делециями [инделами]) оценивали путем глубокого секвенирования области-мишени ZFN с использованием платформы секвенирования MiSeq, как описано выше. Данные показаны на фиг. 10 для образцов крови 8-й недели и 12-й недели, а также на фиг. 11 для образцов костного мозга 12-й недели и отсортированных линий, полученных из образцов клеток костного мозга 12-й недели клеток обработанных 63014/65722. Для сравнения, процентные доли инделов, измеренные через 2 дня после трансфекции (как указано в табл. 5), также показаны на графиках фиг. 10 и 11.

Кроме того, для обоих наборов доноров HSPC, обработанных 63014/65722, было обнаружено сохранение генной модификации в энхансере BCL11A в различные моменты времени и в различных линиях. Сопоставимая модификация наблюдалась как в BCL11A-зависимых (В-клетках, 'CD19'; примитивных предшественниках, 'CD38H'), так и в BCL11A-независимых (миелоидных 'CD33') линиях. Хотя уровни вводимой генной модификации были выше в образце донора PB-MR-003, чем в образце донора PB-MR-004, клетки полученные из PB-MR-004 последовательно демонстрируют более высокие уровни модификации, т.е. лучшее сохранение модификации у мышей, чем те, которые получены из PB-MR-003.

В целом, наблюдаемое сохранение генной модификации у энхансера BCL11A у мышей соответствовало тому, что наблюдалось в предыдущих экспериментах на мышах с использованием ряда ZFN, нацеленных на различные генные мишени.

Кроме того, поскольку эритроидные предшественники человека не способны дифференцироваться у мышей, чтобы определить количество, нацеленных на BCL11A, генных модификаций, которые имели место в этих клетках, клетки костного мозга были изъяты из мышей и дифференцированы *in vitro*. В этих экспериментах человеческие клетки, полученные из костного мозга, были изъяты из умерщвленных мышей на 12 неделе после приживления и дифференцированы *in vitro*, как описано выше. Модификацию гена-мишени BCL11A измеряли с помощью высокопроизводительного Miseq секвенирования ДНК, выделенной из клеток, на 14-й день дифференцировки.

Данные модификации (инделами) представлены на фиг. 12, которая показывает уровни модификации на 14-й день эритроидной дифференцировки. Проценты инделов на 14-й день дифференцировки *in vitro* заметно варьируют для каждой культуры, полученной из клеток, выделенных от одной мыши, что отражает олигоклеточную природу экспансии, полученной в этих условиях. Эти данные показывают, что модификация энхансера BCL11A, опосредованная ZFN 63014/65722, не была заметно изменена во время эритроидной дифференцировки. Как наблюдалось в образцах крови и костного мозга, образцы эритроидного потомства клеток, полученных из PB-MR-004, показали более высокие средние уровни модификации, чем эритроидное потомство клеток, полученных из PB-MR-003.

Относительные уровни различных мРНК глобина определяли с помощью ОТ-ПЦР-анализа РНК, выделенной из клеток, на 14-й день эритроидной дифференцировки *in vitro*, и данные представлены на фиг. 13А, на которой относительные отношения мРНК γ -глобина к мРНК β -глобина и мРНК γ -глобина к мРНК α -глобина (фиг. 13В) усреднены для 5 эритроидных культур из каждой группы. Как в нетрансфицированных, так и в обработанных 63014/65722 образцах отношения мРНК γ -глобина к мРНК β -глобина и мРНК γ -глобина к мРНК α -глобина сильно различаются между эритроидными потомками отдельных мышей из той же группы. Культуры, полученные из донора PB-MR-004, демонстрируют в среднем более низкие отношения γ -глобина, чем культуры, полученные из донора PB-MR-003, что согласуется с лучшим созреванием, которое наблюдается для образцов, полученных из PB-MR-004. Однако, несмотря на эту вариабельность, средние значения для образцов, обработанных ZFN, показывают ~1,5-2-кратное увеличение уровней мРНК γ -глобина по сравнению с их соответствующими нетрансфицированными аналогами.

