

(19)



**Евразийское
патентное
ведомство**

(11) **045196**

(13) **B1**

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

(45) Дата публикации и выдачи патента

2023.10.31

(21) Номер заявки

201891212

(22) Дата подачи заявки

2016.12.15

(51) Int. Cl. **C07K 14/74** (2006.01)

C07K 14/725 (2006.01)

C12N 5/0783 (2010.01)

C12N 15/09 (2006.01)

C12N 5/10 (2006.01)

C12N 5/12 (2006.01)

(54) **АДРЕСНАЯ ДЕЗОРГАНИЗАЦИЯ КЛЕТОЧНОГО РЕЦЕПТОРА ГКГС**

(31) **62/269,410; 62/305,097; 62/329,439**

(32) **2015.12.18; 2016.03.08; 2016.04.29**

(33) **US**

(43) **2019.01.31**

(86) **PCT/US2016/066988**

(87) **WO 2017/106537 2017.06.22**

(71)(73) Заявитель и патентовладелец:

САНГАМО ТЕРАПЬЮТИКС, ИНК.

(US)

(72) Изобретатель:

Ли Гэри К., Пацен Дэвид, Чжан Лей

(US)

(74) Представитель:

Нилова М.И. (RU)

(56) **WO-A1-2015136001**

WO-A2-2007127999

WO-A2-2011012691

US-A1-20140093913

US-A1-20050064474

US-A1-20150128309

(57) В изобретении раскрыты способы и композиции для инактивации генов ГКГС (МНС) с применением нуклеаз с мотивом цинкового пальца (ZFN), содержащих белок с мотивом цинкового пальца и расщепляющий домен или расщепляющий полудомен, в условиях, позволяющих сохранять жизнеспособность клеток. Кроме того, предложены полинуклеотиды, кодирующие ZFN, векторы, содержащие полинуклеотиды, кодирующие ZFN, и клетки, содержащие полинуклеотиды, кодирующие ZFN, и/или клетки, содержащие ZFN.

045196

B1

045196

B1

Ссылки на родственные заявки

Настоящая заявка испрашивает приоритет предварительной заявки на патент США № 62/269410, поданной 18 декабря 2015 г.; предварительной заявки на патент США № 62/305097, поданной 8 марта 2016 г.; и предварительной заявки на патент США № 62/329439, поданной 29 апреля 2016 г., содержание которых полностью включено в настоящий документ посредством ссылок.

Область техники

Настоящее изобретение относится к области геномной модификации клеток человека, в том числе лимфоцитов и стволовых клеток.

Область техники

Генная терапия создает широкие возможности для новой эры терапевтических средств для человека. Эти методологии дадут возможность лечения состояний, неизлечимых в рамках стандартной медицинской практики. Генная терапия может включать различные варианты способов редактирования генома, например, дезорганизацию или коррекцию генного локуса, инсерцию экспрессируемого трансгена, который может находиться под управлением специфического экзогенного промотора, объединенного с трансгеном, или эндогенного промотора, находящегося в сайте инсерции в геноме. Доставка и инсерция трансгена являются примерами препятствий, которые необходимо решить для практического осуществления данной технологии. Например, несмотря на большое количество способов доставки генов, потенциально доступных для терапевтического применения, все они имеют существенные недостатки с точки зрения безопасности, стабильности и уровня экспрессии. Способы, позволяющие получить трансген в виде эписомы (например, системы на основе обычного аденовируса (Ad), аденоассоциированного вируса (AAV) и плазмид) обычно безопасны и позволяют достичь высокого начального уровня экспрессии, однако эти способы не обеспечивают надежную репликацию эписомы, что может ограничивать продолжительность экспрессии в митотически активных тканях. В противоположность этому, способы доставки, приводящие к случайному встраиванию желательного трансгена (например, встраивающийся лентивирус (LV)), обеспечивают более продолжительную экспрессию, однако вследствие неспецифической природы случайной экспрессии могут вызывать нерегулируемый рост клеток-реципиентов, что может привести к появлению злокачественных новообразований за счет активации онкогенов в непосредственной близости от кассеты трансгена, встроенной случайным образом. Кроме того, хотя встраивание трансгена позволяет избежать потерь, обусловленных репликацией, она не предотвращает возможный сайленсинг экзогенного промотора, объединенного с трансгеном. Со временем такой сайленсинг приводит к снижению экспрессии трансгена в большинстве случаев неспецифической инсерции. Кроме того, встраивание трансгена редко происходит в каждой клетке-мишени, что затрудняет достижение достаточно высокого уровня экспрессии требуемого трансгена для получения желательного терапевтического эффекта.

В последние годы разработана новая стратегия встраивания трансгена, в которой применяют расщепление сайт-специфическими нуклеазами (например, нуклеазами с мотивом цинкового пальца (ZFN), нуклеазами с эффекторным доменом, подобным активатору транскрипции (TALEN), системами CRISPR/Cas с рекомбинантной CRISPR-ПНК (crRNA/транс-активируемой CRISPR-ПНК (tracrRNA) ("одиночной направляющей ПНК") и системами Cpf1/CRISPR для управления специфическим расщеплением и т.д.) для смещения инсерции в выбранный локус генома. См., например, патенты США № 9255250; 9045763; 9005973; 8956828; 8945868; 8703489; 8586526; 6534261; 6599692; 6503717; 6689558; 7067317; 7262054; 7888121; 7972854; 7914796; 7951925; 8110379; 8409861; публикации патентов США 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983, 20130177960; и 20150056705. Кроме того, разрабатываются нуклеазы адресного действия на основе системы Argonaute (например, T. thermophilus, называемая "TtAgo", см. Swarts et al. (2014), Nature, 507(7491):258-261), которые также могут обладать определенным потенциалом для редактирования генома и генной терапии.

Подход к встраиванию трансгена, опосредованный нуклеазами, открывает перспективы улучшения экспрессии трансгена, повышения безопасности и продолжительности экспрессии по сравнению с классическими подходами к встраиванию, поскольку он обеспечивает точное позиционирование трансгена с минимальным риском сайленсинга гена или активации близлежащих онкогенов. Т-клеточный рецептор (TCR) представляет собой важнейшую часть селективной активации Т-клеток. Обладая некоторым сходством с антителом, фрагмент TCR, обеспечивающий распознавание антигена, обычно состоит из двух цепей - α и β , совместно образующих гетеродимер. Сходство с антителом лежит в основе механизма, по которому происходит объединение одиночных генов, кодирующих альфа- и бета-комплекс TCR. Каждая из альфа- и бета-цепей TCR (TCR α и TCR β) состоит из двух областей - С-концевой константной области и N-концевой варибельной области. Локусы генома, кодирующие альфа- и бета-цепи TCR, напоминают локусы, кодирующие антитела, в том смысле, что ген TCR α содержит V- и J-сегменты, в то время как локус β -цепи помимо V- и J-сегментов содержит D-сегменты. β -локус TCR дополнительно содержит две различные константные области, отбираемые в процессе отбора. При развитии Т-клетки происходит рекомбинация различных сегментов, так что каждая Т-клетка содержит уникальный варибельный фрагмент альфа- и бета-цепей TCR, называемый участком, определяющим комплементарность (CDR), а орга-

низм содержит обширный репертуар Т-клеток, которые благодаря своим уникальным CDR могут взаимодействовать с уникальными антигенами, экспонируемыми антиген-презентирующими клетками. После реаранжировки α - или β -гена TCR экспрессия второго соответствующего TCR α или TCR β репрессуруется, так что каждая Т-клетка экспрессирует только одну уникальную структуру TCR в ходе процесса, называемого "аллельным исключением рецептора антигена" (см. Brady et al. (2010), J. Immunol., 185:3801-3808).

Во время активации Т-клеток TCR взаимодействует с антигенами, экспонируемыми в виде пептидов на главном комплексе гистосовместимости (ГКГС, МНС) антиген-презентирующей клетки. Распознавание комплекса "антиген-ГКГС" рецептором TCR приводит к стимуляции Т-клетки, что, в свою очередь, ведет к дифференцировке как Т-хелперов ($CD4^+$), так и цитотоксических Т-лимфоцитов ($CD8^+$) в клетки памяти и эффекторные лимфоциты. Затем эти клетки могут размножаться путем клонирования, образуя активированную субпопуляцию в общей популяции Т-клеток, способную реагировать с одним конкретным антигеном.

Белки ГКГС подразделяются на два класса - I и II. Белки ГКГС I класса представляют собой гетеродимеры двух белков - α -цепи, представляющей собой трансмембранный белок, кодируемый генами ГКГС I класса, и $\beta 2$ -микроглобулиновую цепь (иногда называемую В2М), представляющую собой небольшой внеклеточный белок, кодируемый геном, расположенным за пределами генного кластера ГКГС. α -цепь образует три глобулярных домена, и при ассоциации $\beta 2$ -микроглобулиновой цепи глобулярный структурный комплекс становится аналогичен комплексу антитела. Чужеродные пептиды презантируются на двух крайних N-концевых доменах, характеризующихся наибольшей вариабельностью. Белки ГКГС II класса также представляют собой гетеродимеры, однако эти гетеродимеры содержат два трансмембранных белка, кодируемые генами, расположенными в пределах комплекса ГКГС. Комплекс "ГКГС I класса-антиген" взаимодействует с цитотоксическими Т-клетками, а ГКГС II класса представляет антиген Т-хелперам. Кроме того, белки ГКГС I класса обычно экспрессируются почти во всех клетках с ядрами и тромбоцитах (а также эритроцитах у мышей), в то время как белки ГКГС II класса экспрессируются более селективно. Обычно белки ГКГС II класса экспрессируются на В-клетках, некоторых макрофагах и моноцитах, клетках Лангерганса и дендритных клетках.

Генный кластер HLA I у человека содержит три основных локуса - B, C и A, а также несколько минорных локусов (в том числе E, G и F, все из которых находятся в области HLA на 6 хромосоме). Кластер HLA II класса также содержит три основных локуса - DP, DQ и DR, причем генные кластеры как I, так и II класса являются полиморфными в том смысле, что в популяции присутствуют несколько различных аллелей генов как I, так и II класса. Кроме того, существует несколько вспомогательных белков, также играющих роль в функционировании HLA. $\beta 2$ -микроглобулин (кодируемый геном В2М, расположенным на 15 хромосоме) действует как шаперон и стабилизирует белок HLA A, B или C, экспрессируемый на поверхности клетки, а также стабилизирует паз для экспонирования антигена в структуре I класса. Он в норме обнаруживается в небольших количествах в сыворотке и моче.

HLA играет основную роль в отторжении трансплантата. Острая фаза отторжения трансплантата может произойти приблизительно в течение 1-3 недель и обычно включает действие Т-лимфоцитов хозяина на донорскую ткань из-за сенсibilизации системы хозяина по отношению к молекулам HLA I и II класса донора. В большинстве случаев иницирующими антигенами являются HLA I класса. Для успешной трансплантации доноров типировать по HLA; они должны в максимально возможной степени совпадать с HLA пациента-реципиента. В то же время, донорство даже между членами семьи, характеризующимися идентичностью значительной доли HLA, часто бывает неудачным. Таким образом, для сохранения ткани трансплантата в организме реципиента пациента часто необходимо подвергать интенсивной терапии с применением иммунодепрессантов в целях профилактики отторжения. Такая терапия может приводить к осложнениям и повышенной заболеваемости из-за инфекций, вызываемых условно-патогенными организмами, с которым организму пациента может быть сложно бороться. Возможна дезорганизация регуляции генов I и II классов в присутствии некоторых опухолей; последствия такой дезорганизации влияют на прогноз у пациента. Например, при метастатическом раке ободочной и прямой кишки обнаружено снижение экспрессии В2М (Shrout et al. (2008), Br. J. Cane, 98:1999). Поскольку В2М играет ключевую роль в стабилизации комплекса ГКГС I класса, существует гипотеза о том, что потеря В2М при некоторых солидных раковых опухолях является механизмом ускользания от иммунологического надзора, опосредованного Т-клетками. Показано, что подавление экспрессии В2М обусловлено супрессией обычной регуляции экспрессии В2М за счет ИФН и/или специфическими мутациями в кодирующей последовательности В2М, приводящими к нокауту гена (Shrout et al., *ibid*). Вместе с тем, повышение уровня В2М также ассоциировано с некоторыми видами рака. Повышенный уровень В2М в моче является предвестником некоторых видов рака, в том числе рака предстательной железы, хронического лимфоцитарного лейкоза (ХЛЛ) и неходжкинских лимфом.

Адоптивная клеточная терапия (АКТ) является развивающейся формой лечения рака, основанной на доставке опухолеспецифических иммунных клеток в организм пациента в целях атаки и уничтожения рака у пациента доставленными клетками. АКТ может включать применение лимфоцитов, инфильтри-

рующих опухоль (TIL), представляющих собой Т-клетки, выделенные из собственной опухолевой массы пациента и размноженные *ex vivo* с целью обратного вливания пациенту. Данный подход перспективен при лечении метастатической меланомы, причем в одном исследовании получена частота долгосрочных ответов >50% (см., например, Rosenberg et al. (2011), *Clin. Canc. Res.*, 17(13):4550). TIL являются перспективным источником клеток, поскольку они представляют собой смешанный набор собственных клеток пациента, содержащих Т-клеточные рецепторы (TCR), специфичные по отношению к опухолеассоциированным антигенам (ТАА), присутствующим в опухоли (Wu et al. (2012), *Cancer J.*, 18(2):160). Другие подходы включают редактирование Т-клеток, выделенных из крови пациента, путем их модификации с целью придания некоторой способности реагировать на опухоль (Kalos et al. (2011), *Sci. Transl. Med.*, 3(95):95ra73).

Химерные рецепторы антигенов (CAR) представляют собой молекулы, разработанные для обеспечения адресного (таргетного) воздействия иммунных клеток на специфические молекулярные мишени, экспрессируемые на поверхности клеток. В простейшем виде они представляют собой рецепторы, внедренные в клетку, объединяющие специфичный домен, экспрессируемый на внешней стороне клетки, с сигнальными путями внутри клетки, так что при взаимодействии специфического домена с его мишенью происходит активация клетки. Другие CAR получают, имитируя функциональные домены Т-клеточных рецепторов (TCR), где антиген-специфичный домен, например, scFv или рецепторы некоторых видов, объединяют с сигнальным доменом, например, НАМ, и другими доменами, обеспечивающими совместное стимулирующее действие. Затем указанные конструкторы внедряют в Т-клетку *ex vivo*, что обеспечивает активацию Т-клетки в присутствии клетки, экспрессирующей антиген-мишень, и приводит к атаке клетки-мишени активированной Т-клеткой независимо от ГКГС (см. Chicaubam et al. (2011), *Int. Rev. Immunol.*, 30:294-311, Kalos *ibid*) при обратном введении Т-клетки в организм пациента. Таким образом, адоптивная клеточная терапия с применением Т-клеток, модифицированных *ex vivo* посредством внедрения рекомбинантного TCR или CAR, представляет собой крайне перспективный клинический подход для лечения некоторых видов заболеваний. Например, раковые заболевания и их антигены, по отношению к которым обеспечивается адресное воздействие, включают фолликулярную лимфому (CD20 или GD2), нейробластому (CD171), неходжкинскую лимфому (CD19 и CD20), лимфому (CD19), глиобластому (IL13R α 2), хронический лимфоцитарный лейкоз, или ХЛЛ, и острый лимфоцитарный лейкоз, или ALL (для этих обоих заболеваний - CD19). Кроме того, разработаны вирус-специфические CAR для атаки клеток, несущих вирусы, например, ВИЧ. Например, начато клиническое исследование с применением CAR, специфичного по отношению к Grp100, для лечения ВИЧ (Chicaubam, *ibid*).

ACTR (Т-клеточные рецепторы, сопряженные с антителом) представляют собой рекомбинантные компоненты Т-клеток, способные связываться с экзогенным антителом. Связывание антитела с плечами ACTR-компонента придает Т-клетке способность взаимодействовать с антигеном, распознаваемыми антителом, и при контакте с антигеном инициируется взаимодействие ACTR-содержащей Т-клетки с антигеном (см. публикацию патента США 20150139943).

В то же время, одним из недостатков адоптивной клеточной терапии является то, что источник клеточного продукта должен быть пациент-специфичным (аутологичным) во избежание потенциального отторжения трансплантированных клеток. Эти привело исследователей к разработке способов редактирования собственных Т-клеток пациента во избежание такого отторжения. Например, Т-клетки или гемопоэтические стволовые клетки пациента подвергают манипуляции *ex vivo* путем добавления рекомбинантного CAR, ACTR и/или Т-клеточного рецептора (TCR), а затем обрабатывают рекомбинантными нуклеазами с целью нокаута ингибиторов ключевых компонентов Т-клеток, например, PD1 и/или CTLA4 (см. публикацию PCT WO 2014/059173). Для применения этой технологии на крупной популяции пациентов желательно разработать универсальную популяцию клеток (аллогенных). Кроме того, нокаут TCR приведет к неспособности клеток вызывать реакцию трансплантат против хозяина (РТПХ) при введении в организм пациента.

Таким образом, существует потребность в способах и композициях, которые можно применять для модификации экспрессии гена MHC (например, нокаута B2M) и/или нокаута экспрессии TCR в Т-клетке.

Краткое описание

В настоящем изобретении описаны композиции и способы частичной или полной инактивации или дезорганизации (disruption) гена B2M и композиции и способы внедрения и экспрессии желательного уровня экзогенных трансгенов TCR, CAR или ACTR в Т-лимфоцитах после дезорганизации эндогенного TCR и/или B2M или одновременно с ней. Кроме того, в настоящем изобретении предложены способы и композиции удаления (инактивации) или репрессии гена B2M для получения Т-клетки, стволовой клетки, ткани или целого организма без HLA I класса, например клетки, не экспрессирующей одного или более из рецепторов HLA на своей поверхности. В некоторых вариантах реализации клетки или ткани без HLA представляют собой клетки или ткани человека, обладающие преимуществами для применения в трансплантатах. В предпочтительных вариантах реализации Т-клетки без HLC получают для применения в адоптивной Т-клеточной терапии. В одном аспекте в настоящем изобретении описана выделенная клетка (например, эукариотическая клетка, например, клетка млекопитающего, включая лимфоидную клетку, стволовую клетку (например, ИПСК, эмбриональную стволовую клетку, МСК или ГСК), или

прогениторную клетку), в которой экспрессию гена бета-2-микроглобулина (B2M) модулируют путем модификации 1 или 2 экзона гена B2M. В некоторых вариантах реализации модификацию вносят в последовательность, как показано в одной или более из SEQ ID NO: 6-48 или 137-205; в пределах 1-5, в пределах 1-10 или в пределах 1-20 пар оснований с каждой стороны (фланкирующей геномной последовательности) SEQ ID NO: 6-48 или 137-205; или в пределах GGCCTTA, TCAAAT, TCAAATT, TTA CTGA и/или AATTGAA. Эту модификацию можно выполнить за счет экзогенной гибридной молекулы, содержащей функциональный домен (например, домен регуляции транскрипции, нуклеазный домен) и ДНК-связывающий домен, включая

(i) клетку, содержащую экзогенный фактор транскрипции, содержащий ДНК-связывающий домен, связывающийся с сайтом-мишенью, как показано в любой из SEQ ID NO: 6-48 или 137-205, и домен регуляции транскрипции, в которой фактор транскрипции модифицирует экспрессию гена B2M; и/или

(ii) клетку, содержащую инсерцию и/или делецию в пределах одной или более из SEQ ID NO: 6-48 или 137-205; в пределах 1-5, в пределах 1-10 или в пределах 1-20 пар оснований с каждой стороны (фланкирующей геномной последовательности) SEQ ID NO: 6-48 или 137-205; или в пределах GGCCTTA, TCAAAT, TCAAATT, TTA CTGA и/или AATTGAA, но не ограничиваясь ими.

Клетка может включать дальнейшие модификации, например, инактивированный ген T-клеточного рецептора, ген PD1, и/или CTLA4, и/или трансген, кодирующий химерный рецептор антигена (CAR), трансген, кодирующий T-клеточный рецептор, сопряженный с антителом (ACTR), и/или трансген, кодирующий антитело. Кроме того, предложены фармацевтические композиции, содержащие любые клетки, описанные в настоящем изобретении, а также способы применения клеток и фармацевтических композиций в терапевтических средствах *ex vivo* для лечения нарушения (например, рака) у субъекта.

Так, в одном аспекте в настоящем изобретении описаны клетки с модулированной (например, активированной, репрессированной или инактивированной) экспрессией гена B2M. В предпочтительных вариантах реализации модулирован 1 или 2 экзон гена B2M. Модуляцию можно выполнить за счет экзогенной молекулы (например, рекомбинантного фактора транскрипции, содержащего ДНК-связывающий домен и домен активации или репрессии транскрипции), связывающейся с геном B2M и регулирующей экспрессию B2M, и/или за счет модификации последовательности гена B2M (например, с помощью нуклеазы, расщепляющей ген B2M и модифицирующей последовательность гена путем инсерций и/или делеций). В некоторых вариантах реализации, кроме того, осуществляют модификацию экспрессии одного или более дополнительных генов (например, гена TCR, например, гена TCRA). В некоторых вариантах реализации описаны клетки, содержащие рекомбинантную нуклеазу с целью нокаута гена B2M, например, с дестабилизацией комплекса HLA I класса. В предпочтительных вариантах реализации дестабилизация HLA I класса приводит к выраженной потере комплекса HLA I класса на поверхности клетки. В других вариантах реализации описаны клетки, которые содержат экзогенную молекулу, например, рекомбинантного фактора транскрипции (TF), что приводит, например, к модуляции гена B2M). В некоторых вариантах реализации клетки представляют собой T-клетки. Кроме того, описаны клетки с модулированной экспрессией гена B2M, причем указанные клетки дополнительно модифицированы и содержат по меньшей мере один экзогенный трансген или дополнительный нокаут по меньшей мере одного эндогенного гена или их комбинацию. Экзогенный трансген можно внедрить в ген TCR (например, при нокауте гена TCR), в ген B2M или в локус, не имеющий отношения к TCR и B2M, например, локус для безопасного носительства. В некоторых случаях экзогенный трансген кодирует ACTR, рекомбинантный TCR и/или CAR. Трансгенный конструкт можно вставить посредством процесса, опосредованного HDR или NHEJ. В некоторых аспектах клетки с модулированной экспрессией B2M не содержат HLA I класса на своей поверхности и дополнительно содержат по меньшей мере один экзогенный ACTR и/или экзогенный CAR. Некоторые клетки, содержащие модулятор B2M, дополнительно включают нокаут одного или более генов ингибиторов ключевых компонентов иммунного ответа. В некоторых вариантах реализации ингибитор ключевого компонента иммунного ответа представляет собой PD1. В других вариантах реализации ингибитор ключевого компонента иммунного ответа представляет собой CTLA4. В дополнительных аспектах клетка с модулированным B2M включает нокаут PD1 и нокаут CTLA4. В некоторых вариантах реализации клетка дополнительно модифицирована по гену TCR. В некоторых вариантах реализации модулированный ген TCR представляет собой ген, кодирующий TCR β (TCRB). В некоторых вариантах реализации это достигается путем адресного расщепления константной области этого гена (константной области TCR β , или TRBC). В некоторых вариантах реализации ген TCR представляет собой ген, кодирующий TCR α (TCRA). В дополнительных вариантах реализации инсерция достигается посредством адресного расщепления константной области гена TCR, в том числе адресного расщепления гена TCR α (в настоящем изобретении называемой последовательности "TRAC"). В некоторых вариантах реализации клетки с модифицированным геном B2M дополнительно модифицированы по гену TCR, генам HLA-A, -B, -C или гену TAP или любой их комбинации. В некоторых вариантах реализации также модифицируют регулятор HLA II класса - STPA. В некоторых вариантах реализации клетки, описанные в настоящем изобретении, включают модификацию (например, делецию и/или инсерцию, связывание рекомбинантного фактора транскрипции с целью репрессии B2M) гена B2M (например, модификацию

экзона 1 или 2). В некоторых вариантах реализации указанная модификация представляет собой модификацию SEQ ID NO: 6-48 и/или 137-205, в том числе модификацию посредством связывания, расщепления, инсерции и/или делеции одного или более из нуклеотидов в пределах любой из этих последовательностей и/или в пределах 1-50 пар оснований (включая любое значение между ними, например, 1-5, 1-10 или 1-20 пар оснований) генных (геномных) последовательностей, фланкирующих эти последовательности в гене В2М. В некоторых вариантах реализации клетки включают модификацию (связывание, расщепление, инсерции или делеции) в пределах одного или более из следующих последовательностей: GGCCTTA, TCAAATT, TCAAAT, TTA CTGA и/или AATTGAA в пределах гена В2М (например, 1 и/или 2 экзона, см. фиг. 1). В некоторых вариантах реализации модификация включает связывание рекомбинантного TF, описанного в настоящем изобретении, например, приводящее к модуляции, например, репрессии или активации экспрессии. В других вариантах реализации модификация представляет собой генетическую модификацию (изменение нуклеотидной последовательности) в или вблизи от сайта(ов) связывания (сайта-мишени) нуклеаз(ы) и/или расщепления, включая модификации последовательностей в пределах 1-300 (или любого количества пар оснований между ними) пар оснований выше, ниже и/или включая 1 или более из пар оснований сайта(ов) расщепления или связывания; модификации в пределах 1-100 пар оснований (или любого количества пар оснований между ними), включающих сайт(ы) связывания и/или расщепления, и/или с любой стороны от него/них; модификации в пределах 1-50 пар оснований (или любого количества пар оснований между ними), включающих сайт(ы) связывания и/или расщепления, и/или с любой стороны от него/них (например, 1-5, 1-10, 1-20 или более пар оснований); и/или модификации одной или более из пар оснований в пределах сайта связывания нуклеазы и/или сайта расщепления, но не ограничиваясь ими. В некоторых вариантах реализации модификация расположена в области или поблизости от (например, 1-300, 1-50, 1-20, 1-10 или 1-5 или более) пар оснований или любого количества пар оснований между ними) последовательности гена В2М, окружающей любую из SEQ ID NO: 6-48 или 137-205. В некоторых вариантах реализации модификация включает модификацию гена В2М в пределах одной или более из последовательностей, показанных в SEQ ID NO: 6-48 или 137-205, или в пределах GGCCTTA, TCAAATT, TCAAAT, TTA CTGA и/или AATTGAA гена В2М (например, 1 и/или 2 экзона), например, модификацию 1 или более из пар оснований одного или более из этих последовательностей. В некоторых вариантах реализации генетические модификации, опосредованные нуклеазами, происходят между спаренными сайтами-мишенями (при использовании димера для расщепления мишени). Генетические модификации, опосредованные нуклеазами, могут включать инсерции и/или делеции любого количества пар оснований, в том числе инсерции некодирующих последовательностей любой длины и/или трансгенов любой длины, и/или делеции от 1 пары оснований до более чем 1000 т.п.о (или любого значения между ними, включая 1-100 пар оснований, 1-50 пар оснований, 1-30 пар оснований, 1-20, 1-10 или 1-5 пар оснований, но не ограничиваясь ими).