Уровни белка глобина оценивали анализом ВЭЖХ. Фиг. 14 демонстрирует анализ белков глобина в образцах, отобранных на 16-й день дифференцировки. Определяли отношение гамма-глобина (сумма пиков Агамма и Ггамма) к альфа-глобину, а также отношения гамма-глобина (сумма пиков Агамма и Ггамма) к бета-подобным глобинам (сумме пиков Агамма, Ггамма, бета и дельта-глобина) и среднее для каждой группы показано над каждым столбцом. Соглашаясь с плохой эритроидной дифференцировкой образцов, полученных из PB-MR-003, уровни гамма-глобина в нетрансфицированных клетках, полученных от этого донора, были очень высокими (~30%), и поэтому обработка ZFN привела только к 1,2-кратному увеличению уровней гамма-глобина. PB-MR-004 демонстрировал более типичные нетрансфицированные уровни (~9%) и показывал ~2-кратное увеличение уровней белка гамма-глобина после 12-

недельной циркуляции через тело мыши.

Считается, что пациенты, которые имеют >8,6% γ -глобина естественным образом, имеют преимущество по сравнению с пациентами с уровнем γ -глобина <8,6% (Platt et al. (1994) *N Engl J Med*, 330:1639-44). Фактически, достижение химерного процента не серповидных эритроцитов на уровне 10-20% посредством приживания отредактированных клеток может привести к клиническому улучшению (Chang et al. (2017) *Mol Ther Methods Clin Dev* 4:137-148. doi:10.1016/j.omtm. 2016.12.009). Таким образом, несмотря на необходимость прохождения процесса эритроидной дифференцировки *in vitro*, процент химерных клеток и уровень детектируемого белка γ -глобина являются индикаторами терапевтической эффективности.

Все патенты, патентные заявки и публикации, упомянутые в данном документе, включены в него во всей их полноте посредством ссылки.

Хотя раскрытие было предложено с некоторыми деталями в качестве иллюстрации и примера для целей ясности понимания, для специалистов в данной области техники будет очевидно, что различные изменения и модификации могут быть осуществлены на практике, не отступая от сущности или объема данного раскрытия. Соответственно, вышеприведенные описания и примеры не должны рассматриваться как ограничивающие.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Пара нуклеаз цинкового пальца (ZFN), содержащая левую и правую ZFN, которые димеризуются в виде пары, где левая ZFN содержит аминокислотную последовательность, показанную в SEQ ID NO: 29, и правая ZFN содержит аминокислотную последовательность, показанную в SEQ ID NO: 31.
2. Полинуклеотид, кодирующий пару нуклеаз цинкового пальца по п.1.
3. Полинуклеотид по п.2, где указанный полинуклеотид представляет собой мРНК.
4. Полинуклеотид по п.3, содержащий SEQ ID NO: 28 и SEQ ID NO: 30.
5. Клетка, содержащая нуклеазы цинкового пальца по п.1 или полинуклеотид по любому из пп.2-4.
6. Клетка по п.5, где указанная клетка представляет собой стволовую клетку или клетку-предшественник.
7. Клетка по п.6, где указанная клетка представляет собой клетку человека.
8. Клетка по любому из пп.5-7, где геном указанной клетки модифицирован нуклеазами цинкового пальца в области +58 энхансерной последовательности BCL11A.
9. Клетка по п.8, где указанная геномная модификация выбрана из группы, состоящей из вставок, делеций и их комбинаций.
10. Клетка, происходящая от клетки по любому из пп.5-9, где указанная клетка получена культивированием клетки по любому из пп.5-9.
11. Клеточная линия, происходящая от клетки по любому из пп.5-9, где указанная клеточная линия получена культивированием клетки по любому из пп.5-9.
12. Частично или полностью дифференцированная клетка, происходящая от клетки по любому из пп.5-10 или клеточной линии по п.11.
13. Клетка по любому из пп.5-10, где указанная клетка проявляет повышенную экспрессию гамма и/или бета-глобина, по сравнению с клеткой без указанной геномной модификации.
14. Фармацевтическая композиция, содержащая пару нуклеаз цинкового пальца по п.1, полинуклеотид по любому из пп.2-4 или клетку по любому из пп.5-10.
15. Способ модификации эндогенной последовательности энхансера BCL11 в клетке, причем способ включает введение нуклеазы цинкового пальца по п.1 или полинуклеотида по любому из пп.2-4 в указанную клетку таким образом, что эндогенная последовательность энхансера BCL11A модифицируется.
16. Способ по п.15, дополнительно включающий введение экзогенной последовательности в указанную клетку таким образом, что указанная экзогенная последовательность вставляется в эндогенную последовательность энхансера BCL11A.
17. Способ по п.15, где указанная модификация содержит делецию.
18. Способ увеличения выработки глобина у субъекта, включающий введение указанному субъекту клетки по любому из пп.5-10.
19. Способ по п.18, где указанным субъектом является человек, а клетка представляет собой стволовую клетку человека или клетку-предшественник человека.
20. Способ по п.19, где указанная клетка вводится пациенту и клетка приживается, дифференцируется и созревает в указанном субъекте.
21. Способ по любому из пп.18-20, где у указанного субъекта имеется гемоглобинопатия.
22. Способ по п.21, где указанная гемоглобинопатия представляет собой бета-талассемию или серповидноклеточную анемию.
23. Способ получения генетически модифицированной клетки, содержащей геномную модификацию в эндогенной энхансерной последовательности BCL11A, причем указанный способ включает ста-