Модифицированные клетки согласно настоящему изобретению могут представлять собой лимфоидную клетку (например, Т-клетку), стволовую клетку/прогениторную клетку (например, индуцированную плюрипотентную стволовую клетку (ИПСК), эмбриональную стволовую клетку (например, ЭСК человека), мезенхимальную стволовую клетку (МСК) или гемопоэтическую стволовую клетку (ГСК). Стволовые клетки могут быть тотипотентными или плюрипотентными (например, частично дифференцированными, как ГСК, которая может представлять собой плюрипотентную миелоидную или лимфоидную стволовую клетку). В других вариантах реализации настоящего изобретения предложены способы получения клеток, обладающих нулевым фенотипом экспрессии HLA. Любую из модифицированных стволовых клеток, описанных в настоящем изобретении (модифицированных по локусу В2М), можно затем дифференцировать с целью получения дифференцированной (*in vivo* или *in vitro*) клетки, происходящей от стволовой клетки, описанной в настоящем изобретении. Любая из модифицированных стволовых клеток, описанных в настоящем изобретении, может содержать дополнительные модификации в других требуемых генах (например, TCRA, TCRB, PD1, CTLA4 и т.д.).

В еще одном аспекте композиции (модифицированных клеток) и способы, описанные в настоящем изобретении, можно применять, например, при лечении или профилактике или облегчении нарушения. Эти способы обычно включают

(а) расщепление или подавление эндогенного гена В2М в выделенной клетке (например, Т-клетке или лимфоците) с помощью нуклеазы (например, ZFN или TALEN) или нуклеазной системы, например, CRISPR/Cas или Cfp1/CRISPR с рекомбинантной РНК crRNA/tracr, или с помощью рекомбинантного фактора транскрипции (например, ZFN-TF, TALE-TF, Cfp1-TF или Cas9-TF), сопровождающееся инактивацией или подавлением гена В2М; и

(б) внедрение клетки в организм субъекта, за счет чего происходит лечение или профилактика нарушения.

В некоторых вариантах реализации, кроме того, происходит инактивация или подавление гена, кодирующего TCR α (TCRA) и/или TCR β (TCRB). В некоторых вариантах реализации инактивация достигается путем адресного расщепления константной области этого гена (константной области TCR β , или TRBC). В предпочтительных вариантах реализации происходит инактивация или подавление гена, коди-

рующего TCR α (TCRA). В других предпочтительных вариантах реализации нарушение представляет собой рак или инфекционное заболевание. В дополнительных предпочтительных вариантах реализации инактивация достигается путем адресного расщепления константной области этого гена (константной области TCR α , сокращенно -TRAC).

Факторы транскрипции и/или нуклеазу(ы) можно внедрить в клетку в виде мРНК, белка и/или последовательности ДНК, кодирующей нуклеазу(ы). В некоторых вариантах реализации выделенная клетка, внедряемая в организм субъекта, содержит дополнительную геномную модификацию, например, внедренную экзогенную последовательность (в расщепленном гене B2M, TCR или другом гене, например, гене или локусе для безопасного носительства), или инактивацию (например, опосредованную нуклеазой) дополнительных генов, например, одного или более из генов HLA и/или TAP. Экзогенную последовательность можно внедрить посредством вектора (например, Ad, AAV, LV), или посредством такой методики, как электропорация. В некоторых вариантах реализации белки внедряют в клетку посредством сжатия клетки (см. Kollmannsperger et al. (2016), Nat. Comm., 7, 10372, DOI: 10.1038/ncomms10372). В некоторых аспектах композиция может содержать фрагменты выделенной клетки и/или дифференцированные (частично или полностью) клетки.

В некоторых аспектах зрелые клетки можно применять для клеточной терапии, например, адоптивного переноса клеток. В других вариантах реализации клетки для применения при трансплантации Т-клеток содержат еще одну требуемую модификацию гена. В одном аспекте Т-клетки содержат вставленный химерный рецептор антигена (CAR), специфичный по отношению к маркеру рака. В дополнительном аспекте вставленный CAR является специфичным по отношению к маркеру CD19, характерному для В-клеточных злокачественных новообразований. Такие клетки можно применять в терапевтической композиции для лечения пациентов без соответствия HLA, и, таким образом, можно применять в качестве "готового" терапевтического средства для любого пациента, нуждающегося в этом.

В еще одном аспекте Т-клетки с модулированным (модифицированным) B2M содержат вставленную донорную последовательность Т-клеточного рецептора, сопряженного с антителом (ACTR). В некоторых вариантах реализации донорную последовательность ACTR вставляют в ген B2M или TCR для дезориентации экспрессии указанного гена после нуклеазного расщепления. В других вариантах реализации донорную последовательность вставляют в локус "безопасного носительства", например, гены AAVS1, HPRT, альбумина и CCR5. В некоторых вариантах реализации последовательность ACTR вставляют путем адресного внедрения, при котором донорная последовательность ACTR содержит фланкирующие гомологичные плечи, характеризующиеся гомологией по отношению к последовательности, фланкирующей сайт расщепления рекомбинантной нуклеазы. В некоторых вариантах реализации донорная последовательность ACTR дополнительно содержит промотор и/или другую последовательность для регуляции транскрипции. В других вариантах реализации донорная последовательность ACTR не содержит промотора. В некоторых вариантах реализации донорный ACTR вставляют в ген, кодирующий TCR β (TCRB). В некоторых вариантах реализации инсерция достигается путем адресного расщепления константной области этого гена (константной области TCR β , или TRBC). В предпочтительных вариантах реализации донорный ACTR вставляют в ген, кодирующий TCR α (TCRA). В дополнительных предпочтительных вариантах реализации инсерция достигается путем адресного расщепления константной области этого гена (константной области TCR α , сокращенно - TRAC). В некоторых вариантах реализации донорную последовательность вставляют в последовательность экзона в TCRA, в то время как в других вариантах реализации донорную последовательность вставляют в последовательность интрона в TCRA. В некоторых вариантах реализации клетки с модулированным TCR дополнительно содержат CAR. В дополнительных вариантах реализации клетки с модулированным B2M дополнительно модулируют по гену HLA или гену ингибитора ключевого компонента иммунного ответа. Кроме того, предложены фармацевтические композиции, содержащие модифицированные клетки, описанные в настоящем изобретении (например, Т-клетки или стволовые клетки с инактивированным геном B2M), или фармацевтические композиции, содержащие одну или более из B2M-связывающих молекул (например, рекомбинантных факторов транскрипции и/или нуклеаз), описанных в настоящем изобретении. В некоторых вариантах реализации фармацевтические композиции дополнительно содержат одно или более из фармацевтически приемлемых вспомогательных веществ. Модифицированные клетки, B2M-связывающие молекулы (или полинуклеотиды, кодирующие эти молекулы) и/или фармацевтические композиции, содержащие эти клетки или молекулы, вводят в организм субъекта за счет способов, известных в данной области техники, например, посредством внутривенного вливания, вливания в специфический сосуд, например, печеночную артерию, или посредством инъекции непосредственно в ткань (например, мышцу). В некоторых вариантах реализации субъект является взрослым человеком с заболеванием или состоянием, которое можно лечить или облегчить с применением указанной композиции. В других вариантах реализации субъект является субъектом детского возраста, а композицию вводят с целью профилактики, лечения или облегчения заболевания или состояния (например, рака, реакции трансплантат против хозяина и т.д.). В некоторых аспектах композиция (клетки с модулированным B2M, содержащие ACTR) дополнительно содержит экзогенное антитело. Кроме того, см. заявку на патент США № 15/357772. В некоторых аспек-

тах указанное антитело можно применять для активации Т-клеток, содержащих АСТР, с целью профилактики или лечения состояния. В некоторых вариантах реализации антитело распознает антиген, ассоциированный с опухолевой клеткой или с процессами, ассоциированными с раком, например, EpCAM, CEA, gpA33, муцины, TAG-72, CAIX, PSMA, фолат-связывающие антитела, CD19, EGFR, ERBB2, ERBB3, MET, IGF1R, EPHA3, TRAILR1, TRAILR2, RANKL, FAP, VEGF, VEGFR, интегрин $\alpha V\beta 3$ и $\alpha 5\beta 1$, CD20, CD30, CD33, CD52, CTLA4 и тенацин (Scott et al. (2012) Nat Rev Cancer 12:278). В других вариантах реализации антитело распознает антиген, ассоциированный с инфекционным заболеванием, например, ВИЧ, ВГС и т.п.

В еще одном аспекте настоящего изобретения предложены домены, связывающие ДНК В2М (например, ZFP, TALE и одиночные направляющие РНК (онРНК)), связывающиеся с сайтом-мишенью в гене В2М. В некоторых вариантах реализации ДНК-связывающий домен содержит ZFP с распознающими спиральными областями, расположенными в порядке, показанном в одиночной строке табл. 1; ДНК-связывающий белок с TAL-эффекторным доменом и RVD, как показано в одиночной строке табл. 2B; и/или онРНК, показанную в одиночной строке табл. 2A. Эти ДНК-связывающие белки можно ассоциировать с доменами регуляции транскрипции с образованием рекомбинантных факторов транскрипции, модулирующих экспрессию В2М. В качестве альтернативы, ДНК-связывающие белки можно ассоциировать с одним или более из нуклеазных доменов с образованием рекомбинантных нуклеаз с мотивом цинкового пальца (ZFN), TALEN и/или систем CRISPR/Cas, связывающихся с геном В2М и расщепляющих его. В некоторых вариантах реализации ZFN, TALEN или одиночные направляющие РНК (онРНК) системы CRISPR/Cas связываются с сайтами-мишенями в гене В2М человека. ДНК-связывающий домен фактора транскрипции или нуклеазы (например, ZFP, TALE, онРНК) может связываться с сайтом-мишенью в гене В2М, содержащем 9, 10, 11 12 или более (например, 13, 14, 15, 16, 17, 18, 19, 20 или более) нуклеотидов из любой из SEQ ID NO: 6-48 или 137-205. Белки с мотивом цинкового пальца могут содержать 1, 2, 3, 4, 5, 6 или более мотивов цинкового пальца, причем каждый мотив цинкового пальца содержит распознающую спираль, специфически контактирующую с субсайтом-мишенью в гене-мишени. В некоторых вариантах реализации белки с мотивом цинкового пальца содержат 4 или 5 или 6 мотивов цинкового пальца (обозначаемых как F1, F2, F3, F4, F5 и F6 и расположенных в порядке от F1 к F4 или F5 или F6 с N-конца к С-концу), например, как показано в табл. 1. В других вариантах реализации одиночные направляющие РНК или ДНК-связывающие домены с TAL-эффектором могут связываться с сайтом-мишенью, показанным в любой из SEQ ID NO: 6-48 или 137-205 (или 12 или более парами оснований в пределах любой из SEQ ID NO: 6-48). Типичные сайты-мишени онРНК показаны в SEQ ID NO: 16-48, а типичные сайты связывания TALEN показаны в табл. 2B (SEQ ID NO: 137-205). Можно сконструировать дополнительные TALEN для сайтов-мишеней, описанных в настоящем изобретении, с применением канонических или неканонических RVD, описанных в патентах США № 8586526 и 9458205. Нуклеазы, описанные в настоящем изобретении (содержащие ДНК-связывающий домен ZFP, TALE или онРНК) могут выполнять генетические модификации в гене В2М, содержащем любую из SEQ ID NO: 6-48 или 137-205, в том числе модификации (инсерции и/или делеции) в пределах любой из этих последовательностей (SEQ ID NO: 6-48 или 137-205) и/или модификации последовательностей гена В2М, фланкирующих последовательности сайта-мишени, показанные в SEQ ID NO: 6-48 или 137-205, например, модификации в пределах 1 или 2 экзона гена В2М в пределах одной или более из следующих последовательностей: GGCCTTA, TCAAATT, TCAAAT, TTAAGTA и/или AATTGAA.

Кроме того, предложена гибридная молекула, содержащая ДНК-связывающий домен, связывающийся с 1 или 2 экзоном гена В2М, и домен регуляции транскрипции или нуклеазный домен, причем ДНК-связывающий домен содержит белок с мотивом цинкового пальца (ZFP), как показано в одиночной строке табл. 1, эффекторный белок TALE, как показано в одиночной строке табл. 2B, или одиночную направляющую РНК (онРНК), как показано в одиночной строке табл. 2A.

Любой из белков, описанных в настоящем изобретении, может дополнительно содержать расщепляющий домен и/или полудомен (например, рекомбинантный расщепляющий полудомен FokI или аналогичный домен дикого типа). Таким образом, нуклеазный домен любой из нуклеаз (например, ZFN, TALEN, систем CRISPR/Cas), описанных в настоящем изобретении, может содержать нуклеазный домен дикого типа или нуклеазный полудомен (например, расщепляющий полудомен FokI). В других вариантах реализации нуклеазы (например, нуклеазы ZFN, TALEN, CRISPR/Cas) содержат рекомбинантные нуклеазные домены или полудомены, например, рекомбинантные расщепляющие полудомены FokI, образующие облигатные гетеродимеры. См., например, публикацию патента США № 20080131962. В еще одном аспекте настоящего изобретения предложен полинуклеотид, кодирующий любой из белков, гибридных молекул и/или их компонентов (например, онРНК или другой ДНК-связывающий домен), описанных в настоящем изобретении. Этот полинуклеотид может входить в состав вирусного вектора, невирусного вектора (например, плазмиды) или находиться в форме мРНК. Любой из полинуклеотидов, описанных в настоящем изобретении, может также содержать последовательности (донорные, гомологичные плечи или последовательности для вставки) для адресной инсерции в ген В2М, TCR α и/или TCR β . В еще одном аспекте предложен вектор для доставки гена, содержащий любой из описанных в настоящем

изобретении полинуклеотидов. В некоторых вариантах реализации вектор представляет собой аденовирусный вектор (например, вектор Ad5/F35) или лентивирусный вектор (LV), включая компетентные или дефектные по встраиванию лентивирусные векторы, или аденоассоциированный вектор (AAV). Таким образом, в настоящем изобретении также предложены вирусные векторы, содержащие последовательность, кодирующую нуклеазу (например, ZFN или TALEN) и/или нуклеазную систему (CRISPR/Cas или Ttago) и/или донорную последовательность для адресного встраивания в ген-мишень. В некоторых вариантах реализации донорная последовательность и последовательности, кодирующие нуклеазу, находятся на разных векторах. В других вариантах реализации нуклеазы представлены в виде полипептидов. В предпочтительных вариантах реализации полинуклеотиды представляют собой мРНК. В некоторых аспектах мРНК можно химически модифицировать (см., например, Kormann et al. (2011), *Nature Biotechnology*, 29(2):154-157). В других аспектах мРНК может содержать кэп ARCA (см. патенты США 7074596 и 8153773). В некоторых аспектах мРНК может содержать кэп, внедренный посредством ферментативной модификации. Кэп, внедренный с помощью ферментов, может содержать Cap0, Cap1 или Cap2 (см., например, Smietanski et al. (2014), *Nature Communications*, 5:3004). В дальнейших аспектах мРНК можно копировать посредством химической модификации. В дополнительных вариантах реализации мРНК может содержать смесь немодифицированных и модифицированных молекул (см. публикацию патента США 2012-0195936). В дополнительных вариантах реализации мРНК может содержать элемент WPRE (см. публикацию патента США 15/141333). В некоторых вариантах реализации мРНК является двуцепочечной (см., например, Kariko et al. (2011), *Nucl. Acid. Res.*, 39:e142).

В еще одном аспекте настоящего изобретения предложена выделенная клетка, содержащая любой из белков, полинуклеотидов и/или векторов, описанных в настоящем изобретении. В некоторых вариантах реализации клетку выбирают из группы, состоящей из стволовой клетки/прогениторной клетки или Т-клетки (например, CD4⁺ Т-клетки). В дополнительном аспекте настоящего изобретения предложена клетка или линия клеток, происходящая от клетки или линии, содержащей любой из белков, полинуклеотидов и/или векторов, описанных в настоящем изобретении, а именно клетка или линия клеток, происходящая (например, в культуре) от клетки, в которой В2М инактивирован одной или более из ZFN, и/или в геном которой стабильно внедрен донорный полинуклеотид (например, ACTR, рекомбинантный TCR и/или CAR). Таким образом, потомство клеток, описанных в настоящем изобретении, само по себе может не содержать белков, полинуклеотидов и/или векторов, описанных в настоящем изобретении, однако в этих клетках по меньшей мере инактивирован ген В2М и/или внедрен в геном и/или экспрессируется донорный полинуклеотид.

В еще одном аспекте в настоящем изобретении описаны способы инактивации гена В2М в клетке путем внедрения одного или более из белков, полинуклеотидов и/или векторов в клетку, как описано в настоящем изобретении в любом из способов, описанных в настоящем изобретении, нуклеазы могут индуцировать адресный мутагенез, делеции последовательностей ДНК в клетке и/или облегчать адресную рекомбинацию в заранее определенном хромосомном локусе. Таким образом, в некоторых вариантах реализации нуклеазы осуществляют делецию или инсерцию одного или более из нуклеотидов из гена-мишени или в него. В некоторых вариантах реализации ген В2М инактивируют посредством расщепления нуклеазой после негомологичного соединения концов. В других вариантах реализации геномную последовательность гена-мишени замещают, например, с применением нуклеазы (или вектора, кодирующего указанную нуклеазу), как описано в настоящем изобретении, и "донорной" последовательности, которую вставляют в ген после адресного расщепления нуклеазой. Донорная последовательность может присутствовать в векторе нуклеазы, в отдельном векторе (например, AAV, Ad или LV-векторе) или, в качестве альтернативы, ее можно внедрить в клетку с помощью другого механизма доставки нуклеиновых кислот.

Кроме того, любой из способов, описанных в настоящем изобретении, можно осуществить *in vitro*, *in vivo* и/или *ex vivo*. В некоторых вариантах реализации указанные способы осуществляют *ex vivo*, например, с целью модификации Т-клеток для возможности их применения в качестве терапевтических средств в аллогенных условиях для лечения субъекта (например, субъекта с раком). Неограничивающие примеры рака, которые можно лечить и/или предотвращать, включают виды карциномы легких, рака поджелудочной железы, рака печени, рака костей, рака молочной железы, рака ободочной и прямой кишки, лейкозы, виды рака яичников, лимфомы, виды рака головного мозга и т.п.

Краткое описание чертежей

Фиг. 1 представляет собой изображение последовательностей 1 и 2 экзона (SEQ ID NO: 1 и 2) гена В2М, подвергающихся адресному воздействию нуклеаз. Прямоугольниками отмечены пять различных областей расщепления (А (GGCCTTA), С (TCAAATT), D (TCAAAT), Е (TTACTGA) и G (AATTGAA)), фланкированные сайтами связывания (мишенями) ZFN.

На фиг. 2А и 2В изображена нуклеазная активность. Фиг. 2А представляет собой гистограмму, на которой изображена доля модификации генов в каждом сайте в Т-клетках, обработанных ZFN, специфичными по отношению к сайтам А, С, D, Е и G в В2М, как показано на фиг. 1, в дозе 2 или 6 мкг. На фиг. 2В показана активность TALEN по отношению к гену В2М в клетках K562.

На фиг. 3 изображена процентная доля HLA-отрицательных Т-клеток после обработки

В2М-специфичными парами ZFN, согласно FACS-анализу.

На фиг. 4А-4Е изображены результаты FACS для клеток, обработанных как В2М-, так и TCRA-специфичными ZFN. На фиг. 4А изображены результаты при отсутствии обработки ZFN, на фиг. 4В показаны результаты после воздействия только TCRA-специфичных ZFN (96% нокаут сигнала CD3), а на фиг. 4С показаны результаты после воздействия только В2М-специфичных ZFN (92% нокаут сигнала HLA). Фиг. 4D представляет собой иллюстрацию, на которой показано расположение клеток с двойным нокаутом (приводящим к потере как HLA-, так и CD3-маркеров). На фиг. 4Е показаны результаты после обработки клеток как TCRA-, так и CD3-специфичными ZFN, демонстрирующие двойной нокаут в 82% клеток.

На фиг. 5 показаны результаты двойного нокаута TRAC (TCRA) и В2М и адресного встраивания донорной последовательности в локус TRAC (TCRA) или В2М.

Подробное описание изобретения

В настоящем изобретении описаны композиции и способы получения клеток, в которых экспрессия гена В2М модулирована таким образом, что указанные клетки более не содержат HLA I класса на своей поверхности. Клетки, модифицированные таким образом, можно использовать в качестве терапевтических средств, например, трансплантатов, поскольку отсутствие экспрессии В2М предотвращает или ослабляет HLA-опосредованный иммунный ответ. Кроме того, в клетки, в которых выполняли манипуляции с геном В2М, можно вставить другие требуемые гены и/или выполнить нокаут других требуемых генов.

Общие сведения.

При реализации способов, а также получении и применении композиций, описанных в настоящем изобретении, если не указано иное, используются стандартные методики в области молекулярной биологии, биохимии, анализа и исследования структуры хроматина, вычислительной химии, культивирования клеток, технологии рекомбинантных ДНК и связанных с ними областях, соответствующие текущему уровню в данных областях техники. Эти методики описаны в литературе в полном объеме. См., например, Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2-е изд., Cold Spring Harbor Laboratory Press, 1989; Third edition, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 и периодические обновления; серии METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, 3-е изд., Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, vol. 304, "Chromatin" (под ред. P.M. Wassarman, A. P. Wolffe), Academic Press, San Diego, 1999; и METHODS IN MOLECULAR BIOLOGY, vol. 119, "Chromatin Protocols" (под ред. P.B. Becker) Humana Press, Totowa, 1999.

Определения.

Термины "нуклеиновая кислота", "полинуклеотид" и "олигонуклеотид" используются взаимозаменяемо и относятся к дезоксирибонуклеотидному или рибонуклеотидному полимеру в линейной или циклической конформации и в одно- или двуцепочечной форме. Для целей настоящего описания не следует считать, что эти термины ограничивают длину полимера. Эти термины могут включать известные аналоги природных нуклеотидов, а также нуклеотиды, модифицированные по основанию, углеводной и/или фосфатной группе (например, с фосфортиоатными каркасами). В общем случае аналог конкретного нуклеотидную обладает той же специфичностью спаривания оснований, т.е. аналог А будет парой для Т. Термины "полипептид", "пептид" и "белок" используются на равных основаниях для обозначения полимера, состоящего из аминокислотных остатков. Данный термин также используется по отношению к аминокислотным полимерам, в которых одна или более из аминокислот представляют собой химические аналоги или модифицированные производные соответствующих природных аминокислот.

Термин "связывание" относится к специфическому по отношению к последовательности нековалентному взаимодействию между макромолекулами (например, между белком и нуклеиновой кислотой). Не все компоненты взаимодействия при связывании должны быть специфичными по отношению к последовательности (например, контактировать с фосфатными остатками в каркасе ДНК), при условии, что это взаимодействие в целом является специфическим по отношению к последовательности. Такие взаимодействия в общем случае характеризуются константой диссоциации (K_d), равной 10^{-6} М⁻¹ или менее. Термин "аффинность" (сородство) относится к силе связывания: повышенная аффинность связывания коррелирует с более низкой K_d . Термин "неспецифическое связывание" относится к неспецифическому взаимодействию между любой требуемой молекулой (например, рекомбинантной нуклеазой) и макромолекулой (например, ДНК), не зависящему от последовательности-мишени.

"ДНК-связывающая молекула" представляет собой молекулу, которая может связываться с ДНК. Такие ДНК-связывающие молекулы могут представлять собой полипептид, домен белка, домен в составе более крупного белка или полинуклеотид. В некоторых вариантах реализации полинуклеотид представляет собой ДНК, в то время как в других вариантах реализации полинуклеотид представляет собой РНК. В некоторых вариантах реализации ДНК-связывающая молекула представляет собой белковый домен нуклеазы (например, домен FokI), в то время как в других вариантах реализации ДНК-связывающая молекула представляет собой направляющий РНК-компонент РНК-зависимой нуклеазы (например, Cas9 или Cfp1).

"Связывающий белок" представляет собой белок, способный нековалентно связываться с другой молекулой. Связывающий белок может связываться, например, с молекулой ДНК (ДНК-связывающий белок), молекулой РНК (РНК-связывающий белок) и/или молекулой белка (белок-связывающий белок). В случае белок-связывающего белка он может связываться сам с собой (с образованием гомодимеров, гомотримеров и т.д.) и/или с одной или более молекулами другого белка или белков. Связывающий белок может характеризоваться более чем одним типом связывающей активности. Например, белки с мотивом цинкового пальца обладают ДНК-связывающей, РНК-связывающей и белок-связывающей активностью. "ДНК-связывающий белок с мотивом цинкового пальца" (или связывающий домен) представляет собой белок или домен в составе более крупного белка, связывающий ДНК в зависимости от последовательности посредством одного или более из мотивов цинкового пальца, которые представляют собой области аминокислотной последовательности в пределах связывающего домена, структура которых стабилизирована за счет координации иона цинка. Термин "ДНК-связывающий белок с мотивом цинкового пальца" часто сокращают как "белок с мотивом цинкового пальца" или ZFP.

"ДНК-связывающий домен TALE" или "TALE" представляет собой полипептид, содержащий один или более из повторяющихся доменов/единиц TALE. Повторяющиеся домены, каждый из которых содержит повторяющийся вариабельный двойной остаток (RVD), участвуют в связывании TALE с его мишенью-последовательностью ДНК. Длина одиночной "повторяющейся единицы" (также называемой "повтором") обычно составляет 33-35 аминокислот, эта единица характеризуется по меньшей мере некоторой гомологией последовательности по отношению к другим повторяющимся последовательностям TALE в составе природного белка TALE. Можно сконструировать белки TALE, связывающиеся с сайтом-мишенью с использованием канонических или неканонических RVD в составе повторяющихся единиц. См., например, патенты США № 8586526 и 9458205, полностью включенные в настоящий документ посредством ссылок. Можно "сконструировать" ДНК-связывающие домены с мотивом цинкового пальца и домены TALE, связывающиеся с заранее определенной нуклеотидной последовательностью, например, посредством конструирования (модификации одной или более из аминокислот) распознающей спиральной области природного белка с мотивом цинкового пальца, или конструирования аминокислот, участвующих в связывании ДНК (области повторного вариабельного двойного остатка или RVD). Таким образом, рекомбинантные белки с мотивом цинкового пальца и белки TALE представляют собой белки, не встречающиеся в природе. Неограничивающие примеры способов конструирования белков с мотивом цинкового пальца и TALE представляют собой проектирование и отбор. Спроектированный белок представляет собой белок, не встречающийся в природе, конструкция/состав которого главным образом обусловлены рациональными критериями. Рациональные критерии проектирования включают применение правил замены и компьютеризированных алгоритмов обработки информации в базе данных, где хранится информация о существующих конструкциях ZFP или TALE (канонических и неканонических RVD) и данные о связывании. См., например, патенты США № 9458205; 8586526; 6140081; 6453242; и 6534261; см. также WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536; и WO 03/016496. "Отобранный" белок с мотивом цинкового пальца, белок TALE или система CRISPR/Cas не встречается в природе, их получают в основном за счет эмпирического процесса, например, фагового дисплея, захвата при взаимодействии или гибридной селекции. См., например, патенты US 5789538; US 5925523; US 6007988; US 6013453; US 6200759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; и WO 02/099084.

"TtAgo" представляет собой прокариотический белок Argonaute, который, как считается, участвует в сайленсинге генов. TtAgo получают из бактерии *Thermus thermophilus*. См., например, Swarts № 4816567; и Morrison et al. (2013), Proc. Natl. Acad. Sci., U.S.A., 111, 652). "Система TtAgo" представляет собой все необходимые компоненты, включая, например, направляющие ДНК для расщепления ферментом TtAgo.

"Рекомбинация" относится к процессу обмена генетической информацией между двумя полинуклеотидами. Для целей настоящего описания "гомологичная рекомбинация (ГР)" относится к специализированной форме такого обмена, имеющей место, например, при репарации двуцепочечных разрывов в клетках посредством механизмов репарации, опосредованных гомологией. Этот процесс требует гомологии нуклеотидных последовательностей, использует "донорную" молекулу для матричной репарации молекулы-"мишени" (т.е. молекулы, подвергшейся двуцепочечному разрыву), и известен под различными названиями - "неперекрестная конверсия генов" или "конверсия генов на коротком участке", поскольку он приводит к переносу генетической информации из донорной молекулы в мишень. Безотносительно к теоретическим представлениям, такой перенос может включать исправление несовпадений в гетеродуплексной ДНК между разорванной мишенью и донорной молекулой, и/или "отжиг цепей в зависимости от синтеза", при котором донорную молекулу используют для повторного синтеза генетической информации, которая войдет в состав мишени, и/или связанные с ними процессы. Такая специализированная ГР часто приводит к модификации последовательности молекулы-мишени, так что последовательность донорного полинуклеотида частично или полностью встраивается в полинуклеотид-мишень.

В способах согласно настоящему изобретению одна или более из специфических нуклеаз, описанных в настоящем изобретении, создает двуцепочечный разрыв (DSB) в заранее определенном сайте (например, требуемом гене или локусе) последовательности-мишени (например, хроматина клетки), а "до-

норный" полинуклеотид, обладающий гомологией по отношению к нуклеотидной последовательности области разрыва, можно внедрить в клетку. Показано, что присутствие DSB облегчает внедрение донорной последовательности. Конструкт необязательно характеризуется гомологией по отношению к нуклеотидной последовательности области разрыва. Донорную последовательность можно физически внедрить или, в альтернативном случае, донорную последовательность используют в качестве матрицы для репарации разрыва посредством гомологичной рекомбинации, что приводит к полному или частичному встраиванию донорной нуклеотидной последовательности в хроматин клетки. Таким образом, первую последовательность в хроматине клетки можно модифицировать и, в некоторых вариантах реализации, можно преобразовать в последовательность, присутствующую в донорном полинуклеотиде. Таким образом, следует понимать, что использование терминов "заменить" или "замена" предназначено для обозначения замены одной нуклеотидной последовательности другой последовательностью (т.е. замены последовательности в информационном смысле) и не обязательно требует физической или химической замены одного полинуклеотида другим. В любом из способов, описанных в настоящем изобретении, можно применять дополнительные пары белков с мотивом цинкового пальца для дополнительного двуцепочечного расщепления дополнительных сайтов-мишеней в пределах клетки.

В некоторых вариантах реализации способов адресной рекомбинации и/или замены и/или модификации последовательности в требуемой области хроматина клетки хромосомную последовательность модифицируют путем гомологичной рекомбинации экзогенной "донорной" нуклеотидной последовательностью. Такую гомологичную рекомбинацию стимулируют присутствием двуцепочечного разрыва в хроматине клетки при наличии последовательностей, гомологичных области разрыва. В любом из способов, описанных в настоящем изобретении, первая нуклеотидная последовательность ("донорная" последовательность) может содержать последовательности, гомологичные, но не идентичные геномным последовательностям в требуемой области, тем самым стимулируя гомологичную рекомбинацию для инсерции неидентичной последовательности в требуемую область. Таким образом, в некоторых вариантах реализации фрагменты донорной последовательности, гомологичные последовательностям в требуемой области, характеризуются идентичностью последовательности, равной приблизительно 80-99% (или любому целому числу между ними), по отношению к заменяемой геномной последовательности. В других вариантах реализации гомология между донорной и геномной последовательностями превышает 99%, например, в случае различия между донорными и геномными последовательностями длиной более 100 пар оснований подряд лишь на 1 нуклеотид. В некоторых случаях негомологичный фрагмент донорной последовательности может содержать последовательности, отсутствующие в требуемой области, так что в требуемую область встраиваются новые последовательности. В этих случаях негомологичная последовательность обычно фланкирована последовательностями, длина которых составляет 50-1000 пар оснований (или любое целое число между ними) или любое количество пар оснований, превышающее 1000, гомологичными или идентичными последовательностям требуемой области. В других вариантах реализации донорная последовательность не гомологична первой последовательности и встраивается в геном посредством механизмов негомологичной рекомбинации.

Любой из способов, описанных в настоящем изобретении, можно применять для частичной или полной инактивации одной или более из последовательностей-мишеней в клетке путем адресного встраивания донорной последовательности, дезорганизующего экспрессию требуемого(ых) гена(ов). Кроме того, предложены линии клеток с частично или полностью инактивированными генами.

Кроме того, способы адресного внедрения, описанные в настоящем изобретении, также можно применять для встраивания одной или более экзогенных последовательностей. Экзогенная нуклеотидная последовательность может содержать, например, один или более генов или молекул кДНК или кодирующие или не кодирующие последовательности любого типа, а также один или более из управляющих элементов (например, промоторов). Кроме того, экзогенная нуклеотидная последовательность может продуцировать одну или более из молекул РНК (например, короткие шпилечные РНК (кшРНК), ингибиторные РНК (RNAi), микроРНК (миРНК) и т.д.).

"Расщепление" относится к разрыву ковалентного каркаса молекулы ДНК. Разрыв можно инициировать посредством различных способов, включая ферментативный или химический гидролиз фосфодиэфирной связи, но не ограничиваясь ими. Возможно как одно-, так и двуцепочечное расщепление, и двуцепочечное расщепление может произойти в результате двух отдельных событий одноцепочечного расщепления. Расщепление ДНК может привести к получению тупых или липких концов. В некоторых вариантах реализации для адресного двухцепочечного расщепления ДНК применяют гибридные полипептиды.

"Расщепляющий полудомен" представляет собой полипептид, который при взаимодействии со вторым полипептидом (идентичным или отличающимся) образует комплекс, обладающий расщепляющей активностью (предпочтительно активностью двуцепочечного расщепления). Термин "первый и второй расщепляющие полудомены", "+ и - расщепляющие полудомены" и "правый и левый расщепляющие полудомены" используются на равных основаниях для обозначения димеризующихся пар расщепляющих полудоменов.

"Рекомбинантный расщепляющий полудомен" представляет собой расщепляющий полудомен, мо-

дифицированный с образованием облигатных гетеродимеров с еще одним расщепляющим полудоменом (например, еще одним рекомбинантным расщепляющим полудоменом). См. также патенты США № 7888121; 7914796; 8034598; 8623618 и публикацию патента США № 2011/0201055, полностью включенные в настоящий документ посредством ссылок.

Термин "последовательность" относится к нуклеотидной последовательности любой длины, которая может представлять собой ДНК или РНК, может быть линейной, циклической или разветвленной, а также одноцепочечной или двуцепочечной. Термин "донорная последовательность" относится к нуклеотидной последовательности, вставляемой в геном. Длина донорной последовательности может быть любой, например, может составлять от 2 до 10000 нуклеотидов (или любое целое число между ними или превышающее их), предпочтительно от приблизительно 100 до 1000 нуклеотидов (или любое целое число между ними), более предпочтительно от приблизительно 200 до 500 нуклеотидов.

"Хроматин" представляет собой нуклеопротеиновую структуру, составляющую геном клетки. Хроматин клетки содержит нуклеиновую кислоту, главным образом ДНК, и белок, в том числе гистоны и негистонные белки хромосом. Большая часть хроматина эукариотической клетки существует в форме нуклеосом, причем ядро нуклеосомы содержит приблизительно 150 пар оснований ДНК, ассоциированные с октамером, содержащим по два гистона H2A, H2B, H3 и H4; и линкерную ДНК (переменной длины в зависимости от организма) между ядрами нуклеосом. Молекула гистона H1 обычно ассоциирована с линкерной ДНК. Для целей настоящего описания подразумевается, что термин "хроматин" включает нуклеопротеины клеток всех видов, как прокариотических, так и эукариотических. Хроматин клетки включает как хромосомный, так и эписомный хроматин.

"Хромосома" представляет собой хроматиновый комплекс, полностью или частично содержащий геном клетки. Геном клетки часто характеризуется ее кариотипом, который представляет собой совокупность всех хромосом, составляющих геном клетки. Геном клетки может содержать одну или более хромосом.

"Эписома" представляет собой реплицирующуюся нуклеиновую кислоту, нуклеопротеиновый комплекс или другую структуру, содержащую нуклеиновую кислоту, не входящую в состав хромосомного кариотипа клетки. Примеры эписом включают плазмиды и некоторые вирусные геномы. "Сайт-мишень" или "последовательность-мишень" представляет собой нуклеотидную последовательность, обозначающую фрагмент нуклеиновой кислоты, с которым связывается связывающая молекула при наличии достаточных условий для связывания. Например, последовательность 5'-GAATTC-3' представляет собой сайт-мишень для эндонуклеазы рестрикции Eco RI.

"Экзогенная" молекула представляет собой молекулу, которая обычно не присутствует в клетке, но которую можно внедрить в клетку с помощью одного или более из генетических, биохимических или других способов. "Обычное присутствие в клетке" определяют по отношению к конкретной стадии развития и условиям в окружении клетки. Так, например, молекула, присутствующая только во время эмбрионального развития мышцы, является экзогенной молекулой по отношению к взрослой мышечной клетке. Аналогичным образом, молекула, индуцируемая тепловым шоком, представляет собой экзогенную молекулу по отношению к клетке, не подвергавшейся действию теплового шока. Экзогенная молекула может содержать, например, функциональную версию неправильно функционирующей эндогенной молекулы или неправильно функционирующую версию нормально функционирующей эндогенной молекулы.

Экзогенная молекула может представлять собой, в числе прочего, низкомолекулярное соединение, например, полученное посредством комбинаторного химического процесса, или макромолекулу, например, белок, нуклеиновую кислоту, углевод, липид, гликопротеин, полисахарид, любое модифицированное производное вышеуказанных молекул или любой комплекс, содержащий одну или более из вышеуказанных молекул. Нуклеиновые кислоты включают ДНК или РНК, могут быть одно- или двуцепочечными, линейными, разветвленными или циклическими, и могут быть любой длины. См., например, патенты США № 8703489 и 9255259. Нуклеиновые кислоты включают нуклеиновые кислоты, способные образовывать двойные структуры, а также нуклеиновые кислоты, образующие тройные структуры. См., например, патенты США № 5176996 и 5422251. Белки включают ДНК-связывающие белки, факторы транскрипции, факторы ремоделирования хроматина, белки, связывающие метилированную ДНК, полимеразы, метилазы, деметилазы, ацетилазы, дезацетилазы, киназы, фосфатазы, интегразы, рекомбиназы, лигазы, топоизомеразы, гиразы и хеликазы, но не ограничиваются ими. Экзогенная молекула может представлять собой молекулу того же типа, что и эндогенная молекула, например, экзогенный белок или нуклеиновая кислота. Например, экзогенная нуклеиновая кислота может содержать геном инфекционного вируса, плазмиду или эписому, внедренную в клетку, или хромосому, обычно не присутствующую в клетке. Способы внедрения экзогенных молекул в клетки известны специалистам в данной области техники и включают перенос, опосредованный липидами (т.е. липосомами, содержащими нейтральные и катионные липиды), электропорацию, прямую инъекцию, слияние клеток, бомбардировку частицами, совместное осаждение с фосфатом кальция, перенос, опосредованный ДЭАЭ-декстраном, и перенос, опосредованный вирусным вектором, но не ограничиваются ими. Кроме того, экзогенная молекула может представлять собой молекулу того же типа, что и эндогенная молекула, но может быть получена

не из того вида организмов, из которого получена клетка. Например, нуклеотидную последовательность человека можно внедрить в клетки линии, изначально происходящей от мыши или хомяка.

В противоположность этому, "эндогенная" молекула представляет собой молекулу, обычно присутствующую в конкретной клетке на конкретной стадии развития в конкретных условиях окружающей среды. Например, эндогенная нуклеиновая кислота может включать хромосому, геном митохондрии, хлоропласта или другой органеллы, или природную эписомную нуклеиновую кислоту. Дополнительные эндогенные молекулы могут включать белки, например, факторы транскрипции и ферменты.

"Гибридная" молекула представляет собой молекулу, в которой соединены две или более молекулы-субъединицы, предпочтительно ковалентно. Молекулы-субъединицы могут являться молекулами одного и того же химического типа или молекулами различных химических типов. Примеры гибридных молекул первого типа включают гибридные белки (например, гибриды между ДНК-связывающим доменом ZFP или TALE и одним или более из активирующих доменов) и гибридные нуклеиновые кислоты (например, нуклеиновую кислоту, кодирующую гибридный белок, описанный выше), но не ограничиваются ими. Примеры гибридных молекул второго типа включают гибриды между нуклеиновой кислотой, образующей тройную структуру, и полипептидом, и гибриды между белком, связывающимся с малой бороздкой, и нуклеиновой кислотой, но не ограничиваются ими. Данный термин также включает системы, в которых полинуклеотидный компонент ассоциируется с полипептидным компонентом с образованием функциональной молекулы (например, систему CRISPR/Cas, в которой одиночная направляющая РНК ассоциируется с функциональным доменом в целях модуляции экспрессии гена).

Экспрессия гибридного белка в клетке может быть обусловлена доставкой гибридного белка в клетку или доставкой полинуклеотида, кодирующего гибридный белок, в клетку, где полинуклеотид транскрибируется, а транскрипт транслируется с получением гибридного белка. Экспрессия белка в клетке также может включать транс-сплайсинг, расщепление полипептида и лигирование полипептида. Способы доставки полинуклеотидов и полипептидов в клетки представлены в других частях данного описания.

Термин "ген" для целей настоящего описания включает область.

ДНК, кодирующую продукт гена (см. ниже), а также области ДНК, регулирующие продукцию продукта гена, независимо от того, располагаются ли такие регуляторные последовательности вблизи от кодирующих и/или транскрибируемых последовательностей. Соответственно, термин "ген" включает последовательности промоторов, терминаторы, регуляторные последовательности трансляции, например, сайты связывания рибосом и внутренние сайты связывания рибосом, энхансеры, сайленсеры, инсуляторы, пограничные элементы, сайты инициации репликации, сайты присоединения к матриксу и локус-контролирующие области, но не обязательно ограничивается ими.

Локус "безопасного носительства" представляет собой локус в составе генома, в который можно вставить ген без вредоносных эффектов для клетки-хозяина. Наиболее благоприятен локус безопасного носительства, в котором экспрессия вставленной последовательности гена не нарушается за счет фоновой экспрессии соседних генов. Неограничивающие примеры локусов безопасного носительства, являющиеся специфическими мишенями нуклеаз, включают CCR5, HPRT, AAVS1, Rosa и альбумин. См., например, патенты США № 8771985; 8110379; 7951925; публикации патентов США № 20100218264; 20110265198; 20130137104; 20130122591; 20130177983; 20130177960; 20150056705; и 20150159172).

"Экспрессия генов" относится к преобразованию информации, содержащейся в гене, в продукт гена. Продукт гена может представлять собой непосредственный продукт транскрипции гена (например, мРНК, тРНК, рРНК, антисмысловую РНК, рибозим, структурную РНК или РНК любого другого типа) или белок, продуцированный путем трансляции мРНК. Продукты генов также включают РНК, модифицированные посредством таких процессов, как кэпирование, полиаденилирование, метилирование и редактирование, и белки, модифицированные посредством, например, метилирования, ацетилирования, фосфорилирования, убиквитинилирования, АДФ-рибозилирования, миристилирования и гликозилирования.

"Модуляция" или "модификация" экспрессии генов относится к изменению активности гена. Модуляция экспрессии может включать активацию и репрессию гена, в том числе путем модификации гена посредством связывания экзогенной молекулы (например, рекомбинантного фактора транскрипции), но не ограничивается ими. Модуляция также может достигаться за счет модификации последовательности гена посредством редактирования генома (например, расщепления, изменения, инактивации, случайной мутации). Инактивация гена относится к любому снижению экспрессии гена по сравнению с клеткой, не модифицированной в соответствии с описанием, приведенным в настоящем документе. Так, инактивация гена может быть частичной или полной.

"Требуемая область" представляет собой любую область хроматина клетки, например, ген или не кодирующую последовательность в составе гена или рядом с ним, с которой является желательным связывание экзогенной молекулы. Связывание может осуществляться в целях адресного расщепления ДНК и/или адресной рекомбинации. Требуемая область может находиться, например, в хромосоме, эписоме, геноме органеллы (например, митохондрии, хлоропласта) или геноме инфекционного вируса. Требуемая область может находиться в составе кодирующей области гена, в составе транскрибируемых некоди-

рующих областей, например, лидерных последовательностей, трейлерных последовательностей или интронов выше или ниже кодирующей области. Длина требуемой области может составлять от одной пары нуклеотидов до 2000 пар нуклеотидов, или любое целое число пар нуклеотидов.

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

Термин "функциональная связь" и "функционально связанный" используются на равных основаниях по отношению к расположению двух и более компонентов (например, элементов последовательности) в непосредственной близости друг от друга, при котором компоненты упорядочены таким образом, что оба компонента функционируют нормально, и допускает возможность того, что по меньшей мере один из компонентов может опосредовать функцию, которая оказывает влияние на по меньшей мере один из других компонентов. В качестве иллюстрации, последовательность для регуляции транскрипции, например, промотор, функционально связана с кодирующей последовательностью, если последовательность для регуляции транскрипции контролирует уровень транскрипции кодирующей последовательности в ответ на присутствие или отсутствие одного или более из регуляторных факторов транскрипции. Последовательность для регуляции транскрипции обычно функционально цис-связана с кодирующей последовательностью, но не обязательно располагается в непосредственной близости от нее. Например, энхансер представляет собой последовательность для регуляции транскрипции, функционально связанную с кодирующей последовательностью, даже несмотря на то, что они не являются смежными. По отношению к гибридным полипептидам термин "функционально связанный" может относиться к тому, что каждый компонент при связи с другим компонентом выполняет ту же функцию, которую он выполнял бы в отсутствие такой связи. Например, по отношению к гибриднему полипептиду, в котором ДНК-связывающий домен (например, ZFP, TALE) объединен с доменом активации, причем ДНК-связывающий домен и домен активации функционально связаны, если ДНК-связывающий фрагмент в гибридном полипептиде может связываться со своим сайтом-мишенью и/или сайтом связывания, а домен активации может стимулировать экспрессию гена. В случае гибридного полипептида, в котором ДНК-связывающий домен объединен с расщепляющим доменом, ДНК-связывающий домен и расщепляющий домен функционально связаны, если ДНК-связывающий фрагмент в гибридном полипептиде может связываться со своим сайтом-мишенью, а расщепляющий домен может расщеплять ДНК в непосредственной близости от сайта-мишени. Аналогичным образом, по отношению к гибриднему полипептиду, в котором ДНК-связывающий домен объединен с доменом активации или репрессии, ДНК-связывающий домен и домен активации или репрессии функционально связаны, если ДНК-связывающий фрагмент в гибридном полипептиде может связываться со своим сайтом-мишенью, а домен активации может стимулировать экспрессию гена, или домен репрессии может подавлять экспрессию гена.

"Функциональный фрагмент" белка, полипептида или нуклеиновой кислоты представляет собой белок, полипептид или нуклеиновую кислоту, последовательности которых не идентичны полноразмерному белку, полипептиду или нуклеиновой кислоте, но сохраняют ту же функцию, что и полноразмерный белок, полипептид или нуклеиновая кислота. Функциональный фрагмент может содержать большее, меньшее или такое же количество остатков, что и соответствующая нативная молекула, и/или может содержать одну или более аминокислотную или нуклеотидную замену. Способы определения функции нуклеиновой кислоты (например, кодирующей функции, способности гибридизоваться с другой нуклеиновой кислотой) хорошо известны в данной области техники. Аналогичным образом, хорошо известны способы определения функции белка. Например, ДНК-связывающую функцию полипептида можно определить, например, связыванием на фильтре, смещением электрофоретической подвижности или при анализе иммунопреципитации. Расщепление ДНК можно анализировать с помощью гель-электрофореза. См. Ausubel et al., *supra*. Способность белка взаимодействовать с другим белком можно определить, например, посредством совместной иммунопреципитации, анализа двойных гибридов или комплементарности, как генетического, так и биохимического. См., например, Fields et al. (1989), *Nature*, 340:245-246; патент США № 5585245; и публикация PCT WO 98/44350.

"Вектор" может переносить последовательности генов в клетки-мишени. Обычно термины "векторный конструкт", "экспрессирующий вектор", "экспрессирующий конструкт", "экспрессирующая кассета" и "вектор для переноса генов" означают любой нуклеотидный конструкт, способный управлять экспрессией требуемого гена и переносить последовательности генов в клетки-мишени. Таким образом, данный термин включает носители для клонирования и экспрессии, а также встраиваемые векторы. "Репортерный ген" или "репортерная последовательность" относится к любой последовательности, продуцирующей белковый продукт, который легко измерить, предпочтительно, но не обязательно путем стандартного анализа. Подходящие репортерные гены включают последовательности, кодирующие белки, опосредующие устойчивость к антибиотикам (например, устойчивость к ампициллину, устойчивость к неомицину, устойчивость к G418, устойчивость к пурамицину), последовательности, кодирующие окрашенные или флуоресцентные или люминесцентные белки (например, зеленый флуоресцентный белок, ультрафиолетовый флуоресцентный белок, красный флуоресцентный белок, люциферазу), и белки, опосредующие усиленный рост клеток и/или амплификацию гена (например, дигидрофолатредуктазы), но не ограничиваются ими. Эпитопные маркеры включают, например, одну или более из копий FLAG,

His, мус, Tap, HA или любой обнаружимой аминокислотной последовательности. "Маркеры экспрессии" включают последовательности, кодирующие репортеры, которые могут быть функционально связаны с последовательностью желательного гена в целях мониторинга экспрессии требуемого гена.

Термины "субъект" и "пациент" используются на равных основаниях и относятся к млекопитающим, например, пациентам-людям и приматам, не являющимся людьми, а также экспериментальным животным, например, кроликам, собакам, кошкам, крысам, мышам и другим животным. Соответственно, термины "субъект" или "пациент" в настоящем изобретении означают любого пациента или субъекта-млекопитающего, которому можно ввести экспрессирующие кассеты согласно настоящему изобретению. Субъекты согласно настоящему изобретению включают субъектов с нарушением или субъектов, подвергающихся риску развития нарушения.