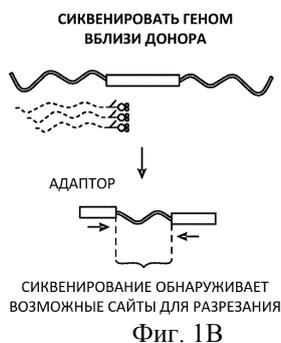
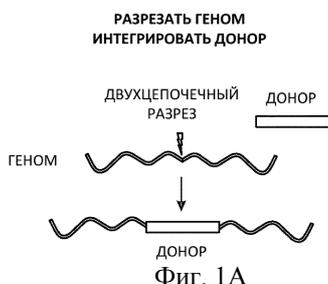
дии:

а) приведение клетки в контакт с полинуклеотидом по любому из пп.2-4;
 б) подвержение клетки условиям, способствующим экспрессии слитого белка из полинуклеотида; а также

в) модифицирование эндогенной энхансерной последовательности BCL11A с помощью экспрессированного слитого белка, достаточное для получения генетически модифицированной клетки.

24. Способ по п.23, дополнительно включающий стимуляцию указанной клетки по меньшей мере одним цитокином.

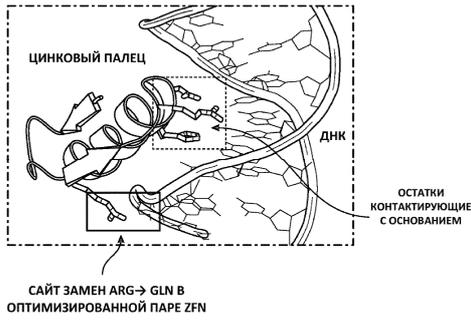
25. Набор, содержащий нуклеазу цинкового пальца по п.1, полинуклеотид по любому из пп.2-4 и/или клетку по любому из пп.5-10.



	ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ	ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ	ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ
BCL11A →	2	60495266	1064	22	20508058	15	17	33940450	9
	8	119866442	939	1	167837598	14	18	206783334	9
	2	82164814	156	1	211855462	14	19	850496	9
	10	132654828	105	9	129811194	14	20	29445862	9
	14	67422056	102	16	46399310	14	1	51299472	8
	4	7829724	93	19	38685086	14	1	201507258	8
	2	23702832	88	14	50504	14	2	16715068	8
	14	73740820	85	14	50504	14	2	20398316	8
	1	29182166	82	2	96873178	13	2	36785314	8
	10	69571176	75	2	107763998	13	2	226916064	8
	16	15141540	68	2	127820122	13	2	237467876	8
	8	94989042	66	6	1595312	13	6	49519624	8
	16	2122340	66	6	35468034	13	6	160655866	8
	19	45406880	64	6	144064118	13	10	26734304	8
	16	67184172	59	8	107386748	13	10	79114636	8
	16	46394744	58	10	98910092	13	12	50505884	8
	13	75949350	57	12	104048816	13	14	73746866	8
	16	34593846	57	14	89468992	13	17	43805774	8
	16	46390726	54	1	21635648	12	17	82718744	8
	1	54506814	53	2	135453242	12	18	20575274	8
	19	55115786	53	6	7999996	12	20	5789510	8
	6	53905856	52	8	13105963	12	21	25237778	8
	17	29524234	52	9	128604032	12	22	19863370	8
	1	26938300	48	11	2992880	12	X	47067946	8
	5	145864342	40	13	91576712	12	1	19916880	7
	20	37107466	39	14	18295748	12	1	143187676	7
	21	33308800	36	16	34587146	12	1	148402426	7
	1	204975770	35	21	25807680	12	1	223224616	7
	5	132776212	35	22	15226546	12	2	119198990	7
	11	33694468	34	1	70765014	11	3	45419400	7
	10	101133376	33	1	120942294	11	4	38364694	7
	17	28854518	33	1	200702950	11	4	99547294	7
	X	30280458	31	3	19942238	11	5	38056212	7
	2	84459580	30	5	2057482	11	6	6715482	7
	7	131503656	30	6	3196959	11	6	15089744	7
	2	112583534	27	9	98755452	11	6	89830096	7
	3	49724756	27	11	30406776	11	6	89888012	7
	12	55928112	27	1	177943758	10	7	113383212	7

ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ	ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ	ХРОМОСОМА	№ ОСНОВАНИЯ	КОЛ-ВО ИНТЕГРАНТОВ
8	66165666	26	2	218419648	10	8	37700430	7
11	119316888	26	6	136793200	10	8	88289848	7
X	134914730	26	8	25969948	10	9	101650500	7
6	9529212	25	10	4849356	10	9	112957352	7
12	12495298	25	10	47957704	10	10	42097318	7
16	10992070	25	10	52321784	10	10	91591818	7
3	47439806	25	10	52321864	10	10	120431936	7
11	70518294	25	11	102665874	10	11	5225468	7
16	10839466	22	12	54411388	10	12	155086	7
X	66004390	22	12	76936636	10	12	51951306	7
1	9267210	21	15	19782342	10	15	64365676	7
9	91430052	21	16	15220070	10	15	65185434	7
16	30692064	21	16	15220090	10	16	34573028	7
1	15579898	20	16	18131698	10	16	34573058	7
2	173018750	20	16	18131716	10	16	46401528	7
16	88559110	20	18	5447734	10	16	81738062	7
14	149709484	19	19	43779827	10	17	836804	7
8	124924782	19	19	51327816	10	17	79814568	7
13	112237576	19	21	3319396	10	19	38478442	7
20	44683654	19	1	28569750	9	19	38478442	7
1	228965770	18	1	57734650	9	19	40578628	7
7	13991304	18	1	115989958	9	19	41165952	7
7	10110844	18	10	148636460	9	20	5858134	7
11	108224088	18	3	51942092	9	22	8638580	7
12	89874464	18	3	149936952	9	22	36423722	7
1	208159440	17	4	84425848	9	1	2914438	6
5	150115792	17	4	138895946	9	1	91089046	6
12	70981936	16	6	3737388	9	1	115328550	6
15	2818428	16	6	16437886	9	1	12516868	6
19	1279936	16	9	93003938	9	1	12516952	6
20	48636010	16	9	104230026	9	1	155610614	6
1	165379386	15	10	67764450	9	1	155746194	6
3	71103574	15	10	69789200	9	1	20164194	6
5	103961570	15	13	98822320	9	2	131443890	6
7	128077200	15	15	31382604	9	3	4323698	6
8	29495584	15	15	34422898	9	3	58414862	6
16	46399748	15	15	34569124	9	3	68129874	6
16	57534378	15	15	78243618	9	3	87800708	6
3	141228736	6	13	79775976	5	9	92277906	4
4	151031930	6	15	40114388	5	9	97934932	4
5	178121148	6	15	68307028	5	9	121568378	4
6	28931180	6	15	82539916	5	9	124392210	4
6	108893514	6	15	86579492	5	9	126591162	4
6	136088656	6	16	21646824	5	9	138212048	4
6	163725808	6	16	53042972	5	10	25357460	4
7	162655	6	16	59873980	5	10	41847448	4
7	35016804	6	16	67841642	5	10	41847304	4
8	30408876	6	16	78808108	5	10	42103514	4
8	91876290	6	17	17025784	5	10	64508254	4
9	119945686	6	17	26884886	5	10	78825618	4
9	4030722	6	17	72902448	5	10	109893420	4
9	104988250	6	18	1146582	5	11	35844902	4
10	72293352	6	18	33538826	5	11	72431102	4
10	79866034	6	19	7158860	5	12	6232920	4
11	18695014	6	16	11923082	5	12	7242432	4
11	121093264	6	18	18492114	5	12	49404046	4
12	3073480	6	19	45886048	5	12	75511410	4
12	26186106	6	20	29881056	5	12	88031364	4
13	43878548	6	20	43558622	5	12	95073610	4
14	41294346	6	20	47571112	5	12	11315332	4
15	58555840	6	22	19545504	5	13	18964476	4
15	60478340	6	22	23284226	5	13	34779902	4
16	3108742	6	14	30894	5	13	113000728	4
16	4518046	6	1	6369676	4	14	92516966	4
16	29292606	6	1	18556502	4	15	33229354	4
16	77989246	6	1	24862574	4	15	64142544	4
16	88266652	6	1	25442520	4	15	67498752	4
16	88732086	6	1	30801948	4	15	75565386	4
17	96838710	6	1	51631166	4	15	78355710	4
19	267174	6	1	72774624	4	15	684294	4
20	39690408	6	1	109213944	4	15	14285910	4
X	40629108	6	1	15360852	4	15	8531250	4
X	40835932	6	1	175878760	4	17	9710280	4
1	1724636	5	1	201154009	4	17	21372082	4
1	6161778	5	1	224887886	4	17	21906776	4
1	17419988	5	2	13774808	4	17	26603948	4
1	36401004	5	2	166910124	4	17	26603992	4
1	143208542	5	2	234668446	4	17	40557500	4
1	150528982	5	3	49802584	4	17	43796376	4
1	154750274	5	3	195856500	4	17	49883740	4
2	86563402	5	3	198148214	4	18	13272890	4
2	10503652	5	4	1886620	4	18	30136846	4
2	117954094	5	4	3480800	4	18	30845084	4
2	130230372	5	4	99869180	4	18	74019902	4
2	176188548	5	4	138090616	4	19	40266218	4
2	2041446	5	4	15887704	4	19	42741976	4
3	16073022	5	5	1072038	4	19	42787202	4
3	75669968	5	5	36318878	4	19	42881190	4
3	178988880	5	5	36322678	4	19	49683046	4
3	184688238	5	5	6562414	4	20	4673782	4
4	49119414	5	5	100709408	4	20	24683668	4
4	73573948	5	5	142193948	4	20	44911634	4
4	78700116	5	5	149564992	4	20	48621130	4
5	7057496	5	5	164896712	4	20	49279476	4
5	11767140	5	5	176740968	4	20	49699412	4
5	55314802	5	6	2643624	4	20	56914278	4
5	13214286	5	6	2732116	4	20	52819166	4
6	58116968	5	6	45166460	4	21	10442640	4
6	60870558	5	6	146750040	4	22	22533154	4
7	128792974	5	6	169090710	4	22	22848222	4
7	133081754	5	7	1214740	4	22	22956152	4
7	92835886	5	7	37357350	4	22	23111094	4
9	131373156	5	7	37357420	4	22	23162790	4
10	50701042	5	7	62358868	4	22	24629320	4
10	87382350	5	7	17679460	4	22	32136074	4
11	49652236	5	7	10678616	4	22	45716526	4
11	61612488	5	7	133077880	4	X	61348666	4
11	62795536	5	8	5753270	4	X	39688532	4
11	67277754	5	8	5753290	4	X	52934730	4
11	82032970	5	8	8244584	4	X	71953192	4
11	86939378	5	8	8371950	4	X	01323678	4
12	6326346	5	8	101955394	4	X	106863278	4
12	104138422	5	8	119000446	4	Y	3900438	4
12	121500124	5	9	38068006	4			