Термин "лечащий" и "лечение" в настоящем изобретении относятся к снижению тяжести и/или частоты симптомов, устранению симптомов и/или основной причины, предотвращению возникновения симптомов и/или их основной причины, и улучшению или ликвидации последствий ущерба. Рак и реакция трансплантат против хозяина представляют собой неограничивающие примеры состояний, которые можно лечить с применением композиций и способов, описанных в настоящем изобретении. Таким образом, термины "лечащий" и "лечение" включают

(i) предотвращение возникновения заболевания или состояния у млекопитающего, в частности, если такое млекопитающее предрасположено к указанному состоянию, но диагноз этого состояния у него еще не установлен;

(ii) ингибирование заболевания или состояния, т.е. остановку его развития;

(iii) облегчение заболевания или состояния, т.е. купирование заболевания или состояния; или

(iv) облегчение симптомов, обусловленных заболеванием или состоянием, т.е. облегчение боли без воздействия на основное заболевание или состояние.

В настоящем изобретении термины "заболевание" или "состояние" могут использоваться на равных основаниях или могут различаться в том смысле, что у конкретного нарушения или состояния может не быть известного этиологического фактора (так что этиология еще не изучена), и поэтому оно еще считается не заболеванием, а лишь нежелательным состоянием или синдромом, при котором врачи выявляют более или менее специфический набор симптомов.

"Фармацевтическая композиция" относится к составу соединения согласно настоящему изобретению и среды, общепринятой в данной области техники для доставки биологически активного соединения в организмы млекопитающих, например, людей. Такая среда включает все фармацевтически приемлемых носители, разбавители или вспомогательные вещества. "Эффективное количество" или "терапевтически эффективное количество" относится к такому количеству соединения согласно настоящему изобретению, которое при введении в организм млекопитающего, предпочтительно человека, является достаточным для лечения млекопитающего, предпочтительно человека. Количество композиции согласно настоящему изобретению, составляющее "терапевтически эффективное количество", может зависеть от соединения, состояния и его тяжести, способа введения и возраста млекопитающего, подлежащего лечению, однако его может определить специалист в данной области техники в рабочем порядке, используя свои знания и настоящее описание.

ДНК-связывающие домены.

В настоящем изобретении описаны композиции, содержащие ДНК-связывающий домен, специфически связывающийся с сайтом-мишенью в любом гене, содержащем ген HLA или регулятор HLA, в том числе гене B2M. В композициях и способах, описанных в настоящем изобретении, можно использовать любой ДНК-связывающий домен, включая ДНК-связывающий домен, содержащий мотив цинкового пальца, ДНК-связывающий домен TALE, ДНК-связывающий фрагмент (онРНК) нуклеазы CRISPR/Cas или ДНК-связывающий домен мегануклеазы, но не ограничиваясь ими. ДНК-связывающий домен может связываться с любой последовательностью-мишенью в составе гена, включая последовательность-мишень из 12 или более нуклеотидов, показанную в любой из SEQ ID NO: 6-48, но не ограничиваясь ими.

В некоторых вариантах реализации ДНК-связывающий домен содержит белок с мотивом цинкового пальца. Указанный белок с мотивом цинкового пальца предпочтительно не встречается в природе в том смысле, что он сконструирован с целью связывания 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, все из которых полностью включены в настоящий документ посредством ссылок. В некоторых вариантах реализации ДНК-связывающий домен содержит белок с мотивом цинкового пальца, описанный в публикации патента США № 2012/0060230 (например, табл. 1), полностью включенной в настоящий документ посредством ссылки.

Рекомбинантный связывающий домен с мотивом цинкового пальца может обладать новой специфичностью связывания по сравнению с природным белком, содержащим мотив цинкового пальца. Способы конструирования включают рациональное проектирование и различные типы отбора, но не ограни-

чиваются ими. Рациональное проектирование включает, например, применение баз данных, содержащих триплетные (или квадриплетные) нуклеотидные последовательности и аминокислотные последовательности отдельных мотивов цинкового пальца, в которых каждая триплетная или квадриплетная нуклеотидная последовательность ассоциирована с одной или более из аминокислотных последовательностей мотивов цинкового пальца, связывающих конкретную триплетную или квадриплетную последовательность. См., например, патенты США № 6453242 и 6534261, полностью включенные в настоящий документ посредством ссылок. Типичные способы отбора, включающие фаговый дисплей и системы двойных гибридов, описаны 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 связывающих доменов, содержащих мотивы цинкового пальца, описано, например, в патенте США № 6794136.

Кроме того, как описано в этом и других источниках, домены цинкового пальца и белки с множественными мотивами цинкового пальца можно соединять друг с другом с применением любых подходящих линкерных последовательностей, включая, например, линкеры длиной 5 или более аминокислот. Типичные линкерные последовательности длиной 6 или более аминокислот см. также в патентах США № 6479626; 6903185; и 7153949. Белки, описанные в настоящем изобретении, могут содержать любую комбинацию подходящих линкеров между отдельными мотивами цинкового пальца в белке. Кроме того, усиление специфичности связывания связывающих доменов, содержащих мотивы цинкового пальца, описано, например, в патенте США № 6794136.

Отбор сайтов-мишеней, ZFP и способы проектирования и конструирования гибридных белков (и полинуклеотидов, кодирующих их) известны специалистам в данной области техники и подробно описаны в патентах США № 6140081; 5789538; 6453242; 6534261; 5925523; 6007988; 6013453; 6200759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536; и WO 03/016496. Кроме того, как описано в данном и других источниках, домены цинкового пальца и белки с множественными мотивами цинкового пальца можно соединять друг с другом с применением любых подходящих линкерных последовательностей, включая, например, линкеры длиной 5 или более аминокислот. Типичные линкерные последовательности длиной 6 или более аминокислот см. также в патентах США № 6479626; 6903185; и 7153949. Белки, описанные в настоящем изобретении, могут содержать любую комбинацию подходящих линкеров между отдельными мотивами цинкового пальца в белке. В некоторых вариантах реализации ДНК-связывающий домен представляет собой рекомбинантный белок с мотивом цинкового пальца, связывающийся (специфично по отношению к последовательности) с сайтом-мишенью в гене В2М или регуляторном гене В2М и модулирующий экспрессию гена В2М. В некоторых вариантах реализации белок с мотивом цинкового пальца связывается с сайтом-мишенью в В2М, в то время как в других вариантах реализации мотив цинкового пальца связывается с сайтом-мишенью в В2М.

Обычно ZFP содержат по меньшей мере три "пальца". Некоторые ZFP содержат четыре, пять или шесть "пальцев". ZFP, содержащие три "пальца", обычно распознают сайт-мишень, содержащий 9 или 10 нуклеотидов; ZFP, содержащие четыре "пальца", обычно распознают сайт-мишень, содержащий 12-14 нуклеотидов; в то время как ZFP, содержащие шесть "пальцев", могут распознавать сайт-мишень, содержащий 18-21 нуклеотидов. ZFP также могут представлять собой гибридные белки, содержащие один или более из регуляторных доменов, которые могут являться доменами активации или репрессии транскрипции.

В некоторых вариантах реализации ДНК-связывающий домен может быть получен из нуклеазы. Например, известны распознающие последовательности эндонуклеаз хоминга и мегануклеаз, например I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII и I-TevIII. См. также патент США № 5420032; патент США № 6833252; Belfort et al. (1997), *Nucleic Acids Res.*, 25:3379-3388; Dujon et al. (1989), *Gene*, 82:115-118; Perler et al. (1994), *Nucleic Acids Res.*, 22, 1125-1127; Jasin (1996), *Trends Genet.*, 12:224-228; Gimble et al. (1996), *J. Mol. Biol.*, 263:163-180; Argast et al. (1998), *J. Mol. Biol.*, 280:345-353; и каталог New England Biolabs. Кроме того, специфичность связывания ДНК эндонуклеаз хоминга и мегануклеаз можно модифицировать с целью связывания неприродных сайтов-мишеней. См., например, Chevalier et al. (2002), *Molec. Cell*, 10:895-905; Epinat et al. (2003), *Nucleic Acids Res.*, 31:2952-2962; Ashworth et al. (2006), *Nature*, 441:656-659; Paques et al. (2007), *Current Gene Therapy*, 7:49-66; публикацию патента США № 20070117128.

В других вариантах реализации ДНК-связывающий домен содержит рекомбинантный домен фактора TAL, аналогичный доменам, полученным из возбудителей заболеваний растений *Xanthomonas* (см. Boch et al. (2009), *Science*, 326:1509-1512; и Moscou and Bogdanove (2009), *Science*, 326:1501) и *Ralstonia* (см. Heuer et al. (2007), *Applied and Environmental Microbiology*, 73(13):4379-4384); заявки на патент США № 20110301073 и 20110145940. Известно, что бактерии-возбудители заболеваний растений рода *Xanthomonas* вызывают многие заболевания важных сельскохозяйственных культур. Патогенность *Xanthomonas* зависит от консервативной системы секреции III типа (T3S), позволяющей вводить более 25 различных эффекторных белков в клетку растения. Среди этих вводимых белков находятся эффекторы, подобные активатору транскрипции (TALE), имитирующие активаторы транскрипции растений и манипулирующие транскриптом растения (см. Kay et al. (2007), *Science*, 318:648-651). Эти белки со-

держат ДНК-связывающий домен и домен активации транскрипции. Один из наиболее изученных TALE - AvrBs3 из *Xanthomonas campestris* pv. *Vesicatoria* (см. Bonas et al. (1989), *Mol. Gen. Genet.*, 218:127-136; и WO 2010079430). TALE содержат централизованный домен из tandemных повторов, каждый из которых содержит приблизительно 34 аминокислоты, который играет ключевую роль в специфичности связывания ДНК этими белками. Кроме того, они содержат последовательность внутриядерной локализации и кислый домен активации транскрипции (см. обзор в статье Schornack S. et al. (2006), *J. Plant. Physiol.*, 163(3):256-272). Кроме того, обнаружено, что два гена фитопатогенной бактерии *Ralstonia solanacearum*, обозначаемые как *brg11* и *hpx17*, являются гомологичными семейству AvrBs3 *Xanthomonas* у штамма GM11000 *R. solanacearum* биовара 1 и штамма RS1000 биовара 4 (см. Heuer et al. (2007), *Appl and Envir Micro*, 73(13):4379-4384). Нуклеотидные последовательности этих генов на 98,9% идентичны друг другу, но отличаются делецией 1575 п.о. в повторяющемся домене *hpx17*. Однако продукты обоих генов характеризуются менее чем 40% идентичностью последовательности по отношению к белкам *Xanthomonas* семейства AvrBs3. Специфичность этих TAL-эффекторов зависит от последовательности, обнаруженной в tandemных повторах. Повторяющаяся последовательность содержит приблизительно 102 пары оснований, и повторы обычно на 91-100% гомологичны друг другу (Bonas et al., *ibid*). Полиморфизм повторов обычно имеет место в положениях 12 и 13; имеет место точное соответствие между идентичностью гипервариабельных остатков (области повторяющихся вариабельных двойных остатков или RVD) в положениях 12 и 13 и идентичностью смежных нуклеотидов в последовательности-мишени эффектора TAL (см. Moscou and Bogdanove (2009), *Science*, 326:1501; и Boch et al. (2009), *Science*, 326:1509-1512). Экспериментальное изучение естественного кода для распознавания ДНК в этих TAL-эффекторах позволило определить, например, что последовательность HD в положениях 12 и 13 (повторяющиеся вариабельные двойные остатки или RVD) приводит к связыванию с цитозином (C), NG связывается с T, NI - с A, C, G или T, NN связывается с A или G, а ING связывается с T. Эти ДНК-связывающие повторы собирают в белки с новыми комбинациями и количеством повторов с целью получения искусственных факторов транскрипции, способных взаимодействовать с новыми последовательностями и активировать экспрессию неэндогенного репортерного гена в клетках растений (Boch et al., *ibid*). Рекombinantные TAL-белки связывают с расщепляющим полудоменом FokI, получая гибридную нуклеазу с TAL-эффекторным доменом (TALEN), в том числе TALEN с атипичными RVD. См., например, патент США 8586526.

В некоторых вариантах реализации TALEN содержит расщепляющий домен эндонуклеазы (например, FokI) или расщепляющий полудомен. В других вариантах реализации TALE-нуклеаза представляет собой мега-TAL. Эти нуклеазы мега-TAL представляют собой гибридные белки, содержащие ДНК-связывающий домен TALE и расщепляющий домен мегануклеазы. Расщепляющий домен мегануклеазы активен в виде мономера и не требует димеризации. (См. Boissel et al. (2013), *Nucl. Acid. Res*, 1-13, DOI: 10.1093/nar/gkt1224).

В дополнительных вариантах реализации нуклеаза содержит компактный TALEN. Эти нуклеазы представляют собой однопочечные гибридные белки, в которых ДНК-связывающий домен TALE присоединен к домену нуклеазы TevI. Этот гибридный белок может действовать как никаза, локализованная в области TALE, или создавать двуцепочечный разрыв, в зависимости от местоположения ДНК-связывающего домена TALE по отношению к домену нуклеазы TevI (см. Beurdeley et al. (2013), *Nat. Comm.*, 1-8, DOI: 10.1038/ncomms2782). Кроме того, домен нуклеазы может проявлять ДНК-связывающую активность. Любые TALEN можно применять в комбинации с дополнительными TALEN (например, один или более TALEN (сTALEN или FokI-TALEN) с одним или более мега-TALE).

Кроме того, как описано в данном и других источниках, домены цинкового пальца и белки с множественными мотивами цинкового пальца или TALE можно соединять друг с другом с применением любых подходящих линкерных последовательностей, включая, например, линкеры длиной 5 или более аминокислот. Типичные линкерные последовательности длиной 6 или более аминокислот см. также в патентах США № 6479626; 6903185; и 7153949. Белки, описанные в настоящем изобретении, могут содержать любую комбинацию подходящих линкеров между отдельными мотивами цинкового пальца в белке. Кроме того, усиление специфичности связывания связывающих доменов, содержащих мотивы цинкового пальца, описано, например, в патенте США № 6794136.

В некоторых вариантах реализации ДНК-связывающий домен входит в состав нуклеазной системы CRISPR/Cas, содержащей одиночную направляющую РНК (онРНК), связывающуюся с ДНК. См., например, патент США № 8697359 и публикации патентов США № 20150056705 и 20150159172. Локус CRISPR (короткие палиндромные повторы, регулярно расположенные группами), кодирующий РНК-компонент системы, и локус *cas* (CRISPR-ассоциированный), кодирующий белки (Jansen et al., 2002, *Mol. Microbiol.*, 43:1565-1575; Makarova et al., 2002, *Nucleic. Acids. Res.*, 30: 482-496; Makarova et al., 2006, *Biol. Direct*, 1: 7; Haft et al., 2005, *PLoS Comput. Biol.*, 1:e60), составляют последовательности генов нуклеазной системы CRISPR/Cas. Локусы CRISPR в микробных клетках-хозяевах содержат комбинацию CRISPR-ассоциированных (*Cas*) генов, а также некодирующих РНК-элементов, способных программировать специфичность CRISPR-опосредованного расщепления нуклеиновых кислот.

CRISPR II типа представляет собой одну из наиболее изученных систем и осуществляет адресный двуцепочечный разрыв ДНК в четыре последовательных этапа. Во-первых, с локуса CRISPR транскри-

бируются две некодирующих РНК - массив предшественника crRNA и tracrRNA. Во-вторых, tracrRNA гибридизуется с повторяющимися областями предшественника crRNA и опосредует процессинг предшественника crRNA в зрелые crRNAs, содержащие отдельные спейсерные последовательности. В-третьих, комплекс "зрелая crRNA:tracrRNA" направляет функциональный домен (например, нуклеазу, например, Cas) к ДНК-мишени за счет спаривания оснований по Уотсону-Крику между спейсером crRNA и протоспейсером в ДНК-мишени рядом с мотивом, прилегающим к протоспейсеру (PAM), что является дополнительным требованием для распознавания мишени. Наконец, Cas9 опосредует расщепление ДНК-мишени с образованием двуцепочечного разрыва в протоспейсере. Активность системы CRISPR/Cas включает три этапа: (i) инсерция чужеродных последовательностей в массив CRISPR во избежание атак в будущем в ходе процесса, называемого "адаптацией", (ii) экспрессия значимых белков, а также экспрессия и процессинг массива, с последующей (iii) РНК-опосредованной интерференцией по чужеродной нуклеиновой кислоте. Таким образом, в бактериальной клетке в естественном функционировании системы CRISPR/Cas участвуют несколько так называемых "Cas"-белков, которые играют роль в, например, инсерции чужеродной ДНК и т.д. В некоторых вариантах реализации белок Cas может представлять собой "функциональное производное" природного белка Cas. "Функциональное производное" нативной полипептидной последовательности представляет собой соединение, характеризующееся качественным биологическим свойством, аналогичным свойству нативной полипептидной последовательности. "Функциональные производные" включают фрагменты нативной последовательности и производные нативной полипептидной последовательности и их фрагменты, при условии, что они обладают биологической активностью, аналогичной активности соответствующей нативной полипептидной последовательности. Биологическая активность, рассматриваемая в настоящем изобретении, представляет собой способность функционального производного гидролизовать субстрат ДНК на фрагменты. Термин "производное" охватывает как варианты аминокислотной последовательности полипептида, так и ковалентные модификации и их гибриды, например, производные белки Cas. Подходящие производные полипептида Cas или его фрагмента включают мутантные формы, гибриды, ковалентные модификации белка Cas или его фрагмента, но не ограничиваются ими. Белок Cas, включая белок Cas или его фрагмент, а также производное белка Cas или его фрагмента, можно получить из клетки или синтезировать химически или получить путем объединения этих двух процедур. Клетка может представлять собой клетку, естественным образом продуцирующую белок Cas, или клетку, естественным образом продуцирующую белок Cas и подвергшуюся генной инженерии с целью продуцирования эндогенного белка Cas при повышенном уровне экспрессии или продуцирования белка Cas с экзогенно внедренной нуклеиновой кислоты, причем указанная нуклеиновая кислота кодирует Cas, аналогичный эндогенному Cas или отличающийся от него. В некоторых случаях клетка не продуцирует белок CAS естественным образом и подверглась генной инженерии с целью продуцирования белка Cas. В некоторых вариантах реализации белок Cas представляет собой небольшой ортолог Cas9 для доставки посредством AAV-вектора (Ran et al. (2015), Nature, 510, p. 186).

В некоторых вариантах реализации ДНК-связывающий домен входит в состав системы TtAgo (см. Swarts et al., *ibid*; Sheng et al., *ibid*). У эукариот сайленсинг генов опосредует семейство белков Argonaute (Ago). В данной парадигме Ago связывается с малыми (19-31 нуклеотидов) РНК. Этот сайленсинг-комплекс "белок-РНК" распознает РНК-мишени посредством спаривания оснований по Уотсону-Крику между малой РНК и мишенью и эндонуклеолитически расщепляет РНК-мишень (Vogel (2014), Science, 344:972-973). В противоположность этому, прокариотические белки Ago связываются с малыми одноцепочечными фрагментами ДНК и, вероятно, обнаруживают и удаляют чужеродную (часто вирусную) ДНК (Yuan et al. (2005), Mol. Cell, 19, 405; Olovnikov, et al. (2013), Mol. Cell, 51, 594; Swarts et al., *ibid*). Типичные прокариотические белки Ago включают белки Aquifex aeolicus, Rhodobacter sphaeroides и Thermus thermophilus.

Один из наиболее изученных прокариотических белков Ago представляет собой белок T. thermophilus (TtAgo; Swarts et al. *ibid*). TtAgo ассоциируется с 15- или 13-25-нуклеотидными одноцепочечными фрагментами ДНК с 5'-фосфатными группами. Эта "направляющая ДНК", связываемая TtAgo, вызывает связывание комплекса "белок-ДНК" с комплементарной по Уотсону-Крику последовательностью ДНК в посторонней молекуле ДНК. После того как информация о последовательности в этих направляющих ДНК позволяет идентифицировать ДНК-мишень, комплекс "TtAgo-направляющая ДНК" расщепляет ДНК-мишень. Такой механизм также поддерживается структурой комплекса "TtAgo-направляющая ДНК" при связывании с ДНК-мишенью (G. Sheng et al., *ibid*). Ago Rhodobacter sphaeroides (RsAgo) обладает аналогичными свойствами (Olovnikov et al. *ibid*).

Экзогенные направляющие ДНК или произвольную последовательность ДНК можно загрузить в белок TtAgo (Swarts et al. *ibid*). Поскольку специфичность расщепления TtAgo зависит от направляющей ДНК, комплекс "TtAgo-ДНК", образованный экзогенной, заданной исследователем направляющей ДНК, должен вызывать расщепление комплементарной заданной исследователем ДНК-мишени TtAgo. Таким образом можно создать адресный двуцепочечный разрыв в ДНК. Применение системы "TtAgo-направляющая ДНК" (или ортологичной системы "Ago-направляющая ДНК" из других организмов) обеспечивает адресное расщепление геномной ДНК в клетках. Такое расщепление может быть одноцепочечным.

чечным или двуцепочечным. Для расщепления геномной ДНК млекопитающего предпочтительно было бы использовать версию TtAgo, кодон-оптимизированную для экспрессии в клетках млекопитающих. Кроме того, могла бы быть предпочтительной обработка клеток комплексом "TtAgo-ДНК", образованным *in vitro*, где белок TtAgo объединен с пептидом, проникающим в клетку. Кроме того, могло бы быть предпочтительным применение версии белка TtAgo, модифицированной посредством мутагенеза в целях улучшения активности при 37°C. Расщепление ДНК, опосредованное Ago-РНК, можно применять для воздействия с различными результатами, включая нокаут генов, адресное добавление генов, исправление генов, адресную делецию генов с применением методик, стандартно используемых в области работы с разрывами ДНК. Таким образом, можно применять любой ДНК-связывающий домен.

Гибридные молекулы.

Кроме того, предложены гибридные молекулы, содержащие ДНК-связывающий домен (например, ZFP или TALE, компоненты CRISPR/Cas, например, одиночная направляющая РНК), описанные в настоящем изобретении и ассоциированные с гетерологичным регуляторным (функциональным) доменом (или его функциональным фрагментом). Распространенные домены включают, например, домены факторов транскрипции (активаторов, репрессоров, совместных активаторов, совместных репрессоров), сайленсеры, онкогены (например, *myc*, *jun*, *fos*, *myb*, *max*, *mad*, *rel*, *ets*, *bcl*, *myb*, члены семейства *mos* и т.д.); ферменты репарации ДНК и их ассоциированные факторы и модификаторы; ферменты реаранжировки ДНК и их ассоциированные факторы и модификаторы; белки, ассоциированные с хроматином, и их модификаторы (например, киназы, ацетилазы и дезацетилазы); и ферменты, модифицирующие ДНК (например, метилтрансферазы, топоизомеразы, хеликазы, лигазы, киназы, фосфатазы, полимеразы, эндонуклеазы) и их ассоциированные факторы и модификаторы. Такие гибридные молекулы включают факторы транскрипции, содержащие ДНК-связывающие домены, описанные в настоящем изобретении, и домен регуляции транскрипции, а также нуклеазы, содержащие ДНК-связывающие домены и один или более из доменов нуклеазы. Подходящие домены для активации (домены активации транскрипции) включают домен активации HSV VP16 (см., например, Hagmann et al., *J. Virol.*, 71, 5952-5962 (1997)), ядерные рецепторы гормонов (см., например, Torchia et al., *Curr. Opin. Cell. Biol.*, 10:373-383 (1998)); субъединицу p65 ядерного фактора каппа-B (Bitko & Bank, *J. Virol.*, 72:5610-5618 (1998); и Doyle & Hunt, *Neuroreport*, 8:2937-2942 (1997)); Liu et al., *Cancer Gene Ther.*, 5:3-28 (1998) или искусственные химерные функциональные домены, например, VP64 (Beerli et al. (1998), *Proc. Natl. Acad. Sci., USA*, 95:14623-33) и дегрон (Molinari et al. (1999), *EMBO J.*, 18, 6439-6447). Дополнительные типичные домены активации включают Oct 1, Oct-2A, Sp1, AP-2 и CTF1 (Seipel et al., *EMBO J.*, 11, 4961-4968 (1992), а также p300, CBP, PCAF, SRC1 PvALF, AtHD2A и ERF-2. См., например, Roby et al. (2000), *Mol. Endocrinol.*, 14:329-347; Collingwood et al. (1999), *J. Mol. Endocrinol.*, 23:255-275; Leo et al. (2000), *Gene*, 245:1-11; Manteuffel-Cymborowska (1999), *Acta Biochim. Pol.*, 46:77-89; McKenna et al. (1999), *J. Steroid Biochem. Mol. Biol.*, 69:3-12; Malik et al. (2000), *Trends Biochem. Sci.*, 25:277-283; и Lemon et al. (1999), *Curr. Opin. Genet. Dev.* 9:499-504. Дополнительные типичные домены активации включают OsGAI, HALF-1, C1, AP1, ARF-5,-6,-7 и -8, CPRF1, CPRF4, MYC-RP/GP и TRAB1, но не ограничиваются ими. См., например, Ogawa et al. (2000), *Gene*, 245:21-29; Okanami et al. (1996), *Genes Cells*, 1:87-99; Goff et al. (1991), *Genes Dev.*, 5:298-309; Cho et al. (1999), *Plant Mol. Biol.*, 40:419-429; Ulmason et al. (1999), *Proc. Natl. Acad. Sci., USA*, 96:5844-5849; Sprenger-Haussels et al. (2000), *Plant J.*, 22:1-8; Gong et al. (1999), *Plant Mol. Biol.*, 41:33-44; и Hobo et al. (1999), *Proc. Natl. Acad. Sci., USA*, 96:15,348-15,353.