Фиг. 2



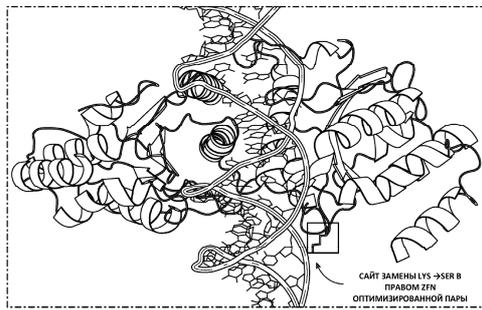
Фиг. 3А

САЙТ ЗАМЕН ARG GLN В ОПТИМИЗИРОВАННОЙ ПАРЕ ZFN

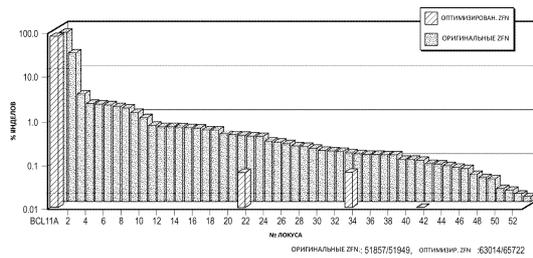
ОСТАТКИ КОНТАКТИРУЮЩИЕ С ОСНОВАНИЕМ

ЛЕВЫЙ ZFN	F1	VPAAMAERP	FQCRICMNF	S	DCSNLRARHIRTH	
	F2	TGEKP	FACDICGRKF	A	RNFSLTMTHTKI	
	F3	TGSQKP	FQCRICMNF	S	STGNLTHIRTH	
	F4	TGEKP	FACDICGRKF	A	TSGSLTRHTKI	
	F5	THPRAPRPK	FQCRICMNF	S	DCSNLRARHIRTH	
	F6	TGEKP	FACDICGRKF	A	AQCQLFHTKI	LRG
ПРАВЫЙ ZFN	F1	VPAAMAERP	FQCRICMNF	A	RNDHRTHTHTKI	
	F2	TGEKP	FQCRICMNF	S	QKALIRHIRTH	
	F3	TGEKP	FACDICGRKF	A	QKGTLGHTKI	
	F4	TGSQKP	FQCRICMNF	S	RGDDLSPHIRTH	
	F5	TGEKP	FACDICGRKF	A	RRDNLHSHTKI	LRG