Для специалиста в данной области техники должно быть ясно, что при образовании гибридного белка (или нуклеиновой кислоты, кодирующей его) между ДНК-связывающим доменом и функциональным доменом, в качестве функционального домена подходит домен активации или молекула, взаимодействующая с доменом активации. В качестве домена активации гибридного белка можно применять практически любую молекулу, способную рекрутировать активирующий комплекс и/или характеризующуюся активирующей активностью (например, ацелированием гистонов) по отношению к гену-мишени. Домены инсуляторов, домены локализации и белки ремоделирования хроматина, например, ISWI-содержащие домены и/или белки с метил-связывающим доменом, подходящие для применения в качестве функциональных доменов гибридных молекул, описаны, например, в патенте США № 7053264.

Типичные домены репрессии включают KRAB A/B, KOX, TGF-бета-индуцибельный ранний ген (TIEG), *v-erbA*, SID, MBD2, MBD3, члены семейства DNMT (например, DNMT1, DNMT3A, DNMT3B), Rb и MeCP2, но не ограничиваются ими. См., например, Bird et al. (1999), *Cell*, 99:451-454; Tyler et al. (1999) *Cell* 99:443-446; Knoepfler et al. (1999) *Cell* 99:447-450; и Robertson et al. (2000) *Nature Genet.* 25:338-342. Дополнительные типичные домены репрессии включают ROM2 и AtHD2A, но не ограничиваются ими. См., например, Chem et al. (1996) *Plant Cell* 8:305-321; и Wu et al. (2000) *Plant J.* 22:19-27.

Гибридные молекулы конструируют с применением способов клонирования и биохимического конъюгирования, хорошо известных специалистам в данной области. Гибридные молекулы содержат ДНК-связывающий домен (например, ZFP, TALE, онРНК) и функциональный домен (например, домен активации или репрессии транскрипции). Гибридные молекулы также необязательно содержат сигналы внутриядерной локализации (например, сигналы среднего Т-антигена SV40) и эпителиальные маркеры (например, FLAG и гемагглютинин). Гибридные белки (и нуклеиновые кислоты, кодирующие их) спроекти-

рованы таким образом, что рамка считывания трансляции сохраняется для всех компонентов гибрида.

Гибриды между полипептидным компонентом функционального домена (или его функциональным фрагментом), с одной стороны, и небелковым ДНК-связывающим доменом (например, антибиотиком, интеркалирующим агентом, агентом, связывающимся с малой бороздкой, нуклеиновой кислотой), с другой стороны, конструируют с применением способов биохимического конъюгирования, известны специалистам в данной области техники. См., например, каталог Pierce Chemical Company (Рокфорд, штат Иллинойс, США).

Описаны способы и композиции для получения гибридов между агентом, связывающимся с малой бороздкой, и полипептидом. Mapp et al. (2000) Proc. Natl. Acad. Sci. USA 97:3930-3935. Кроме того, односторонние направляющие РНК систем CRISPR/Cas ассоциируют с функциональным доменом с образованием активных регуляторов транскрипции и нуклеаз.

В некоторых вариантах реализации сайт-мишень присутствует в доступной области хроматина клетки. Доступные области можно определить согласно описанию, например, в патентах США № 7217509 и 7923542. Если сайт-мишень не присутствует в доступной области хроматина клетки, можно получить одну или более из доступных областей, как описано в патентах США № 7785792 и 8071370. В дополнительных вариантах реализации ДНК-связывающий домен гибридной молекулы может связываться с хроматином клетки независимо от того, находится ли сайт-мишень в доступной области или нет. Например, такие ДНК-связывающие домены могут связываться с линкерной ДНК и/или нуклеосомной ДНК. Примеры "пионерных" ДНК-связывающих доменов этого типа находятся в некоторых рецепторах стероидов и ядерном факторе 3 гепатоцитов (HNF3) (Cordingley et al. (1987) Cell 48:261-270; Pina et al. (1990) Cell 60:719-731; и Cirillo et al. (1998) EMBO J. 17:244-254). Можно получить состав гибридной молекулы с фармацевтически приемлемым носителем, известным специалистам в данной области техники. См., например, Remington's Pharmaceutical Sciences, 17th ed., 1985; и патенты США № 6453242 и 6534261.

Функциональный компонент/домен гибридной молекулы можно выбрать из любого из множества различных компонентов, способных влиять на транскрипцию гена при связывании гибридной молекулы с последовательностью-мишенью за счет своего ДНК-связывающего домена. Следовательно, функциональный компонент может включать различные домены факторов транскрипции, например, активаторы, репрессоры, совместные активаторы, совместные репрессоры и сайленсеры, но не ограничивается ими.

Дополнительные типичные функциональные домены описаны, например, в патентах США № 6534261 и 6933113.

Кроме того, можно выбрать функциональные домены, регулируемые экзогенными низкомолекулярными соединениями или лигандами. Например, можно использовать технологию RheoSwitch®, при которой функциональный домен принимает активную конформацию только в присутствии внешнего лиганда RheoChem™ (см., например, US 20090136465). Таким образом, ZFP может быть функционально связан с регулируемым функциональным доменом, причем активность полученного ZFP-TF контролируется внешним лигандом.

Нуклеазы.

В некоторых вариантах реализации гибридная молекула содержит ДНК-связывающий домен, ассоциированный с расщепляющим (нуклеазным) доменом. Фактически, модификации гена можно достичь с помощью нуклеазы, например, рекомбинантной нуклеазы. Технология рекомбинантных нуклеаз основана на конструировании природных ДНК-связывающих белков. Например, описано конструирование эндонуклеаз хоминга с адаптированной специфичностью связывания ДНК. Chames et al. (2005) Nucleic Acids Res 33(20):e178; Arnould et al. (2006) J. Mol. Biol. 355:443-458. Кроме того, описано конструирование ZFP. См., например, патенты США № 6534261; 6607882; 6824978; 6979539; 6933113; 7163824 и 7013219.

Кроме того, ZFP и/или TALE можно объединять с доменами нуклеаз с образованием ZFN и TALEN - функциональной единицы, способной распознавать заданную нуклеотидную мишень за счет своего рекомбинантного ДНК-связывающего домена (ZFP или TALE) и вызывать расщепление ДНК вблизи от сайта связывания ДНК за счет нуклеазной активности. Таким образом, способы и композиции, описанные в настоящем изобретении, можно широко использовать, и они могут включать любую требуемую нуклеазу. Неограничивающие примеры нуклеаз включают мегануклеазы, TALEN и нуклеазы с мотивом цинкового пальца. Нуклеазы могут содержать гетерологичные домены связывания и расщепления ДНК (например, нуклеазы с мотивом цинкового пальца; ДНК-связывающие домены мегануклеаз с доменами гетерологичного расщепления) или, в качестве альтернативы, можно модифицировать ДНК-связывающий домен природной нуклеазы с целью связывания выбранного сайта-мишени (например, мегануклеазу, модифицированную с целью связывания с сайтом, отличающимся от ее исходного сайта связывания).

Любая из нуклеаз, описанных в настоящем изобретении, может содержать рекомбинантный ДНК-связывающий домен TALE и домен нуклеазы (например, домен эндонуклеазы и/или мегануклеазы), также называемый нуклеазой TALEN. Опубликованы способы и композиции для конструирования этих белков TALEN для надежного сайт-специфического взаимодействия с последовательностью-мишенью, выбранной пользователем (см. патент США № 8586526). В некоторых вариантах реализации TALEN

содержит расщепляющий домен эндонуклеазы (например, FokI) или расщепляющий полудомен. В других вариантах реализации TALE-нуклеаза представляет собой мега-TAL. Эти нуклеазы мега-TAL представляют собой гибридные белки, содержащие ДНК-связывающий домен TALE и расщепляющий домен мегануклеазы. Расщепляющий домен мегануклеазы активен в виде мономера и не требует димеризации. (См. Voissel et al. (2013) *Nucl Acid Res.*: 1-13, DOI: 10.1093/nar/gkt1224). Кроме того, домен нуклеазы может проявлять ДНК-связывающую активность.

В дополнительных вариантах реализации нуклеаза содержит компактный TALEN (сTALEN). Эти нуклеазы представляют собой одноцепочечные гибридные белки, в которых ДНК-связывающий домен TALE присоединен к домену нуклеазы *TevI*. Этот гибридный белок может действовать как никаза, локализованная в области TALE, или создавать двуцепочечный разрыв, в зависимости от местоположения ДНК-связывающего домена TALE по отношению к домену нуклеазы *TevI* (см. Beurdeley et al. (2013) *Nat Comm*: 1-8 DOI: 10.1038/ncomms2782). Любые TALEN можно применять в комбинации с дополнительными TALEN (например, один или более TALEN (сTALEN или FokI-TALEN) с одним или более мега-TAL) или другими ДНК-расщепляющими ферментами.

В некоторых вариантах реализации нуклеаза включает мегануклеазу (эндонуклеазу хоминга) или ее фрагмент, проявляющий расщепляющую активность. Природные мегануклеазы распознают сайт расщепления длиной 15-40 пар оснований и обычно подразделяются на четыре семейства: семейство LAGLIDADG, семейство GIY-YIG, семейство His-Cyst box и семейство HNH. Типичные эндонуклеазы хоминга включают I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII и I-TevIII. Известны распознаваемые ими последовательности. См. также патент США № 5420032; патент США № 6833252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 и каталог New England Biolabs.

ДНК-связывающие домены природных мегануклеаз, в первую очередь семейства LAGLIDADG, применяют для сайт-специфической модификации генома растений, дрожжей, *Drosophila*, клеток млекопитающих и мышей, однако этот подход ограничен модификацией гомологичных генов с консервативной последовательностью, распознаваемой мегануклеазой (Monet et al. (1999), *Biochem. Biophysics. Res. Common.* 255: 88-93) или предварительно модифицированными геномами, в которые внедрили распознаваемую последовательность (Route et al. (1994), *Mol. Cell. Biol.* 14: 8096-106; Chilton et al. (2003), *Plant Physiology.* 133: 956-65; Puchta et al. (1996), *Proc. Natl. Acad. Sci. USA* 93: 5055-60; Rong et al. (2002), *Genes Dev.* 16: 1568-81; Gouble et al. (2006), *J. Gene Med.* 8(5):616-622). Соответственно, предпринимаются попытки конструирования мегануклеаз, характеризующихся новой специфичностью связывания с сайтами, имеющими значение для медицины или биотехнологии (Porteus et al. (2005), *Nat. Biotechnol.* 23: 967-73; Sussman et al. (2004), *J. Mol. Biol.* 342: 31-41; Epinat et al. (2003), *Nucleic Acids Res.* 31: 2952-62; Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66; публикации патентов США № 20070117128; 20060206949; 20060153826; 20060078552; и 20040002092). Кроме того, природные или рекомбинантные ДНК-связывающие домены мегануклеаз можно функционально связать с расщепляющим доменом гетерологичной нуклеазы (например, FokI), и/или расщепляющие домены мегануклеаз можно функционально связать с гетерологичным ДНК-связывающим доменом (например, ZFN или TALE).

В других вариантах реализации нуклеаза представляет собой нуклеазу с мотивом цинкового пальца (ZFN) или гибрид ДНК-связывающего домена TALE и нуклеазы (TALEN). ZFN и TALEN содержат ДНК-связывающий домен (ДНК-связывающий домен белка с мотивом цинкового пальца или TALE), модифицированный с целью связывания сайта-мишени в выбранном гене, и расщепляющий домен или расщепляющий полудомен (например, рестриктазы и/или мегануклеазы, описанных в настоящем изобретении). Как подробно описано выше, можно сконструировать связывающие домены цинкового пальца и ДНК-связывающие домены TALE, связывающие выбранную последовательность. 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. Рекомбинантный связывающий домен с мотивом цинкового пальца или белок TALE может обладать новой специфичностью связывания по сравнению с природным белком. Способы конструирования включают рациональное проектирование и различные типы отбора, но не ограничиваются ими. Рациональное проектирование включает, например, применение баз данных, содержащих триплетные (или квадриплетные) нуклеотидные последовательности и аминокислотные последовательности отдельных мотивов цинкового пальца или TALE, в которых каждая триплетная или квадриплетная нуклеотидная последовательность ассоциирована с одной или более из аминокислотных последовательностей мотивов цинкового пальца или повторяющихся единиц TALE, связывающих конкретную триплетную или квадриплетную последовательность. См., например, патенты США 6453242 и 6534261, полностью включенные в настоящий документ посредством ссылок. Отбор сайтов-мишеней и способы проектирования и конструирования гибридных белков (и полинуклеотидов, кодирующих их)

известны специалистам в данной области техники и подробно описаны в патентах США № 7888121 и 8409861, полностью включенных в настоящий документ посредством ссылок.

Кроме того, как описано в данном и других источниках, домены цинкового пальца, TALE и/или белки с множественными мотивами цинкового пальца можно соединять друг с другом с применением любых подходящих линкерных последовательностей, включая, например, линкеры длиной 5 или более аминокислот. Типичные линкерные последовательности длиной 6 или более аминокислот см., например, в патентах США № 6479626; 6903185; и 7153949. Белки, описанные в настоящем изобретении, могут содержать любую комбинацию подходящих линкеров между отдельными мотивами цинкового пальца в белке. См., например, патент США № 8772453.

Таким образом, такие нуклеазы, как ZFN, TALEN и/или мегануклеазы могут содержать любой ДНК-связывающий домен и любой нуклеазный (расщепляющий) домен (расщепляющий домен, расщепляющий полудомен). Как отмечено выше, расщепляющий домен может представлять собой гетерологичный домен по отношению к ДНК-связывающему домену, например, ДНК-связывающий домен с мотивом цинкового пальца или TAL-эффектора и расщепляющий домен нуклеазы, или ДНК-связывающий домен мегануклеазы и расщепляющий домен другой нуклеазы. Гетерологичные расщепляющие домены можно получить из любой эндонуклеазы или экзонуклеазы. Типичные эндонуклеазы, из которых можно получить расщепляющий домен, включают эндонуклеазы рестрикции и эндонуклеазы хоминга, но не ограничиваются ими. См., например, каталог 2002-2003, New England Biolabs, Beverly, MA; и Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388. Известны дополнительные ферменты, расщепляющие ДНК (например, нуклеаза S1; нуклеаза маша; панкреатическая ДНКаза I; нуклеаза микрококков; эндонуклеаза НО дрожжей; кроме того, см. Linn et al. (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Один или более из этих белков (или их функциональных фрагментов) можно использовать в качестве источника расщепляющих доменов и расщепляющих полудоменов.

Аналогичным образом, расщепляющий полудомен можно получить из любой нуклеазы или ее фрагмента, описанных выше, которые требуют димеризации для проявления расщепляющей активности. В общем случае, если гибридные белки содержат расщепляющие полудомены, для расщепления требуется два гибридных белка. В качестве альтернативы, можно использовать единственный белок, содержащий два расщепляющих полудомена. Эти два расщепляющих полудомена могут происходить от одной и той же эндонуклеазы (или ее функциональных фрагментов), или каждый расщепляющий полудомен может происходить от своей эндонуклеазы (или ее функциональных фрагментов). Кроме того, сайты-мишени двух гибридных белков предпочтительно расположены по отношению друг к другу таким образом, что связывание двух гибридных белков с соответствующими сайтами-мишенями помещает расщепляющие полудомены в такую пространственную ориентацию по отношению друг к другу, которая позволяет расщепляющим полудоменам образовывать функциональный расщепляющий домен, например, путем димеризации. Так, в некоторых вариантах реализации границы сайтов-мишеней разделены 5-8 нуклеотидами или 15-18 нуклеотидами. В то же время между двумя сайтами-мишенями может находиться любое число нуклеотидов или пар нуклеотидов (например, от 2 до 50 пар нуклеотидов или более). В общем случае сайт расщепления находится между сайтами-мишенями, однако сайт расщепления может находиться на расстоянии 1 или более т.п.о., в том числе 1-50 пар оснований (или любого значения между ними, в том числе 1-5, 1-10 и 1-20 пар оснований), 1-100 пар оснований (или любого значения между ними), 100-500 пар оснований (или любого значения между ними), от 500 до 1000 пар оснований (или любого значения между ними) или даже более чем 1 т.п.о.

Эндонуклеазы рестрикции (ферменты рестрикции) присутствуют у многих видов организмов и могут связываться с ДНК специфичным по отношению к последовательности образом (в распознаваемом сайте) и расщеплять ДНК в распознаваемом сайте или вблизи от него. Некоторые ферменты рестрикции (например, типа IIS) расщепляют ДНК в сайтах, удаленных от распознаваемого сайта, и содержат отдельные связывающие и расщепляющие домены. Например, фермент Fok I типа IIS катализирует двуцепочечное расщепление ДНК на расстоянии 9 нуклеотидов от распознаваемого сайта на одной цепи и 13 нуклеотидов от распознаваемого сайта 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.* 269:31978-31982. Таким образом, в одном варианте реализации гибридные белки содержат расщепляющий домен (или расщепляющий полудомен) по меньшей мере одного фермента рестрикции типа IIS и один или более из связывающих доменов с мотивом цинкового пальца, которые могут являться рекомбинантными или не рекомбинантными.

Типичный фермент рестрикции типа IIS, расщепляющий домен которого можно отделить от связывающего домена, представляет собой FokI. Данный конкретный фермент активен в виде димера. Vitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 10,570-10,575. Соответственно, для целей настоящего описания фрагмент фермента FokI, используемый в описанных гибридных белках, считается расщепляющим полудоменом. Таким образом, для адресного двуцепочечного расщепления и/или адресной замены последовательностей в клетке с использованием гибридов "мотив цинкового пальца-FokI" с целью восстановления активного расщепляющего домена можно использовать два гибридных белка, каждый из которых

содержит расщепляющий полудомен FokI. Кроме того, в качестве альтернативы можно использовать единственную полипептидную молекулу, содержащую связывающий домен с мотивом цинкового пальца, и два расщепляющих полудомена FokI. Параметры адресного расщепления и адресной модификации последовательности с использованием гибридов "мотив цинкового пальца-FokI" представлены в других частях настоящего описания.

Расщепляющий домен или расщепляющий полудомен могут представлять собой любой фрагмент белка, сохраняющий расщепляющую активность или способность к мультимеризации (например, димеризации) с образованием функционального расщепляющего домена.

Типичные ферменты рестрикции типа IIS описаны в международной публикации WO 07/014275, полностью включенной в настоящий документ посредством ссылки. Дополнительные ферменты рестрикции также содержат разделяемые связывающий и расщепляющий домены, и их можно рассматривать в настоящем описании. См., например, Roberts et al. (2003) *Nucleic Acids Res.* 31:418-420.

В некоторых вариантах реализации расщепляющий домен содержит один или более из рекомбинантных расщепляющих полудоменов (также называемых мутантными димеризующимися доменами), что минимизирует или предотвращает гомодимеризацию, как описано, например, в патентах США № 7914796; 8034598 и 8623618; и публикации патента США № 20110201055, описания которых полностью включены в настоящий документ посредством ссылок. Все аминокислотные остатки в положениях 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 и 538 FokI являются мишенями для влияния на димеризацию расщепляющих полудоменов Fok I. Типичные рекомбинантные расщепляющие полудомены FokI, образующие облигатные гетеродимеры, включают пару, в которой первый расщепляющий полудомен содержит мутации в аминокислотных остатках в положениях 490 и 538 FokI, а второй расщепляющий полудомен содержит мутации в аминокислотных остатках 486 и 499.

Так, в одном варианте реализации, мутация в положении 490 замещает Glu (E) остатком Lys (K); мутация в положении 538 замещает Iso (I) остатком Lys (K); мутация в положении 486 замещает Gln (Q) остатком Glu (E); а мутация в положении 499 замещает Iso (I) остатком Lys (K). Конкретнее, рекомбинантные расщепляющие полудомены, описанные в настоящем изобретении, получили путем мутирования положений 490 (E→K) и 538 (I→K) в одном расщепляющем полудомене с получением рекомбинантного расщепляющего полудомена, обозначаемого как "E490K:1538K", и путем мутирования положений 486 (Q→E) и 499 (I→L) в другом расщепляющем полудомене с получением рекомбинантного расщепляющего полудомена, обозначаемого как "Q486E:I499L". Рекомбинантные расщепляющие полудомены, описанные в настоящем изобретении, являются облигатными гетеродимерными мутантами, в которых минимизировано или устранено аномальное расщепление. См., например, патенты США 7914796 и 8034598, описания которых полностью включены в настоящую заявку посредством ссылок для всех целей. В некоторых вариантах реализации рекомбинантный расщепляющий полудомен содержит мутации в положениях 486, 499 и 496 (пронумерованных по отношению к FokI дикого типа), например, мутации, замещающие остаток Gln (Q) дикого типа в положении 486 остатком Glu (E), остаток Iso (I) дикого типа в положении 499 остатком Leu (L) и остаток Asn (N) дикого типа в положении 496 остатком Asp (D) или Glu (E) (также называемые доменами "ELD" и "ELE", соответственно). В других вариантах реализации рекомбинантный расщепляющий полудомен содержит мутации в положениях 490, 538 и 537 (пронумерованных по отношению к FokI дикого типа), например, мутации, замещающие остаток Glu (E) дикого типа в положении 490 остатком Lys (K) и остаток His (H) дикого типа в положении 537 остатком Lys (K) или Arg (R) (также называемые доменами "KKK" и "KKR", соответственно). В других вариантах реализации рекомбинантный расщепляющий полудомен содержит мутации в положениях 490 и 537 (пронумерованных по отношению к FokI дикого типа), например, мутации, замещающие остаток Glu (E) дикого типа в положении 490 остатком Lys (K), а остаток His (H) дикого типа в положении 537 - остатком Lys (K) или Arg (R) (также называемые доменами "KIK" и "KIR", соответственно). См., например, патенты США № 7914796; 8034598 и 8623618, описания которых полностью включены в настоящую заявку посредством ссылок для всех целей. В других вариантах реализации рекомбинантный расщепляющий полудомен содержит "Sharkey" и/или мутации "Sharkey" (см. Guo et al. (2010) *J. Mol. Biol.* 400(1):96-107).

В качестве альтернативы, нуклеазы можно собирать *in vivo* в сайте-мишени нуклеиновой кислоты с использованием так называемой технологии "разделенного фермента" (см., например, публикацию патента США № 20090068164). Компоненты таких разделенных ферментов можно экспрессировать на отдельных экспрессирующих конструктах или соединить в одной открытой рамке считывания, где отдельные компоненты разделены, например, саморасщепляющимся пептидом 2A или последовательностью IRES. Компоненты могут представлять собой связывающие домены с мотивами цинкового пальца или домены НК-связывающего домена мегануклеазы. Нуклеазы (например, ZFN и/или TALEN) можно подвергать скринингу перед использованием, например, в системе на основе хромосом дрожжей, как описано в патенте США № 8563314.

В некоторых вариантах реализации нуклеаза содержит систему CRISPR/Cas. Локус CRISPR (короткие палиндромные повторы, регулярно расположенные группами), кодирующий РНК-компонент систе-

мы, и локус Cas (CRISPR-ассоциированный), кодирующий белки (Jansen et al., 2002. Mol. Microbiol. 43: 1565-1575; Makarova et al., 2002. Nucleic Acids Res. 30: 482-496; Makarova et al., 2006. Biol. Direct 1: 7; Haft et al., 2005. PLoS Comput. Biol. 1: e60), составляют последовательности генов нуклеазной системы CRISPR/Cas. Локусы CRISPR в микробных клетках-хозяевах содержат комбинацию CRISPR-ассоциированных (Cas) генов, а также некодирующих РНК-элементов, способных программировать специфичность CRISPR-опосредованного расщепления нуклеиновых кислот.

CRISPR II типа представляет собой одну из наиболее изученных систем и осуществляет адресный двуцепочечный разрыв ДНК в четыре последовательных этапа. Во-первых, с локуса CRISPR транскрибируются две некодирующие РНК - массив предшественника crRNA и tracrRNA. Во-вторых, tracrRNA гибридизуется с повторяющимися областями предшественника crRNA и опосредует процессинг предшественника crRNA в зрелые crRNAs, содержащие отдельные спейсерные последовательности. В-третьих, комплекс "зрелая crRNA:tracrRNA" направляет Cas9 к ДНК-мишени за счет спаривания оснований по Уотсону-Крику между спейсером crRNA и протоспейсером в ДНК-мишени рядом с мотивом, прилегающим к протоспейсеру (PAM), что является дополнительным требованием для распознавания мишени. Наконец, Cas9 опосредует расщепление ДНК-мишени с образованием двуцепочечного разрыва в протоспейсере. Активность системы CRISPR/Cas включает три этапа: (i) инсерция чужеродных последовательностей в массив CRISPR во избежание атак в будущем в ходе процесса, называемого "адаптацией", (ii) экспрессия значимых белков, а также экспрессия и процессинг массива, с последующей (iii) РНК-опосредованной интерференцией по чужеродной нуклеиновой кислоте. Таким образом, в бактериальной клетке в естественном функционировании системы CRISPR/Cas участвуют несколько так называемых "Cas"-белков, которые играют роль в, например, инсерции чужеродной ДНК и т.д. В некоторых вариантах реализации используют систему CRISPR-Cpf1. Система CRISPR-Cpf1, выявленная в *Francisella* spp, представляет собой систему CRISPR-Cas 2 класса, опосредующую надежную РНК-интерференцию в клетках человека. Хотя Cpf1 и Cas9 являются функционально консервативными, они различаются с точки зрения многих аспектов, в том числе направляющих РНК и субстратной специфичности (см. Fagerlund et al. (2015) Genom Bio 16:251). Основное различие между белками Cas9 и Cpf1 состоит в том, что Cpf1 не использует tracrRNA, и поэтому требует только crRNA. Длина crRNA FnCpf1 составляет 42-44 нуклеотида (19-нуклеотидный повтор и 23-25-нуклеотидный спейсер); они содержат одиночную структуру "стебель-петля", устойчивую к изменениям последовательности, позволяющим сохранять вторичную структуру. Кроме того, crRNA Cpf1 значительно короче, чем ~100-нуклеотидные рекомбинантные онРНК, необходимые для Cas9, и требования PAM к FnCpf1 представляют собой 5'-TTN-3' и 5'-CTA-3' на замещаемой цепи. Хотя как Cas9, так и Cpf1 осуществляют двуцепочечные разрывы в ДНК-мишени, Cas9 использует свои RuvC- и HNH-подобные домены для получения разрезов с тупыми концами в последовательности-затравке направляющей РНК, в то время как Cpf1 использует RuvC-подобный домен для получения липких концов за пределами затравки. Поскольку Cpf1 получает липкие концы за пределами критической области затравки, в сайте-мишени не происходит нарушения NHEJ, что обеспечивает возможность продолжения расщепления того же сайта с помощью Cpf1 до осуществления желательного события HDR-рекомбинации. Таким образом, следует понимать, что в способах и композициях, описанных в настоящем изобретении, термин "Cas" включает как белок Cas9, так и белок Cpf1. Таким образом, в настоящем изобретении термин "система CRISPR/Cas" относится как к системе CRISPR/Cas, так и к системе CRISPR/Cpf1, включая как нуклеазные системы, так и системы факторов транскрипции.