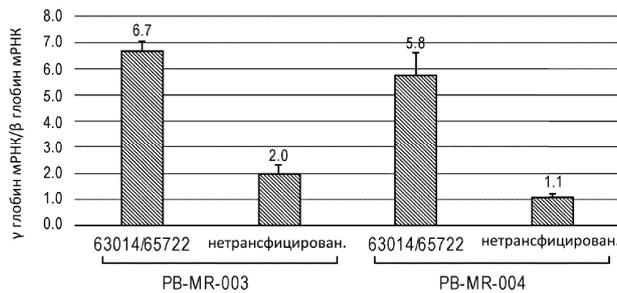
Фиг. 3В



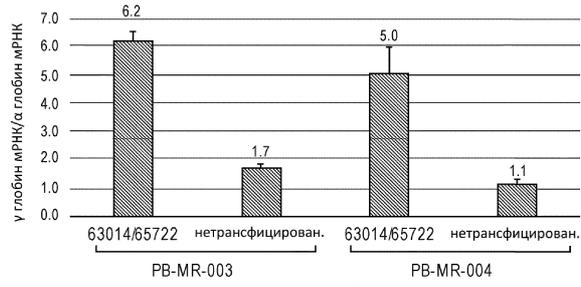
Фиг. 3С



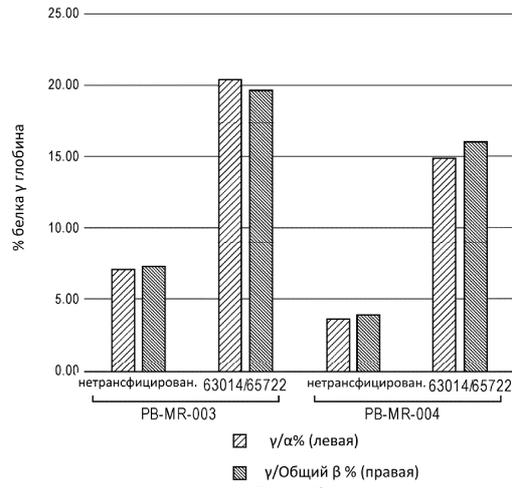
Фиг. 4



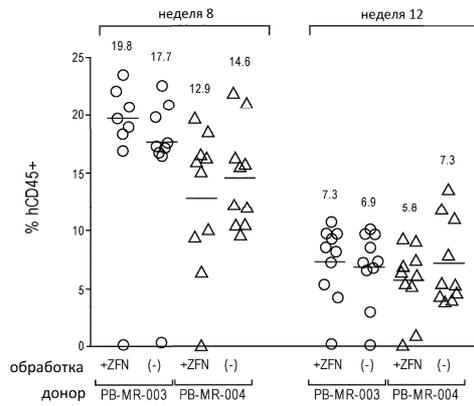
Фиг. 5А



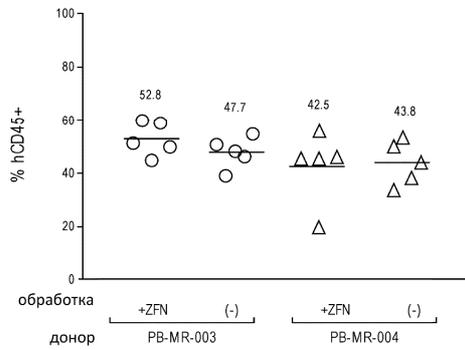
Фиг. 5В



Фиг. 6

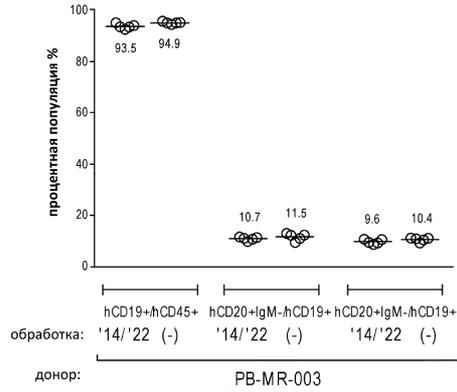


Фиг. 7



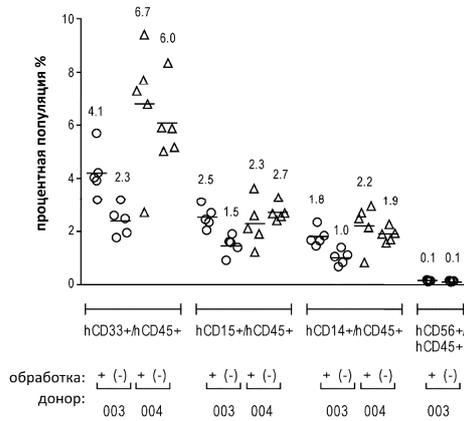
Фиг. 8

Лимфоидные (В клетки)