В некоторых вариантах реализации белок Cas может представлять собой "функциональное производное" природного белка Cas. "Функциональное производное" нативной полипептидной последовательности представляет собой соединение, характеризующееся качественным биологическим свойством, аналогичным свойству нативной полипептидной последовательности. "Функциональные производные" включают фрагменты нативной последовательности и производные нативной полипептидной последовательности и их фрагменты, при условии, что они обладают биологической активностью, аналогичной активности соответствующей нативной полипептидной последовательности. Биологическая активность, рассматриваемая в настоящем изобретении, представляет собой способность функционального производного гидролизовать субстрат ДНК на фрагменты. Термин "производное" охватывает как варианты аминокислотной последовательности полипептида, так и ковалентные модификации и их гибриды. Подходящие производные полипептида Cas или его фрагмента включают мутантные формы, гибриды, ковалентные модификации белка Cas или его фрагмента, но не ограничиваются ими. Белок Cas, включая белок Cas или его фрагмент, а также производное белка Cas или его фрагмента, можно получить из клетки или синтезировать химически или получить путем объединения этих двух процедур. Клетка может представлять собой клетку, естественным образом продуцирующую белок Cas, или клетку, естественным образом продуцирующую белок Cas и подвергшуюся генной инженерии с целью продуцирования эндогенного белка Cas при повышенном уровне экспрессии или продуцирования белка Cas с экзогенно введенной нуклеиновой кислоты, причем указанная нуклеиновая кислота кодирует Cas, аналогичный эндогенному Cas или отличающийся от него. В некоторых случаях клетка не продуцирует белка CAS естественным образом и подверглась генной инженерии с целью продуцирования белка Cas.

Типичные нуклеазные системы CRISPR/Cas, специфичные по отношению к генам TCR и другим

генам, описаны, например, в публикации патента США № 20150056705.

Нуклеаза(ы) могут выполнять один или более из двуцепочечных и/или одноцепочечных разрезов в сайте-мишени. В некоторых вариантах реализации нуклеаза содержит каталитически неактивный расщепляющий домен (например, белок FokI и/или Cas). См., например, патент США № 9200266; 8703489 и Guillinger et al. (2014) *Nature Biotech.* 32(6):577-582. Каталитически неактивный расщепляющий домен в комбинации с каталитически активным доменом может действовать в качестве никазы, выполняя одноцепочечный разрез. Таким образом, можно использовать две никазы в комбинации для получения двуцепочечного разреза в конкретной области. Кроме того, в данной области техники известны дополнительные никазы, например, McCaffery et al. (2016) *Nucleic Acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Epub 2015 Oct 19.

Доставка.

Белки (например, факторы транскрипции, нуклеазы, молекулы TCR и CAR), полинуклеотиды или композиции, содержащие белки и/или полинуклеотиды, описанные в настоящем изобретении, можно доставлять в клетку-мишень с помощью любых подходящих средств, включающих, например, инъекцию белковых и/или мРНК-компонентов.

Подходящие клетки включают эукариотические и прокариотические клетки и/или линии клеток, но не ограничиваются ими. Неограничивающие примеры таких клеток или линий клеток, полученных из таких клеток, включают Т-клетки, COS, клетки CHO (например, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHO-K1SV), VERO, MDCK, WI38, V79, B14AF28-G3, ВНК, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (например, HEK293-F, HEK293-H, HEK293-T) и perC6, а также клетки насекомых, например, *Spodoptera fugiperda* (Sf), или клетки грибов, например, *Saccharomyces*, *Pichia* и *Schizosaccharomyces*. В некоторых вариантах реализации линия клеток представляет собой линию клеток CHO-K1, MDCK или HEK293. Подходящие клетки также включают стволовые клетки, например, эмбриональные стволовые клетки, в том числе плюрипотентные стволовые клетки (ИПС-клетки), гемопоэтические стволовые клетки, нервные стволовые клетки и мезенхимальные стволовые клетки.

Способы доставки белков, содержащих ДНК-связывающие домены, описанные в настоящем изобретении, описаны, например, в патентах США № 6453242; 6503717; 6534261; 6599692; 6607882; 6689558; 6824978; 6933113; 6979539; 7013219; и 7163824, описание которых полностью включено в настоящий документ посредством ссылки.

Кроме того, можно выполнять доставку ДНК-связывающих доменов и гибридных белков, содержащих эти ДНК-связывающие домены, описанных в настоящем изобретении, с помощью векторов, содержащих последовательности, кодирующие один или более из ДНК-связывающих белков. Кроме того, с помощью этих векторов можно выполнять доставку дополнительных нуклеиновых кислот (например, донорных). Можно использовать любые векторные системы, включая плазмидные векторы, ретровирусные векторы, лентивирусные векторы, аденовирусные векторы, поксвирусные векторы, герпесвирусные векторы и аденоассоциированные вирусные векторы и т.д., но не ограничиваясь ими. Кроме того, см. патенты США № 6534261; 6607882; 6824978; 6933113; 6979539; 7013219; и 7163824, полностью включенные в настоящий документ посредством ссылки. Кроме того, очевидно, что в соответствующих случаях любой из этих векторов может содержать одну или более из последовательностей, кодирующих ДНК-связывающий белок, и/или дополнительных нуклеиновых кислот. Так, при внедрении в клетку одного или более из ДНК-связывающих белков, описанных в настоящем изобретении, и, в соответствующих случаях, дополнительных ДНК, они могут находиться на одном и том же векторе или различных векторах. При использовании множественных векторов каждый вектор при желании может содержать последовательность, кодирующую один или несколько из ДНК-связывающих белков, и дополнительные нуклеиновые кислоты. Для внедрения в клетки (например, клетки млекопитающих) и ткани-мишени нуклеиновых кислот, кодирующих рекомбинантные ДНК-связывающие белки, и, при желании, совместного внедрения дополнительных нуклеиновых кислот можно применять традиционные способы переноса генов, основанные на применении вирусов или не основанные на применении вирусов. Кроме того, такие способы можно применять для введения нуклеиновых кислот (например, кодирующих ДНК-связывающие белки и/или донорных нуклеиновых кислот) в клетки *in vitro*. В некоторых вариантах реализации нуклеиновые кислоты вводят для применения в генной терапии *in vivo* или *ex vivo*. Системы доставки, не основанные на вирусных векторах, включают ДНК-плазмиды, свободные нуклеиновые кислоты и комплексы нуклеиновых кислот с носителем для доставки, например, липосомой, липидной наночастицей или поллоксамером. Системы доставки на основе вирусных векторов включают ДНК-или РНК-вирусы, геномы которых после доставки в клетку существуют в виде эписомы или встраиваются в геном клетки. Обзор процедур генной терапии см. в работах Anderson, *Science* 256:808-813 (1992); Nabel & Feigner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al., in *Current Topics in Microbiology and Immunology Doerfler and Bohm* (eds.) (1995); и Yu et al., *Gene Therapy* 1:13-26 (1994).

Способы доставки нуклеиновых кислот без применения вирусов включают электропорацию, липо-

фекцию, микроинъекцию, баллистическую трансфекцию, виросомы, липосомы, липидные наночастицы, иммунолипосомы, конъюгаты нуклеиновых кислот с поликатионами или липидами, свободную ДНК, мРНК, искусственные вирионы и поглощение ДНК, усиленное за счет агентов. Кроме того, для доставки нуклеиновых кислот можно использовать сонопорацию с помощью, например, системы Sonitron 2000 (Rich-Mar). В предпочтительном варианте реализации одну или более из нуклеиновых кислот доставляют в виде мРНК. Кроме того, предпочтительным является применение экзипированной мРНК для повышения эффективности трансляции и стабильности мРНК. Особенно предпочтительны экзипы ARCA (антиреверсивный аналог экзипа) или их варианты. См. патенты США № 7074596 и 8153773, включенные в настоящий документ посредством ссылок.

Дополнительные типичные системы доставки нуклеиновых кислот включают системы, предложенные Amara Biosystems (Кельн, Германия), Maxcyte, Inc. (Роквилл, штат Мэриленд, США), BTX Molecular Delivery Systems (Холлистон, штат Массачусетс, США) и Copernicus Therapeutics Inc. (см., например, US6008336). Липофекция описана, например, в патентах US 5049386, US 4946787 и US 4897355), а реагенты для липофекции доступны в продаже (например, Transfectam™, Lipofectin™ и Lipofectamine™ RNAiMAX). Катионные и нейтральные липиды, подходящие для эффективной рецептор-распознающей липофекции полинуклеотидов, включают липиды, описанные в Feigner, WO 91/17424, WO 91/16024. Доставка можно осуществлять в клетки (введение *ex vivo*) или ткани-мишени (введение *in vivo*).

Получение комплексов "липид-нуклеиновая кислота", в том числе специфических липосом, например, иммунолипидных комплексов, хорошо известно специалистам 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).

Дополнительные способы доставки включают применение упаковки доставляемых нуклеиновых кислот в носители EnGeneIC для доставки (EDV). Эти EDV специально спроектированы для адресного воздействия на ткани-мишени с использованием биспецифических антител, в которых одно плечо обладает специфичностью по отношению к ткани-мишени, а другое - по отношению к EDV. Антитело переносит EDV к поверхности клетки-мишени, а затем клетка поглощает EDV посредством эндоцитоза. После проникновения в клетку происходит высвобождение содержимого (см. MacDiarmid et al. (2009) Nature Biotechnology 27(7) p. 643).

Применение систем на основе ДНК- или РНК-содержащих вирусов для доставки нуклеиновых кислот, кодирующих рекомбинантные ДНК-связывающие белки, и/или донорных последовательностей (например, CAR или ACTR) желательным образом использует преимущества хорошо разработанных процессов адресного воздействия вируса на специфические клетки в организме и переноса загрузки вируса в ядро. Вирусные векторы можно вводить непосредственно в организмы пациентов (*in vivo*) или применять для обработки клеток *in vitro* с введением модифицированных клеток пациентам (*ex vivo*). Традиционные системы доставки нуклеиновых кислот на основе вирусов включают ретровирусные, лентивирусные, аденовирусные векторы, векторы на основе аденоассоциированных вирусов, вируса коровьей оспы и вируса простого герпеса для переноса генов. Встраивание в геном хозяина возможно для способов переноса генов на основе ретровирусов, лентивирусов и аденоассоциированных вирусов; оно часто приводит к долгосрочной экспрессии вставленного трансгена. Кроме того, для многих видов клеток и тканей-мишеней наблюдается высокая эффективность трансдукции. Тропизм ретровируса можно менять путем встраивания чужеродных белков оболочки, что расширяет потенциальную популяцию клеток-мишеней. Лентивирусные векторы представляют собой ретровирусные векторы, способные трансдуцировать или инфицировать неделящиеся клетки и обычно позволяют получить высокие титры вируса. Выбор ретровирусной системы переноса генов зависит от ткани-мишени. Ретровирусные векторы содержат цис-действующие длинные концевые повторы с возможностью упаковки чужеродной последовательности длиной до 6-10 т.п.о. Минимальные цис-действующие ДКП являются достаточными для репликации и упаковки векторов, которые затем применяют для встраивания терапевтического гена в клетку-мишень с целью получения устойчивой экспрессии трансгена. Широко используемые ретровирусные векторы включают векторы на основе вируса лейкоза мышей (MuLV), вируса лейкоза гиббонов (GaLV), вируса иммунодефицита обезьян (ВИО), вируса иммунодефицита человека (ВИЧ) и их комбинации (см., например, Buchscher et al., J. Virol. 66:2731-2739 (1992); Johann et al., J. Virol. 66:1635-1640 (1992); Sommerfelt et al., Virol. 176:58-59 (1990); Wilson et al., J. Virol. 63:2374-2378 (1989); Miller et al., J. Virol. 65:2220-2224 (1991); PCT/US94/05700).

В вариантах применения, где предпочтительна временная экспрессия, можно применять системы на основе аденовирусов. Векторы на основе аденовирусов обладают крайне высокой эффективностью трансдукции в клетках многих типов и не требуют деления клеток. Такие векторы позволяют получить высокий титр и высокий уровень экспрессии. Этот вектор можно получить в больших количествах в относительно простой системе. Векторы на основе аденоассоциированных вирусов ("AAV") также применяют для трансдукции клеток нуклеиновыми кислотами-мишенями, например, при продукции нуклеиновых кислот и пептидов *in vitro*, и для процедур геновой терапии *in vivo* и *ex vivo* (см., например, West et al.,

Virology 160:38-47 (1987); патент США № 4797368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). Конструирование рекомбинантных AAV-векторов описано в ряде публикаций, в том числе 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 USA 81:6466-6470 (1984); и Samulski et al., J. Virol. 63:03822-3828 (1989). В настоящее время в клинических исследованиях, использующих комплементацию дефектных векторов генами, вставленными в линии клеток-помощников для получения трансдуцирующего агента, существует по меньшей мере шесть подходов к генной терапии на основе вирусных векторов. pLASN и MFG-S являются примерами ретровирусных векторов, применяемых в клинических исследованиях (Dunbar et al., Blood 85:3048-305 (1995); Kohn et al., Nat. Med. 1:1017-102 (1995); Malech et al., PNAS USA 94:22 12133-12138 (1997)). PA317/pLASN представляет собой первый терапевтический вектор, применявшийся в исследовании генной терапии. (Blaise et al., Science 270:475-480 (1995)). Для упакованных векторов MFG-S наблюдается эффективность трансдукции 50% и выше. (Ellem et al., Immunol Immunother. 44(1):10-20 (1997); Dranoff et al., Hum. Gene Ther. 1:111-2 (1997). Векторы на основе рекомбинантных аденоассоциированных вирусов (rAAV) представляют собой перспективные альтернативные системы доставки генов на основе дефектного и непатогенного аденоассоциированного парвовируса 2 типа. Все векторы происходят от плазмиды, сохранившей только инвертированные концевые повторы AAV длиной 145 п.о., фланкирующие экспрессирующую кассету трансгена. Эффективный перенос генов и стабильная доставка трансгена из-за встраивания в геном трансдуцированной клетки являются ключевыми особенностями этой векторной системы. (Wagner et al., Lancet 351:9117 1702-3 (1998), Kearns et al., Gene Ther. 9:748-55 (1996)). Кроме того, в соответствии с настоящим изобретением можно применять AAV других серотипов, в том числе AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV8.2, AAV9 и AAVrh10 и псевдотипированные AAV, например, AAV2/8, AAV2/5 и AAV2/6.

Рекомбинантные аденовирусные векторы (Ad), дефектные по репликации, можно получить в высоких титрах и легко использовать для инфицирования ряда клеток различных типов. Большинство аденовирусных векторов сконструированы таким образом, что трансген заменяет гены Ad E1a, E1b и/или E3; затем вектор, дефектный по репликации, размножают в клетках 293 человека, обеспечивающих функции удаленных генов в транс-режиме. Ad-векторы могут трансдуцировать *in vivo* ткани различных типов, в том числе неделящиеся, дифференцированные клетки, например, находящиеся в печени, почках и мышцах. Традиционные Ad-векторы обладают большой несущей емкостью. Пример применения Ad-вектора в клиническом исследовании включает полинуклеотидное терапевтическое средство для противоопухолевой иммунизации посредством внутримышечной инъекции (Sternan et al., Hum. Gene Ther. 7:1083-9 (1998)). Дополнительные примеры применения аденовирусного вектора для переноса генов в клинических исследованиях включают работы Rosenecker et al., Infection 24:1 5-10 (1996); Sternan et al., Hum. Gene Ther. 9:7 1083-1089 (1998); Welsh et al., Hum. Gene Ther. 2:205-18 (1995); Alvarez et al., Hum. Gene Ther. 5:597-613 (1997); Topf et al., Gene Ther. 5:507-513 (1998); Sternan et al., Hum. Gene Ther. 7:1083-1089 (1998). Упаковывающие клетки применяют для образования вирусных частиц, способных инфицировать клетку-хозяина. Такие клетки включают клетки 293, упаковывающие аденовирусы и AAV, в также клетки ψ 2 или клетки PA317, упаковывающие ретровирусы. Вирусные векторы, применяемые в генной терапии, обычно получают с использованием продуцирующей линии клеток, упаковывающей нуклеотидный вектор в вирусную частицу. Векторы обычно содержат минимальные последовательности вируса, необходимые для упаковки и последующего встраивания в геном хозяина (в соответствующих случаях), а другие последовательности вируса замещают экспрессирующей кассетой, кодирующей белок, подлежащий экспрессии. Отсутствующие функции вируса обеспечивают клетки упаковывающей линии в транс-режиме. Например, AAV-векторы, используемые в генной терапии, обычно содержат только последовательности инвертированных концевых повторов (ИКП) из генома AAV, необходимые для упаковки и встраивания в геном хозяина.

Вирусную ДНК упаковывают в линии клеток, содержащих плазмиду-помощник, кодирующую другие гены AAV, а именно гер и сар, но не содержащую последовательностей ИКП. Клетки этой линии также инфицируют аденовирусом, используемым в качестве вируса-помощника. Вирус-помощник стимулирует репликацию AAV-вектора и экспрессию генов AAV с плазмиды-помощника. Плазмида-помощник не упаковывается в значительных количествах из-за отсутствия последовательностей ИКП. Заражение аденовирусом можно снизить, например, путем термической обработки, к которой аденовирус более чувствителен, чем AAV. Кроме того, AAV можно получать с помощью бакуловирусной системы (см., например, патенты США 6723551 и 7271002).

Выделение частиц AAV из клеток 293 или бакуловирусной системы обычно включает рост клеток, продуцирующих вирус, с последующим сбором вирусных частиц из надосадочной жидкости клеток или лизированием клеток и сбором вируса из необработанного лизата. Затем AAV очищают с помощью способов, известных в данной области техники, в том числе ионообменной хроматографии (см., например, патенты США 7419817 и 6989264), ионообменной хроматографии и центрифугирования в градиенте CsCl (например, публикация PCT WO2011094198A10), иммуноаффинной хроматографии (например, WO2016128408) или очистки на AVB-сепарозе (например, GE Healthcare Life Sciences).

При многих генно-терапевтических вариантах применения желательна доставка вектора для генной терапии в ткань конкретного типа с высокой степенью специфичности. Соответственно, вирусный вектор можно модифицировать, придав ему специфичность по отношению к клеткам данного типа путем экспрессии лиганда в виде гибридного белка с белком оболочки вируса на внешней поверхности вируса. Выбирают лиганд с аффинностью к рецептору, заведомо присутствующему на клетке требуемого типа. Например, Han et al. (Proc. Natl. Acad. Sci. USA 92:9747-9751 (1995)) сообщили, что вирус лейкоза мышей Молони можно модифицировать, придав ему способность к экспрессии херегулина человека, объединенного с gp70, и этот рекомбинантный вирус инфицирует некоторые клетки рака молочной железы человека, экспрессирующие рецептор эпидермального фактора роста человека. Этот принцип можно распространить на другие пары "вирус - клетка-мишень", в которых клетка-мишень экспрессирует рецептор, а вирус экспрессирует гибридный белок, содержащий лиганд рецептора поверхности клетки. Например, можно модифицировать нитевидный фаг с целью экспонирования фрагментов антител (например, Fv или Fv), обладающих аффинностью специфического связывания к практически любому выбранному клеточному рецептору. Хотя вышеприведенное описание применимо в первую очередь к вирусным векторам, тот же принцип можно применять к невирусным векторам. Такие векторы можно конструировать, включая в них специфические последовательности для поглощения, благоприятствующие поглощению специфическими клетками-мишенями.

Доставку векторов для генной терапии можно выполнять *in vivo* путем введения отдельному пациенту, обычно путем системного введения (например, внутривенного, внутрибрюшинного, внутримышечного, субдермального или внутрочерепного вливания) или наружного применения, как описано ниже. В качестве альтернативы, доставку векторов можно выполнять в клетки *ex vivo*, например, в клетки, эксплантированные из отдельного пациента (например, лимфоциты, аспираты костного мозга, биоптаты тканей), или универсальные донорские гемопоэтические стволовые клетки с последующей имплантацией клеток обратно в организм пациента, обычно после отбора клеток, в которые внедрили вектор.

Трансфекция клеток *ex vivo* для диагностики, исследования, трансплантации или генной терапии (например, путем вливания трансфицированных клеток обратно в организм хозяина) хорошо известна специалистам в данной области техники. В предпочтительном варианте реализации клетки выделяют из организма субъекта, трансфицируют нуклеиновой кислотой ДНК-связывающих белков (геном или кДНК) и вливают обратно в организм субъекта (например, пациента). Клетки различных типов, подходящие для трансфекции *ex vivo*, хорошо известны специалистам в данной области техники (обсуждение способов выделения и культивирования клеток из организмов пациентов см., например, в пособии Freshney et al., *Culture of Animal Cells, A Manual of Basic Technique* (3rd ed. 1994)) и источниках, цитируемых в настоящем документе).

В одном варианте реализации в процедурах трансфекции клеток и генной терапии *ex vivo* используют стволовые клетки. Преимущество использования стволовых клеток состоит в том, что их можно дифференцировать в клетки других типов *in vitro* или ввести в организм млекопитающего (например, донора клеток), где они внедрятся в костный мозг. Известны способы дифференцировки CD34⁺-клеток *in vitro* в клинически значимые виды клеток иммунной системы с использованием цитокинов, например, ГМКСФ, ИФН- γ и ФНО- α (см. Inaba et al., *J. Exp. Med.* 176:1693-1702 (1992)).

Стволовые клетки выделяют для трансдукции и дифференцировки с помощью известных способов. Например, стволовые клетки выделяют из клеток костного мозга путем пэннинга клеток костного мозга с помощью антител, связывающих нежелательные клетки, например, CD4⁺ и CD8⁺ (Т-клетки), CD45⁺ (все В-клетки), GR-1 (гранулоциты) и Iad (дифференцированные антиген-презентирующие клетки) (см. Inaba et al., *J. Exp. Med.* 176:1693-1702(1992)).

Кроме того, в некоторых вариантах реализации можно использовать модифицированные стволовые клетки. Например, нервные стволовые клетки, которым придали устойчивость к апоптозу, можно использовать в качестве терапевтических композиций, где стволовые клетки также содержат ZFP TF согласно настоящему изобретению. Устойчивость к апоптозу можно получить, например, путем нокаута BAX и/или BAK с использованием BAX- или BAK-специфичных ZFN (см. публикацию патента США № 12/456,043) в стволовых клетках или клетках, дезорганизованных по каспазе, например, с использованием ZFN, специфичных по отношению к каспазе-6. Эти клетки можно трансфицировать ZFP TF, заведомо регулирующих TCR.

Кроме того, векторы (например, ретровирусы, аденовирусы, липосомы и т.д.), содержащие терапевтические ДНК-связывающие белки (или нуклеиновые кислоты, кодирующие эти белки), можно вводить непосредственно в организм для трансдукции клеток *in vivo*. В качестве альтернативы, можно вводить свободную ДНК. Введение любым путем обычно используют для внедрения молекулы в непосредственном контакте с кровью или клетками ткани, включая инъекцию, вливание, наружное применение и электропорацию, но не ограничиваясь ими. Подходящие способы введения таких нуклеиновых кислот доступны и хорошо известны специалистам в данной области техники и, несмотря на возможность использования более чем одного пути для введения конкретной композиции, конкретный путь часто обеспечивает более быструю и более эффективную реакцию, чем другой путь. Способы внедрения ДНК в гемопоэтические стволовые клетки описаны, например, в патенте США № 5928638. Векторы, которые можно

применять для внедрения трансгенов в гемопоэтические стволовые клетки, например, CD34⁺-клетки, включают аденовирус 35 типа.

Векторы, подходящие для внедрения трансгенов в клетки иммунной системы (например, Т-клетки), включают нестраивающиеся лентивирусные векторы. См., например, Ory et al. (1996) Proc. Natl. Acad. Sci. USA 93:11382-11388; Dull et al. (1998)7. Virology 72:8463-8471; Zuffery et al. (1998) 7. Virology 72:9873-9880; Follenzi et al. (2000) Nature Genetics 25:217-222. Фармацевтически приемлемые носители частично определяются конкретной вводимой композицией, а также конкретным способом, используемым для введения композиции. Соответственно, существует широкий спектр подходящих составов фармацевтических композиций, описанных ниже (см., например, Remington's Pharmaceutical Sciences, 17th ed., 1989). Как отмечалось выше, описанные способы и композиции можно использовать по отношению к клеткам любого типа, включая прокариотические клетки, клетки грибов, клетки архей, клетки растений, клетки насекомых, клетки животных, клетки позвоночных, клетки млекопитающих и клетки человека, в том числе Т-клетки и стволовые клетки любого типа, но не ограничиваясь ими. Подходящие линии клеток для экспрессии белков известны специалистам в данной области техники и включают клетки COS, клетки CHO (например, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11), VERO, MDCK, WI38, V79, B14AF28-G3, ВНК, HaKK, NS0, SP2/0-Ag14, HeLa, HEK293 (например, HEK293-F, HEK293-H, HEK293-T) и perC6, клетки насекомых, например, Spodoptera fugiperda (Sf), или клетки грибов, например, Saccharomyces, Pichia и Schizosaccharomyces, но не ограничиваются ими. Кроме того, можно использовать потомство, варианты и производные этих линий клеток.

Варианты применения.

Описанные композиции и способы можно применять для любого варианта применения, при котором желательна модуляция экспрессии и/или функционала В2М, включая терапевтические и исследовательские варианты применения, при которых желательна модуляция В2М. Например, описанные композиции можно применять *in vivo* и/или *ex vivo* (клеточные терапевтические средства) для дезорганизации экспрессии эндогенного В2М в Т-клетках, модифицированных для адоптивной клеточной терапии с целью экспрессии одного или более из экзогенных CAR, экзогенных TCR, экзогенного ASTR или других рак-специфических рецепторных молекул, тем самым выполняя лечение и/или профилактику рака. Кроме того, при таких условиях модуляция экспрессии В2М в клетке может устранить или существенно снизить риск нежелательной перекрестной реакции со здоровой неспецифической тканью (т.е. реакции трансплантат против хозяина).

Способы и композиции также включают композиции стволовых клеток, в которых ген В2М в стволовых клетках модулирован (модифицирован), а клетки дополнительно содержат ASTR и/или CAR и/или выделенный или рекомбинантный TCR. Например, аллогенные гемопоэтические стволовые клетки, модулированные путем нокаута или нокдауна В2М, можно внедрить в организм пациента с несовпадающим HLA после абляции костного мозга. Эти модифицированные ГСК должны обеспечивать вторичную колонизацию пациента, но не должны вызывать потенциальной РТПХ. Кроме того, внедренные клетки могут содержать другие модификации которые должны способствовать последующей терапии (например, устойчивость к химиотерапии) при лечении основного заболевания. Клетки без HLA также можно применять в качестве "готового" терапевтического средства в неотложных ситуациях у пациентов с травмами.

Кроме того, способы и композиции согласно настоящему изобретению можно применять при проектировании и внедрении моделей *in vitro* и *in vivo*, например, моделей В2М и/или HLA и связанных с ними нарушений у животных, что позволит изучать эти нарушения. Все патенты, патентные заявки и публикации, упомянутые в настоящем документе, полностью включены в данное описание посредством ссылок.

Хотя настоящее описание представлено в некоторых подробностях с помощью иллюстраций и примеров с целью облегчения его понимания, для специалистов очевидно, что в настоящее изобретение можно внести некоторые изменения и модификации без выхода за рамки прилагаемой формулы изобретения. Соответственно, вышеизложенное описание и следующие примеры не следует интерпретировать как ограничивающие.

Примеры

Пример 1. Проектирование В2М-специфичных нуклеаз.

В2М-специфичные ZFN конструировали с целью сайт-специфического получения двуцепочечных разрывов в гене В2М. ZFNs проектировали в основном в соответствии с описанием в статьях Urnov et al. (2005) Nature 435(7042):646-651, Lombardo et al. (2007) Nat Biotechnol. Nov;25(11):1298-306, а также публикациях патентов США 2008-0131962, 2015-016495, 2014-0120622, 2014-0301990 и патенте США 8956828. Пары ZFN оказывали адресное воздействие на различные сайты в константной области гена В2М (см. фиг. 1). Распознающие спирали типичных пар ZFN, а также последовательности-мишени показаны ниже в табл. 1. Сайты-мишени конструировали с мотивами цинкового пальца, специфичных по отношению к В2М, показаны в первом столбце. Нуклеотиды в сайте-мишени, подвергающиеся адресному воздействию распознающих спиралей ZFP, указаны прописными буквами; неспецифические нуклеотиды указаны строчными буквами. Кроме того, показаны линкеры, использованные для соединения домена

нуклеазы FokI и ДНК-связывающего домена ZFP (см. публикацию патента США 20150132269). Например, аминокислотная последовательность доменного линкера L0 представляет собой ДНК-связывающий домен-QLVKS-домен нуклеазы FokI (SEQ ID NO: 3). Аналогичным образом, аминокислотные последовательности доменного линкера N7a представляет собой домен нуклеазы FokI-SGTPHEVGVYTL-ДНК-связывающий домен (SEQ ID NO: 4), а N6a представляет собой домен нуклеазы FokI-SGAQGSTLDF-ДНК-связывающий домен (SEQ ID NO: 5).

Таблица 1

Конструкты с мотивом цинкового пальца, специфичные по отношению к В2М

Название ZFN последовательность	F1	F2	F3	F4	F5	F6	Доменный линкер
Ть-мишень							
SBS57071 5' gcCACGGA gCGAGACA TCTCGgcc cga (SEQ ID NO:6)	RSDDLK (SEQ ID NO:49)	DSSARKK (SEQ ID NO:50)	DRSNLSR (SEQ ID NO:51)	QRTHLRD (SEQ ID NO:52)	QSGHLAR (SEQ ID NO:53)	DSSNREA (SEQ ID NO:54)	L0
SBS57531 5'gaGTAG CGcGAGCA CAGCtaag gccacg (SEQ ID NO:7)	AQCCLFH (SEQ ID NO:55)	DQSNLRA (SEQ ID NO:56)	RSANLNR (SEQ ID NO:57)	RSDDLTR (SEQ ID NO:58)	QSGSLTR (SEQ ID NO:59)	H/O	N6a
SBS57362 5'tcCAGC AGAGAATG GAAAGTca aat (SEQ ID NO:8)	LNHHLQQ (SEQ ID NO:60)	QSGNLAR (SEQ ID NO:61)	RSDTLA (SEQ ID NO:62)	QNAHRKT (SEQ ID NO:63)	RSDNLSE (SEQ ID NO:64)	KPYNLRT (SEQ ID NO:65)	N6a
SBS57376 5'ttTCCCT GAATTGCT ATGTGTct gggttt (SEQ ID NO:9)	TRDHLST (SEQ ID NO:66)	RSDARTN (SEQ ID NO:67)	QSSDLR (SEQ ID NO:68)	HRSSLKN (SEQ ID NO:69)	QSSHLTR (SEQ ID NO:70)	DSSDRKK (SEQ ID NO:71)	L0
SBS57017 5'tgTCGG ATgGATGA AACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	QSGNLAR (SEQ ID NO:61)	TSGNLTR (SEQ ID NO:73)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N7a
SBS57327 5' taGCAATT CAGGAAaT TTGACTtt ccat (SEQ ID NO:11)	DRSNLSR (SEQ ID NO:51)	ARWYLDK (SEQ ID NO:75)	QSGNLAR (SEQ ID NO:61)	AKWNLDA (SEQ ID NO:76)	QQHVLQN (SEQ ID NO:77)	QNATRTK (SEQ ID NO:78)	L0
SBS57328 5'taGCAA TTCAGGAA ATTTgact ttccat (SEQ ID NO:11)	TNQLHW (SEQ ID NO:79)	QSGNLAR (SEQ ID NO:61)	RSDNLRE (SEQ ID NO:80)	ASHVLNA (SEQ ID NO:81)	QNATRTK (SEQ ID NO:78)	H/O	L0

SBS57332 5'tgTCGG ATgGATGA AACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	QSGNLAR (SEQ ID NO:61)	TSANLSR (SEQ ID NO:82)	TSGNLTR (SEQ ID NO:73)	RTEDRLA (SEQ ID NO:83)	N6a
SBS57469 5'tgTCGG ATGGATGA aACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	YTSSLCY (SEQ ID NO:84)	QSGHLSR (SEQ ID NO:85)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N7a
SBS57331 5'tgTCGG ATGGATGA aACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRKN (SEQ ID NO:86)	QSGNLAR (SEQ ID NO:61)	TSANLSR (SEQ ID NO:82)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N6a
SBS57326 5'taGCAA TTCAGGAA aTTTGACT ttccat (SEQ ID NO:11)	DRSNLSR (SEQ ID NO:51)	ARWYLDK (SEQ ID NO:75)	QSGNLAR (SEQ ID NO:61)	AKWNLDA (SEQ ID NO:76)	TTPVLVQ (SEQ ID NO:87)	QNATRTK (SEQ ID NO:78)	L0
SBS55822 5'caTCCG ACATTGAA GTTGACTt actgaa (SEQ ID NO:12)	DRSNLSR (SEQ ID NO:51)	FPGSRTR (SEQ ID NO:88)	QSGNLAR (SEQ ID NO:61)	WRISLAA (SEQ ID NO:89)	DRSNLSR (SEQ ID NO:51)	DSSDRKK (SEQ ID NO:71)	N7a
SBS57511 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLLR (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	HLGLRD (SEQ ID NO:91)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57509 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLLR (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	QSAHRKN (SEQ ID NO:94)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57482 5'caTCCG ACATTGAA GTTGACTt actgaa (SEQ ID NO:12)	DRSNLSR (SEQ ID NO:51)	FPGSRTR (SEQ ID NO:88)	QSGNLAR (SEQ ID NO:61)	HKLSLSI (SEQ ID NO:95)	DRSNLSR (SEQ ID NO:51)	DSSDRKK (SEQ ID NO:71)	N7a
SBS57347 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLLR (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	QNAHRKT (SEQ ID NO:63)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57296 5'ctGAAG AATGGAGA GAGaattg	RSANLTR (SEQ ID NO:57)	QSAHRKN (SEQ ID NO:94)	RHSHLTS (SEQ ID NO:96)	QSGNLAR (SEQ ID NO:61)	QSNQLAV (SEQ ID NO:97)	H/O	N7a

aaaaag (SEQ ID NO:14)							
SBS57322 5' aaAAAG TGGAGCAT TCAGACTt gtcttt (SEQ ID NO:15)	DRSNLSR (SEQ ID NO:51)	QSADRTK (SEQ ID NO:98)	TNQNRI (SEQ ID NO:99)	RSANLTR (SEQ ID NO:57)	RSDSLV (SEQ ID NO:100)	QNANRKT (SEQ ID NO:101)	L0
SBS57323 5' aaAAAGTG GAGCATTC AGACTtgt cttt (SEQ ID NO:15)	DRSNLSR (SEQ ID NO:51)	QSADRTK (SEQ ID NO:98)	LKQNLDA (SEQ ID NO:103)	RSANLTR (SEQ ID NO:57)	RSDSLV (SEQ ID NO:100)	QNANRKT (SEQ ID NO:101)	L0
SBS57447 SBS57296 5' ctGAAG AATGGAGA GAGaattg aaaaag (SEQ ID NO:14)	RSANLTR (SEQ ID NO:57)	QSAHRKN (SEQ ID NO:94)	RHSHLTS (SEQ ID NO:96)	QSGNLAR (SEQ ID NO:61)	QRGNLWT (SEQ ID NO:102)	H/O	N7a

Все ZFN протестировали и обнаружили, что они связываются со своими сайтами-мишенями и обладают нуклеазной активностью. Кроме того, сконструировали направляющие РНК для системы CRISPR/Cas9 *S. pyogenes*, адресно воздействующие на ген B2M. Ниже в табл. 2А указаны последовательности-мишени в гене B2M, а также последовательности направляющих РНК. Все направляющие РНК протестировали в системе CRISPR/Cas9 и подтвердили их активность. Строчная буква "g" на 5'-конце некоторых направляющих последовательностей обозначает добавленный нуклеотид G для использования в последовательности PAM.

Таблица 2А

Направляющие РНК для константной области B2M человека

Название	Цепь	Мишень (5' → 3')	нРНК (5' → 3')
B2M- Gf1073	п	GGCCGAGATGTCTCGCTCCGTGG (SEQ ID NO:16)	GGCCGAGATGTCTCGCTCCG (SEQ ID NO:104)
B2M- Gr1074	о	CGCGAGCACAGCTAAGGCCACGG (SEQ ID NO:17)	gCGCGAGCACAGCTAAGGCCA (SEQ ID NO:105)
B2M- Gr1080	о	GAGTAGCGGAGCACAGCTAAGG (SEQ ID NO:18)	GAGTAGCGGAGCACAGCTA (SEQ ID NO:106)
B2M- Gf1107	п	CTCGCGCTACTCTCTTTCTGG (SEQ ID NO:19)	gCTCGCGCTACTCTCTTTCT (SEQ ID NO:107)
B2M- Gf1112	п	GCTACTCTCTTTCTGGCCTGG (SEQ ID NO:20)	GCTACTCTCTTTCTGGCC (SEQ ID NO:108)
B2M- Gf1115	п	ACTCTCTTTCTGGCCTGGAGG (SEQ ID NO:21)	gACTCTCTTTCTGGCCTGG (SEQ ID NO:109)
B2M- Gr1114	о	ACTCACGCTGGATAGCCTCCAGG (SEQ ID NO:22)	gACTCACGCTGGATAGCCTCC (SEQ ID NO:110)
B2M- Gr1126	о	AGGGTAGGAGAGACTCACGCTGG (SEQ ID NO:23)	gAGGGTAGGAGAGACTCACGC (SEQ ID NO:111)
B2M- Gr4942	о	CGTGAGTAAACCTGAATCTTTGG (SEQ ID NO:24)	gCGTGAGTAAACCTGAATCTT (SEQ ID NO:112)
B2M- Gf4948	п	CTCAGGTACTCCAAAGATTCAGG (SEQ ID NO:25)	gCTCAGGTACTCCAAAGATTC (SEQ ID NO:113)

B2M-Gr4969	o	TTTGA ^c CTTTCCATTCTCTGCTGG (SEQ ID NO:26)	gTTTGA ^c CTTTCCATTCTCTGC (SEQ ID NO:114)
B2M-Gf4976	π	TCACGTCATCCAGCAGAGAATGG (SEQ ID NO:27)	gTCACGTCATCCAGCAGAGAA (SEQ ID NO:115)
B2M-Gr4995	o	ACCCAGACACATAGCAATTCAGG (SEQ ID NO:28)	gACCCAGACACATAGCAATTC (SEQ ID NO:116)
B2M-Gf5009	π	TTCCTGAATTGCTATGTGTCTGG (SEQ ID NO:29)	gTTCCTGAATTGCTATGTGTC (SEQ ID NO:117)
B2M-Gf5010	π	TCCTGAATTGCTATGTGTCTGGG (SEQ ID NO:30)	gTCCTGAATTGCTATGTGCT (SEQ ID NO:118)
B2M-Gr5023	o	AAGTCAACTTCAATGTCCGGATGG (SEQ ID NO:31)	gAAGTCAACTTCAATGTCCGGA (SEQ ID NO:119)
B2M-Gr5027	o	CAGTAAGTCAACTTCAATGTCGG (SEQ ID NO:32)	gCAGTAAGTCAACTTCAATGT (SEQ ID NO:120)
B2M-Gf5051	π	GAAGTTGACTTACTGAAGAATGG (SEQ ID NO:33)	GAAGTTGACTTACTGAAGAA (SEQ ID NO:121)
B2M-Gf5071	π	TGGAGAGAGAATTGAAAAAGTGG (SEQ ID NO:34)	gTGGAGAGAGAATTGAAAAAG (SEQ ID NO:122)
B2M-Gf5098	π	TTCAGACTTGTCTTTCAGCAAGG (SEQ ID NO:35)	gTTCAGACTTGTCTTTCAGCA (SEQ ID NO:123)
B2M-Gf5103	π	ACTTGTCTTTCAGCAAGGACTGG (SEQ ID NO:36)	gACTTGTCTTTCAGCAAGGAC (SEQ ID NO:124)
B2M-Gr5141	o	ATACTCATCTTTTTTCAGTGGGG (SEQ ID NO:37)	gATACTCATCTTTTTTCAGTGG (SEQ ID NO:125)
B2M-Gr5142	o	CATACTCATCTTTTTTCAGTGGGG (SEQ ID NO:38)	gCATACTCATCTTTTTTCAGTG (SEQ ID NO:126)
B2M-Gr5143	o	GCATACTCATCTTTTTTCAGTGGG (SEQ ID NO:39)	GCATACTCATCTTTTTTCAGT (SEQ ID NO:127)
B2M-Gr5144	o	GGCATACTCATCTTTTTTCAGTGG (SEQ ID NO:40)	GGCATACTCATCTTTTTTCAG (SEQ ID NO:128)
B2M-Gr5165	o	AGTCACATGGTTCACACGGCAGG (SEQ ID NO:41)	gAGTCACATGGTTCACACGGC (SEQ ID NO:129)
B2M-Gr5169	o	ACAAAGTCACATGGTTCACACGG (SEQ ID NO:42)	gACAAAGTCACATGGTTCACA (SEQ ID NO:130)
B2M-Gr5178	o	TGGGCTGTGACAAAGTCACATGG (SEQ ID NO:43)	gTGGGCTGTGACAAAGTCACA (SEQ ID NO:131)
B2M-Gr5197	o	TTACCCCACTTAACTATCTTGGG (SEQ ID NO:44)	gTTACCCCACTTAACTATCTT (SEQ ID NO:132)
B2M-Gr5198	o	CTTACCCCACTTAACTATCTTGG (SEQ ID NO:45)	gCTTACCCCACTTAACTATCT (SEQ ID NO:133)
B2M-Gf5208	π	CACAGCCCAAGATAGTTAAGTGG (SEQ ID NO:46)	gCACAGCCCAAGATAGTTAAG (SEQ ID NO:134)
B2M-Gf5209	π	ACAGCCCAAGATAGTTAAGTGGG (SEQ ID NO:47)	gACAGCCCAAGATAGTTAAGT (SEQ ID NO:135)
B2M-Gf5210	π	CAGCCCAAGATAGTTAAGTGGGG (SEQ ID NO:48)	gCAGCCCAAGATAGTTAAGTG (SEQ ID NO:136)

TALEN, полученные для адресного воздействия на локус B2M, показаны ниже в табл. 2B. Все TALEN протестировали в клетках K562; обнаружено, что они были активными (см. табл. 2C и фиг. 2B).

TALEN, специфичные по отношению к В2М

SBS#	Сайт-мишень 5' -> 3'	RVD N -> C
103049	atTCGGGCCGAGATGTCTCgc (SEQ ID NO:137)	NG-HD-NN-NN-NN-HD-HD-NN-NI-NN-NI-NG-NN-NG-HD-NG-HD
103050	gtAGCGCGAGCACAGCTAAgg (SEQ ID NO:138)	NI-NN-HD-NN-HD-NN-NI-NN-HD-NI-HD-NI-NN-HD-NG-NI-NI
103051	ctCCGTGGCCTTAGCTGTGct (SEQ ID NO:139)	HD-HD-NN-NG-NN-NN-HD-HD-NG-NG-NI-NN-HD-NG-NN-NG-NK
103052	ctCCAGGCCAGAAAGAGAgg (SEQ ID NO:140)	HD-HD-NI-NN-NN-HD-HD-NI-NN-NI-NI-NI-NN-NI-NN-NI-NK
103053	gtGGCCTTAGCTGTGCTCGcg (SEQ ID NO:141)	NN-NN-HD-HD-NG-NG-NI-NN-HD-NG-NN-NG-NN-HD-NG-HD-NK
103054	atAGCCTCCAGGCCAGAAga (SEQ ID NO:142)	NI-NN-HD-HD-NG-HD-HD-NI-NN-NN-HD-HD-NI-NN-NI-NI-NI
103055	ctTAGCTGTGCTCGCGCTAct (SEQ ID NO:143)	NG-NI-NN-HD-NG-NN-NG-NN-HD-NG-HD-NN-HD-NN-HD-NG-NI
103056	ctGGATAGCCTCCAGGCCAga (SEQ ID NO:144)	NN-NN-NI-NG-NI-NN-HD-HD-NG-HD-HD-NI-NN-NN-HD-HD-NI
103057	ctGTGCTCGCGCTACTCTctc (SEQ ID NO:145)	NN-NG-NN-HD-NG-HD-NN-HD-NN-HD-NG-NI-HD-NG-HD-NG-HD
103058	ctCACGCTGGATAGCCTCCag (SEQ ID NO:146)	HD-NI-HD-NN-HD-NG-NN-NN-NI-NG-NI-NN-HD-HD-NG-HD-HD
103059	ctACTCTCTCTTTCTGGCctg (SEQ ID NO:147)	NI-HD-NG-HD-NG-HD-NG-HD-NG-NG-NG-HD-NG-NN-NN-HD-HD
103060	gtAGGAGAGACTCACGCTGga (SEQ ID NO:148)	NI-NN-NN-NI-NN-NI-NN-NI-HD-NG-HD-NI-HD-NN-HD-NG-NK
103061	gtGTCTTTTCCCGATATTCct (SEQ ID NO:149)	NN-NG-HD-NG-NG-NG-NG-HD-HD-HD-NN-NI-NG-NI-NG-NG-HD
103062	gtGAGTAAACCTGAATCTTtg (SEQ ID NO:150)	NN-NI-NN-NG-NI-NI-NI-HD-HD-NG-NN-NI-NI-NG-HD-NG-NG
103063	ttTTCCCGATATTCCTCAGgt (SEQ ID NO:151)	NG-NG-HD-HD-HD-NN-NI-NG-NI-NG-NG-HD-HD-NG-HD-NI-NK

103064	atGACGTGAGTAAACCTGAat (SEQ ID NO:152)	NN-NI-HD-NN-NG-NN-NI-NN-NG-NI-NI-NI-HD-HD-NG-NN-NI
103065	ttCCTCAGGTACTCCAAAGat (SEQ ID NO:153)	HD-HD-NG-HD-NI-NN-NN-NG-NI-HD-NG-HD-HD-NI-NI-NI-NK
103066	ctCTGCTGGATGACGTGAGta (SEQ ID NO:154)	HD-NG-NN-HD-NG-NN-NN-NI-NG-NN-NI-HD-NN-NG-NN-NI-NK
103067	ctCAGGTACTCCAAAGATTca (SEQ ID NO:155)	HD-NI-NN-NN-NG-NI-HD-NG-HD-HD-NI-NI-NI-NN-NI-NG-NG
103068	atTCTCTGCTGGATGACGTga (SEQ ID NO:156)	NG-HD-NG-HD-NG-NN-HD-NG-NN-NN-NI-NG-NN-NI-HD-NN-NG
103069	gtACTCCAAGATTCAGGttt (SEQ ID NO:157)	NI-HD-NG-HD-HD-NI-NI-NI-NN-NI-NG-NG-HD-NI-NN-NN-NG
103070	ttTCCATTCTCTGCTGGATga (SEQ ID NO:158)	NG-HD-HD-NI-NG-NG-HD-NG-HD-NG-NN-HD-NG-NN-NN-NI-NG
103071	ctCCAAAGATTCAGGTTTAct (SEQ ID NO:159)	HD-HD-NI-NI-NI-NN-NI-NG-NG-HD-NI-NN-NN-NG-NG-NG-NI
103072	ttGACTTTCCATTCTCTGctg (SEQ ID NO:160)	NN-NI-HD-NG-NG-NG-HD-HD-NI-NG-NG-HD-NG-HD-NG-NN-HD
103073	ctCACGTCATCCAGCAGAGaa (SEQ ID NO:161)	HD-NI-HD-NN-NG-HD-NI-NG-HD-HD-NI-NN-HD-NI-NN-NI-NK
103074	atAGCAATTCAGGAAATTTga (SEQ ID NO:162)	NI-NN-HD-NI-NI-NG-NG-HD-NI-NN-NN-NI-NI-NI-NG-NG-NG
103075	ttCCCTGAATTGCTATGTGTct (SEQ ID NO:163)	HD-HD-NG-NN-NI-NI-NG-NG-NN-HD-NG-NI-NG-NN-NG-NN-NG
103076	gtCAACTTCAATGTCGGATgg (SEQ ID NO:164)	HD-NI-NI-HD-NG-NG-HD-NI-NI-NG-NN-NG-HD-NN-NN-NI-NG
103077	ctATGTGTCTGGGTTTCATcc (SEQ ID NO:165)	NI-NG-NN-NG-NN-NG-HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG
103078	ttCTTCAGTAAGTCAACTTca (SEQ ID NO:166)	HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD-NI-NI-HD-NG-NG
103079	atGTGTCTGGGTTTCATCCat (SEQ ID NO:167)	NN-NG-NN-NG-HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG-HD-HD
103080	atTCTTCAGTAAGTCAACTtc (SEQ ID NO:168)	NG-HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD-NI-NI-HD-NG
103081	gtCTGGGTTTCATCCATCCga (SEQ ID NO:169)	HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG-HD-HD-NI-NG-HD-HD
103082	ctCCATTCTTCAGTAAGTCaa (SEQ ID NO:170)	HD-HD-NI-NG-NG-HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD
103083	ttTCATCCATCCGACATTGaa (SEQ ID NO:171)	NG-HD-NI-NG-HD-HD-NI-NG-HD-HD-NN-NI-HD-NI-NG-NG-NK