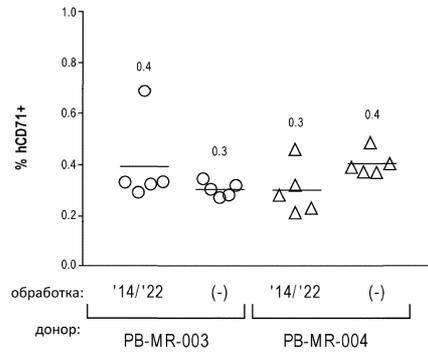
Фиг. 9А

Миелоидные



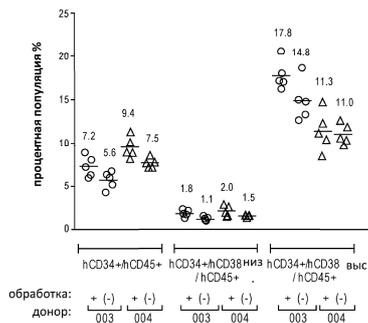
Фиг. 9В

Эритроидные

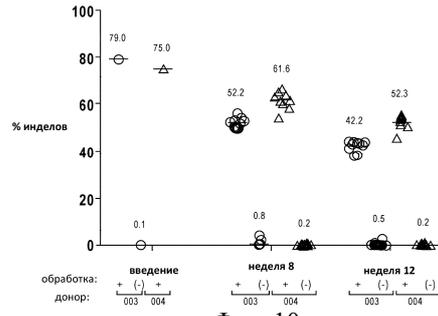


Фиг. 9С

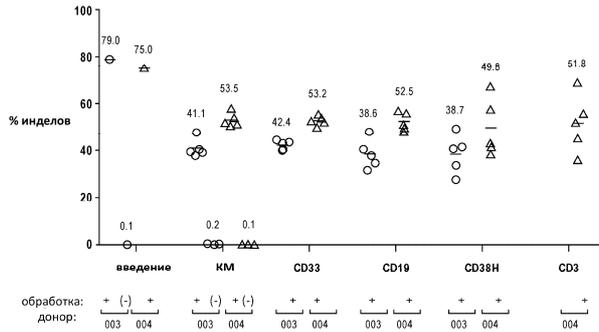
hSPC



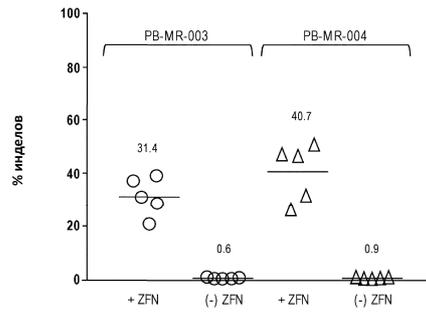
Фиг. 9D



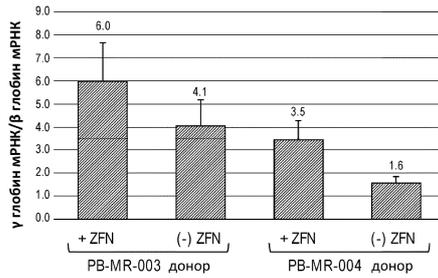
Фиг. 10



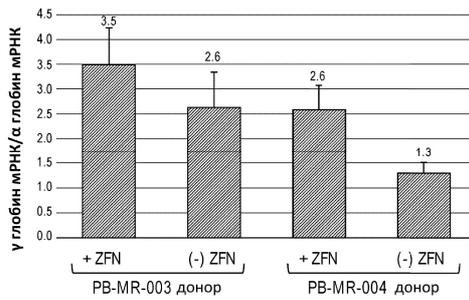
Фиг. 11



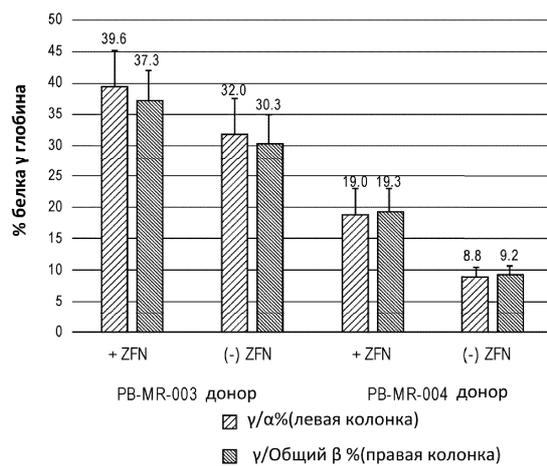
Фиг. 12



Фиг. 13А



Фиг. 13В



Фиг. 14