103084	ttCTCTCTCCATTCTTCAGta (SEQ ID NO:172)	HD-NG-HD-NG-HD-NG-HD-HD-NI-NG-NG-HD-NG-NG-HD-NI-NK
103085	atCCATCCGACATTGAAGTtg (SEQ ID NO:173)	HD-HD-NI-NG-HD-HD-NN-NI-HD-NI-NG-NG-NN-NI-NI-NN-NG
103086	ttCAATTCTCTCTCCATTctt (SEQ ID NO:174)	HD-NI-NI-NG-NG-HD-NG-HD-NG-HD-NG-HD-HD-NI-NG-NG-HD
103087	atCCGACATTGAAGTTGACTt (SEQ ID NO:175)	HD-HD-NN-NI-HD-NI-NG-NG-NN-NI-NI-NN-NG-NG-NN-NI-HD
103088	ctTTTTCAATTCTCTCTCCat (SEQ ID NO:176)	NG-NG-NG-NG-HD-NI-NI-NG-NG-HD-NG-HD-NG-HD-NG-HD-HD
103089	ttGAAGTTGACTTACTGAAga (SEQ ID NO:177)	NN-NI-NI-NN-NG-NG-NN-NI-HD-NG-NG-NI-HD-NG-NN-NI-NI
103090	atGCTCCACTTTTTCAATTct (SEQ ID NO:178)	NN-HD-NG-HD-HD-NI-HD-NG-NG-NG-NG-NG-NG-HD-NI-NI-NG-NG
103091	gtTGACTTACTGAAGAATGga (SEQ ID NO:179)	NG-NN-NI-HD-NG-NG-NI-HD-NG-NN-NI-NI-NN-NI-NI-NG-NK
103092	gtCTGAATGCTCCACTTTTtc (SEQ ID NO:180)	HD-NG-NN-NI-NI-NG-NN-HD-NG-HD-HD-NI-HD-NG-NG-NG-NG
103093	atGGAGAGAGAATTGAAAaag (SEQ ID NO:181)	NN-NN-NI-NN-NI-NN-NI-NN-NI-NI-NG-NG-NN-NI-NI-NI-NI
103094	ctTGCTGAAAGACAAGTCTga (SEQ ID NO:182)	NG-NN-HD-NG-NN-NI-NI-NI-NN-NI-HD-NI-NI-NN-NG-HD-NG
103095	ttCAGACTTGTCTTTCAGCaa (SEQ ID NO:183)	HD-NI-NN-NI-HD-NG-NG-NN-NG-HD-NG-NG-NG-HD-NI-NN-HD
103096	gtGTAGTACAAGAGATAGAaa (SEQ ID NO:184)	NN-NG-NI-NN-NG-NI-HD-NI-NI-NN-NI-NN-NI-NG-NI-NN-NI
103097	ctTGCTTTTCAGCAAGGACTg (SEQ ID NO:185)	NG-NN-NG-HD-NG-NG-NG-HD-NI-NN-HD-NI-NI-NN-NN-NI-HD
103098	atTCAGTGTAGTACAAGAGat (SEQ ID NO:186)	NG-HD-NI-NN-NG-NN-NG-NI-NN-NG-NI-HD-NI-NI-NN-NI-NK
103099	ctTTCAGCAAGGACTGGTctt (SEQ ID NO:187)	NG-NG-HD-NI-NN-HD-NI-NI-NN-NN-NI-HD-NG-NN-NN-NG-HD
103100	gtGAATTCAGTGTAGTACAag (SEQ ID NO:188)	NN-NI-NI-NG-NG-HD-NI-NN-NG-NN-NG-NI-NN-NG-NI-HD-NI
103101	ctGGTCTTTCTATCTCTTgta (SEQ ID NO:189)	NN-NN-NG-HD-NG-NG-NG-HD-NG-NI-NG-HD-NG-HD-NG-NG-NK
103102	ttTTTCAGTGGGGTGAATtc (SEQ ID NO:190)	NG-NG-NG-HD-NI-NN-NG-NN-NN-NN-NN-NG-NN-NI-NI-NG
103103	ctATCTCTTGTACTACTga (SEQ ID NO:191)	NI-NG-HD-NG-HD-NG-NG-NN-NG-NI-HD-NG-NI-HD-NI-HD-NG

103104	atACTCATCTTTTTCAGTGgg (SEQ ID NO:192)	NI-HD-NG-HD-NI-NG-HD-NG-NG-NG-NG-NG-HD-NI-NN-NG-NK
103105	ctACACTGAATTCACCCCac (SEQ ID NO:193)	NI-HD-NI-HD-NG-NN-NI-NI-NG-NG-HD-NI-HD-HD-HD-HD-HD
103106	ttCACACGGCAGGCATACTca (SEQ ID NO:194)	HD-NI-HD-NI-HD-NN-NN-HD-NI-NN-NN-HD-NI-NG-NI-HD-NG
103107	atTCACCCCCACTGAAAAga (SEQ ID NO:195)	NG-HD-NI-HD-HD-HD-HD-HD-NI-HD-NG-NN-NI-NI-NI-NI-NI
103108	gtCACATGGTTCACACGGCag (SEQ ID NO:196)	HD-NI-HD-NI-NG-NN-NN-NG-NG-HD-NI-HD-NI-HD-NN-NN-HD
103109	atTCACCCCCACTGAAAAga (SEQ ID NO:195)	NG-HD-NI-HD-HD-HD-HD-HD-NI-HD-NG-NN-NI-NI-NI-NI-NI
103110	gtCACATGGTTCACACGGCag (SEQ ID NO:186)	HD-NI-HD-NI-NG-NN-NN-NG-NG-HD-NI-HD-NI-HD-NN-NN-HD
103111	ctGAAAAAGATGAGTATGCct (SEQ ID NO:197)	NN-NI-NI-NI-NI-NI-NN-NI-NG-NN-NI-NN-NG-NI-NG-NN-HD
103112	ctGTGACAAAGTCACATGGtt (SEQ ID NO:198)	NN-NG-NN-NI-HD-NI-NI-NI-NN-NG-HD-NI-HD-NI-NG-NN-NK
103113	atGAGTATGCCTGCCGTGga (SEQ ID NO:199)	NN-NI-NN-NG-NI-NG-NN-HD-HD-NG-NN-HD-HD-NN-NG-NN-NG
103114	ctATCTTGGGCTGTGACAAag (SEQ ID NO:200)	NI-NG-HD-NG-NG-NN-NN-NN-HD-NG-NN-NG-NN-NI-HD-NI-NI
103115	gtATGCCTGCCGTGTAACca (SEQ ID NO:201)	NI-NG-NN-HD-HD-NG-NN-HD-HD-NN-NG-NN-NG-NN-NI-NI-HD
103116	ttAACTATCTTGGGCTGTGac (SEQ ID NO:202)	NI-NI-HD-NG-NI-NG-HD-NG-NG-NN-NN-NN-HD-NG-NN-NG-NK
103117	gtGTGAACCATGTGACTTTgt (SEQ ID NO:203)	NN-NG-NN-NI-NI-HD-HD-NI-NG-NN-NG-NN-NI-HD-NG-NG-NG
103118	ttACCCCACTTAACTATCTtg (SEQ ID NO:204)	NI-HD-HD-HD-HD-NI-HD-NG-NG-NI-NI-HD-NG-NI-NG-HD-NG
103119	atGTGACTTTGTCACAGCCca (SEQ ID NO:205)	NN-NG-NN-NI-HD-NG-NG-NG-NN-NG-HD-NI-HD-NI-NN-HD-HD

TALEN из табл. 2B протестировали в трех различных концентрациях - 25, 100 или 400 нг каждого TALEN на реакцию. Обнаружено, что все протестированные TALEN связывались со своими сайтами-мишенями и обладали нуклеазной активностью; типичные данные показаны в табл. 2C и на фиг. 2B.

Активность пар TALEN в клетках K562

SBS#	% инсерций/делеций - 25 нг	% инсерций/делеций - 100 нг	% инсерций/делеций - 400 нг
103049:103050	2,0	7,3	19,5
103051:103052	18,3	38,7	63,8
103053:103054	12,0	17,4	32,4
103055:103056	6,3	12,7	25,1
103057:103058	15,7	24,5	46,2
103061:103062	3,1	5,7	22,9
103063:103064	1,9	4,0	14,8
103065:103066	7,8	13,7	41,2
103067:103068	14,1	25,6	49,3
103069:103070	2,4	5,0	27,9
103071:103072	1,5	3,6	13,3
103073:103074	0,1	0,5	3,0
103075:103076	0,3	0,5	1,5
103077:103078	2,0	5,8	17,1
103079:103080	15,3	30,1	42,3
103081:103082	5,2	16,2	27,5
103083:103084	7,3	12,2	32,2
103085:103086	0,3	1,3	3,9
103087:103088	0,7	4,4	10,5
103089:103090	1,3	8,7	16,1
103091:103092	14,3	33,5	48,5
103093:103094	2,4	7,6	20,2
103095:103096	12,0	23,5	42,0
103097:103098	10,0	28,3	52,0
103099:103100	1,9	7,5	15,3
103101:103102	3,3	7,0	15,5
103103:103104	15,8	29,3	44,9
103105:103106	2,2	5,9	14,2
103107:103108	1,0	2,1	4,2
103109:103110	1,0	3,8	8,0
103111:103112	11,3	30,2	26,5
103113:103114	13,5	22,2	26,6
103115:103116	29,8	41,0	66,1
103117:103118	5,8	20,6	45,7
103119:103120	14,5	40,9	57,7

Таким образом, нуклеазы, описанные в настоящем изобретении (например, нуклеазы, содержащие ДНК-связывающий домен ZFP, TALE или онРНК) связывались со своими сайтами-мишенями и расщепляли ген B2M, тем самым внося генетические модификации в пределах гена B2M, содержащего любую из SEQ ID NO: 6-48 или 137-205, в том числе модификации (инсерции и/или делеции) в пределах любой из этих последовательностей (SEQ ID NO: 6-48 или 137-205); модификации в пределах 1-50 (например, 1-10) пар оснований последовательностей этого гена; модификации между сайтами-мишенями парных сайтов-мишеней (для димеров); и/или модификации в пределах одной или более из следующих последовательностей: GGCCTTA, TCAAATT, TCAAAT, TTACTGA и/или AATTGAA (см. фиг. 1).

Кроме того, все ДНК-связывающие домены (ZFP, TALE и онРНК) связывались со своими сайтами-мишенями, а также были получены в виде активных рекомбинантных факторов транскрипции при ассоциации с одним или более из доменов регуляции транскрипции.

Пример 2. B2M-специфическая активность ZFN в Т-клетках.

B2M-специфические пары ZFN протестировали в Т-клетках человека на предмет нуклеазной активности. мРНК, кодирующие ZFN, трансфицировали в очищенные Т-клетки. Вкратце, Т-клетки получили из продукта лейкофереза и очистили с помощью системы Miltenyi CliniMACS (двойной отбор CD4 и CD8). Затем эти клетки активировали с использованием Dynabeads (ThermoFisher) в соответствии с протоколом производителя. Через 3 дня после активации клетки трансфицировали двумя дозами мРНК (в

общей сложности 2 или 6 мкг двух ZFN) с использованием 96-луночного электропоратора ВТХ (ВТХ) в соответствии со стандартными протоколами. Затем клетки размножали в течение еще 7 дней, в общей сложности в течение 10 дней после активации. Клетки извлекали на 7 день и анализировали на предмет модификации мишени В2М посредством глубокого секвенирования (Miseq, Illumina), а также извлекали на 10 день для FACS-анализа с использованием окрашивания HLA-A, -B и -C.

Все В2М-специфические пары ZFN являлись активными в Т-клетках и вызвали в среднем 89% и 83% эффект для доз 6 мкг и 2 мкг мРНК, соответственно (см. фиг. 2). Используемые пары и положения (показанные на фиг. 1) перечислены ниже в табл. 3.

Таблица 3
В2М-специфические пары ZFN и сайты-мишени

Пара SBS	Сайт, пара
57071 и 57531	A1
57362 и 57376	C1
57017 и 57327	D1
57017 и 57328	D2
57332 и 57327	D3
57469 и 57327	D4
57469 и 57328	D5
57331 и 57326	D6
55822 и 57511	E1
55822 и 57509	E2
57482 и 57511	E3
55822 и 57347	E4
57296 и 57322	G1
57296 и 57323	G2
57447 и 57322	G3

Аналогичным образом, Т-клетки, обработанные ZFN, теряли экспрессию HLA A, B и C, причем FACS-анализ показал, что среднее количество HLA-отрицательных Т-клеток составляло 81 и 67% при дозах 6 мкг и 2 мкг РНК соответственно (см. фиг. 3).

Пример 3. Активность направляющих РНК против В2М *in vitro*.

В этих экспериментах Cas9 вводили на плазмиде pVAX, а онРНК вводили на плазмиде под контролем промотора U6. Плазмиды смешивали по 100 нг или по 400 нг и смешивали с 2e5 клеток на цикл эксперимента. Клетки трансфицировали с использованием системы Амаха. Вкратце, использовали набор для трансфекции Амаха и трансфицировали нуклеиновые кислоты с использованием стандартного челночного протокола Амаха. После трансфекции клетки оставляли на 10 мин при комнатной температуре, а затем ресуспендировали в подогретой RPMI. Затем клетки выращивали в стандартных условиях при 37°C. Геномную ДНК выделяли через 7 дней после трансфекции и подвергали анализу на MiSeq.

В данных, показанных ниже (в табл. 4), указан процент инсерций и делеций, обнаруженный при двух дозах направляющих РНК, а также показано, что различные направляющие РНК индуцировали расщепление в сайте-мишени. Цифры представляют собой среднее по двум экспериментам. Все направляющие РНК являлись активными.

Таблица 4

Активность системы CRISPR/Cas по отношению к B2M

Направляющая РНК	ср. для 100 нг (% инсерций/делеций)	ср. для 400 нг (% инсерций/делеций)	GFP (% инсерций/делеций)
B2M-Gf1073	31,12	55,85	0,20
B2M-Gr1074	35,21	55,86	0,24
B2M-Gr1080	24,83	62,22	0,17
B2M-Gf1107	4,32	52,68	0,19
B2M-Gf1112	9,99	22,09	0,11
B2M-Gf1115	26,52	28,97	0,16
B2M-Gr1114	16,93	53,88	0,12
B2M-Gr1126	15,25	55,92	0,04
B2M-Gr4942	3,05	48,28	0,15
B2M-Gf4948	1,62	12,18	0,11
B2M-Gr4969	1,68	10,11	0,16
B2M-Gf4976	7,47	12,65	0,14
B2M-Gr4995	1,14	31,89	0,16
B2M-Gf5009	0,93	8,87	0,10
B2M-Gf5010	0,30	5,84	0,16
B2M-Gr5023	2,66	2,41	0,22
B2M-Gr5027	14,90	13,13	0,19
B2M-Gf5051	12,98	42,24	0,25
B2M-Gf5071	5,16	47,59	0,44
B2M-Gf5098	4,17	25,60	0,35
B2M-Gf5103	4,00	24,85	0,45
B2M-Gr5141	3,53	18,55	0,32
B2M-Gr5142	3,26	18,22	0,35
B2M-Gr5143	1,79	20,51	0,35
B2M-Gr5144	7,04	6,99	0,34
B2M-Gr5165	6,48	26,44	0,56
B2M-Gr5169	4,62	34,59	0,44
B2M-Gr5178	3,62	30,98	0,61
B2M-Gr5197	2,32	23,26	0,52
B2M-Gr5198	1,87	20,26	0,53
B2M-Gf5208	6,20	17,64	0,55
B2M-Gf5209	10,09	35,92	0,17
B2M-Gf5210	8,47	21,72	0,21

Пример 4. Двойной нокаут B2M и TCR в первичных Т-клетках.

Пары B2M, описанные в настоящем изобретении, также протестировали в комбинации с ZFN, специфичной по отношению к TCRA (см. табл. 5 ниже и предварительные заявки на патент США 62/269365 и 62/306500). Клетки получали и обрабатывали, как описано в примере 2. мРНК, кодирующие пары ZFN (SBS#57017/SBS#57327 для B2M и SBS#55254/SBS#55248 для TCRA), внедрили в клетки посредством электропорации с помощью прибора Maxcyte в соответствии с инструкциями производителя. Вкратце, Т-клетки активировали в 0 день и обрабатывали ZFN-кодирующей мРНК на 3 день, когда плотность клеток составляла 3×10^7 клеток/мл. Электропорацию выполняли с использованием холодового шока при 30°C в течение ночи после электропорации. На 4 день клетки подсчитывали и анализировали на предмет жизнеспособности, разбавляли до 0,5еб клеток/мл и переносили на 37°C. На 7 день клетки подсчитывали, повторно анализировали и повторно разбавляли до 0,5еб клеток/мл. На 10 и 14 день отбирали порции клеток для FACS и глубокого секвенирования на MiSeq.

Таблица 5

ZFN TCRA

Название ZFN последовательность-мишень	F1	F2	F3	F4	F5	F6	Доменный линкер
SBS55266 5' tcAAGCTGG TCGAGaAAAGC Tttgaaac (SEQ ID NO:206)	QSSDLR (SEQ ID NO:68)	QSGNRRT (SEQ ID NO:208)	RSANLAR (SEQ ID NO:209)	DRSALAR (SEQ ID NO:210)	RSDVLSE (SEQ ID NO:211)	KHSTRRV (SEQ ID NO:212)	N7c
SBS53853 5' aaCAGGTAA GACAGGGGTCT Agcctggg (SEQ ID NO:207)	TMHQEVE (SEQ ID NO:213)	TSGHLSR (SEQ ID NO:214)	RSDHLTQ (SEQ ID NO:215)	DSANLSR (SEQ ID NO:216)	QSGSLTR (SEQ ID NO:59)	AKWNLDA (SEQ ID NO:76)	L0
SBS55254 5' ctCCTGAAA GTGGCCGGgtt taatctgc (SEQ ID NO:217)	RSDHLST (SEQ ID NO:219)	DRSHLAR (SEQ ID NO:220)	LKQHLNE (SEQ ID NO:221)	QSGNLAR (SEQ ID NO:61)	HNSSLKD (SEQ ID NO:222)	H/O	L0
SBS55248 5' agGATTCGG AACCCAATCAC tgacaggt (SEQ ID NO:218)	DQSNLRA (SEQ ID NO:56)	TSSNRKT (SEQ ID NO:223)	LQOTLAD (SEQ ID NO:224)	QSGNLAR (SEQ ID NO:61)	RREDLIT (SEQ ID NO:225)	TSSNLSR (SEQ ID NO:226)	L0

Клетки разделили на четыре группы: Контроль без ZFN, только ZFN TCRA, только ZFN B2M и ZFN TCRA + ZFN B2M. При FACS-анализе наборы ZFN против только TCRA (180 мкг/мл мРНК ZFN) или только B2M (180 мкг/мл мРНК ZFN) позволили получить высокую частоту расщепления (96% CD3-маркеров для ZFN TCRA и 92% нокаут HLA-маркеров для ZFN B2M (фиг. 4)). При обработке клеток парами ZFN обоих типов (по 180 мкг/мл) 82% клеток теряли и CD3, и HLA-маркеры.

Кроме того, аналогичные группы клеток обрабатывали различным количеством TCRA-специфическими ZFN, как показано (60- 250 мкг/мкл) и ZFN B2M (в концентрации 60 мкг/мл), и на 10 и 14 дни подвергали глубокому секвенированию на MiSeq (Illumina) и FACS-анализу; результаты показаны ниже в табл. 6. Результаты показывают, что при использовании этих ZFN наблюдали высокую частоту двойного нокаута, согласно обнаружению с помощью NHEJ-опосредованных инсерций и делеций.

Таблица 6

FACS- и miSeq-анализ при двойном нокауте TCRA/B2M

мкг/мкл	FACS						MiSeq			
	D10			D14			D10		D14	
	TCRA-	B2M-	ДВОЙНЫЕ-	TCRA-	B2M-	ДВОЙНЫЕ-	TCRA-	B2M-	TCRA	B2M-
фиктивная группа	2,3	2,9	2,1	3,0	3,8	2,1	0,6	1,8	0,2	1,4
TRAC 180	67,4	1,8	1,5	40,1	1,0	0,5	66,0	0,8	52,4	2,9
B2M 180	3,4	90,5	3,1	1,5	89,5	1,4	0,9	64,6	0,2	74,7
T180: B180	60,4	84,3	57,3	43,9	78,8	41,8	63,2	73,0	51,1	63,4
T240: B120	73,3	85,3	68,3	56,9	80,2	53,2	73,0	79,5	66,0	66,7
T120: B240	79,4	80,5	71,2	70,6	73,9	63,2	79,0	72,1	69,2	61,3

Таким образом, данные демонстрируют, что двойной нокаут B2M и TCRA инактивировал оба гена в области мишени или вблизи от нее и/или в пределах 1-50 (например, 1-10) пар оснований (в том числе между парными сайтами-мишенями) последовательностей-мишеней и/или сайтов расщепления нуклеазами, описанных в настоящем изобретении (в том числе последовательности B2M TCAAAT (сайт D на

фиг. 1) и последовательности TCRA CCTTC, между двумя последовательностями-мишенями для TCRA-специфичной пары SBS#55254/SBS#55248).

Пример 5. Двойной нокаут B2M и TCR при адресном внедрении.

Нуклеазы, описанные выше (см. пример 4), использовали для инактивации B2M и TCRA (см. пример 5) и для адресного внедрения донорной последовательности (трансгена) в локус TCRA или B2M. В данном эксперименте TCRA-специфическая пара ZFN представляла собой SBS#55266/SBS#53853, включая последовательность TTGAAA между сайтами-мишенями TCRA-специфической ZFN (табл. 5), а пара B2M представляла собой SBS#57332/SBS#57327 (табл. 1), включая последовательность TCAAAT между сайтами-мишенями B2M-специфической ZFN. Вкратце, Т-клетки (АС-ТС-006) размораживали и активировали с использованием парамагнитных гранул CD3/28 dynabeads (при соотношении клетки:гранулы 1:3) в среде X-vivo15 для культивирования Т-клеток (0 день). Через 2 дня культивирования (2 день) к культуре клеток добавляли AAV-донорную последовательность (содержащую трансген GFP и плечи, гомологичные гену TCRA или B2M), за исключением контрольных групп, которые поддерживали без добавления донорной последовательности. На следующий день (3 день) добавляли ZFN TCRA и B2M посредством доставки мРНК в следующих 5 группах:

(a) группа 1 (только ZFN TCRA и B2M, без донорной последовательности): TCRA, 120 мкг/мл; только B2M, 60 мкг/мл;

(b) группа 2 (ZFN TCRA и B2M и донорная последовательность с плечами, гомологичными TCRA): TCRA, 120 мкг/мл; B2M, 60 мкг/мл, и AAV (TCRA-hPGK-eGFP-клон E2) 1E5vg/клетку;

(c) группа 3 (ZFN TCRA и B2M и донорная последовательность с плечами, гомологичными TCRA): TCRA, 120 мкг/мл; B2M, 60 мкг/мл, и AAV (TCRA-hPGK-eGFP-клон E2) 3E4vg/клетку;

(d) группа 4 (ZFN TCRA и B2M и донорная последовательность с плечами, гомологичными B2M): TCRA, 120 мкг/мл; B2M, 60 мкг/мл, и AAV (pAAV сайт D B2M hPGK GFP) 1E5vg/клетку;

(e) группа 5 (ZFN TCRA и B2M и донорная последовательность с плечами, гомологичными B2M): TCRA, 120 мкг/мл; B2M, 60 мкг/мл, и AAV (pAAV сайт D B2M hPGK GFP) 3E4vg/клетку.

Все эксперименты выполнили при плотности клеток 3×10^7 клеток/мл с использованием протокола, описанного в заявке на патент США № 15/347182 (сильный холодовой шок), и культивировали до холодового шока при 30°C в течение ночи после электропорации.

На следующий день (4 день) клетки разбавляли до 0,5е6 клеток/мл и переносили для культивирования при 37°C. Через три дня (7 день) клетки повторно разбавляли до 0,5е6 клеток/мл. Еще через три и семь дней культивирования (10 и 14 день, соответственно), клетки собирали для FACS- и MiSeq-анализа (при разбавлении до 0,5е6 клеток/мл).

Как показано на фиг. 5, экспрессия GFP указывала на успешное внедрение в мишень и получение генетически модифицированных клеток, содержащих модификации B2M и TCRA (инсерции и/или делеции) в сайтах нуклеаз (или в пределах 1-50 пар оснований от сайтов-мишеней нуклеаз, в том числе в пределах TTGAAA и TCAAAT и/или между парными сайтами-мишенями).

Кроме того, выполнили эксперименты, в которых в локус B2M и/или TCRA внедряли трансген CAR с получением клеток с двойным нокаутом B2M/TCRA, экспрессирующих CAR.

Все патенты, патентные заявки и публикации, упомянутые в настоящем документе, полностью включены в настоящий документ посредством ссылок.

Хотя настоящее описание представлено в некоторых подробностях с помощью иллюстраций и примеров с целью облегчения его понимания, для специалистов очевидно, что в настоящее изобретение можно внести некоторые изменения и модификации без выхода за рамки прилагаемой формулы изобретения. Соответственно, вышеизложенное описание и примеры не следует интерпретировать как ограничивающие.

ФОРМУЛА ИЗОБРЕТЕНИЯ

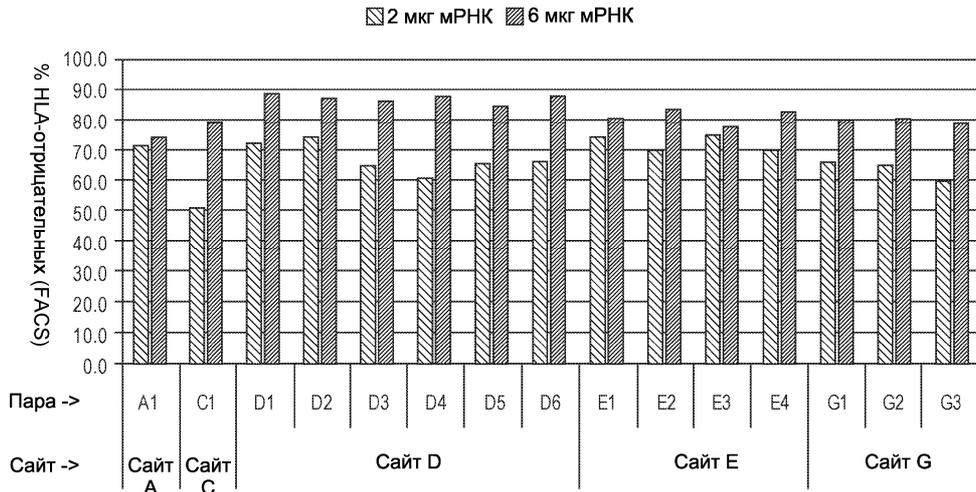
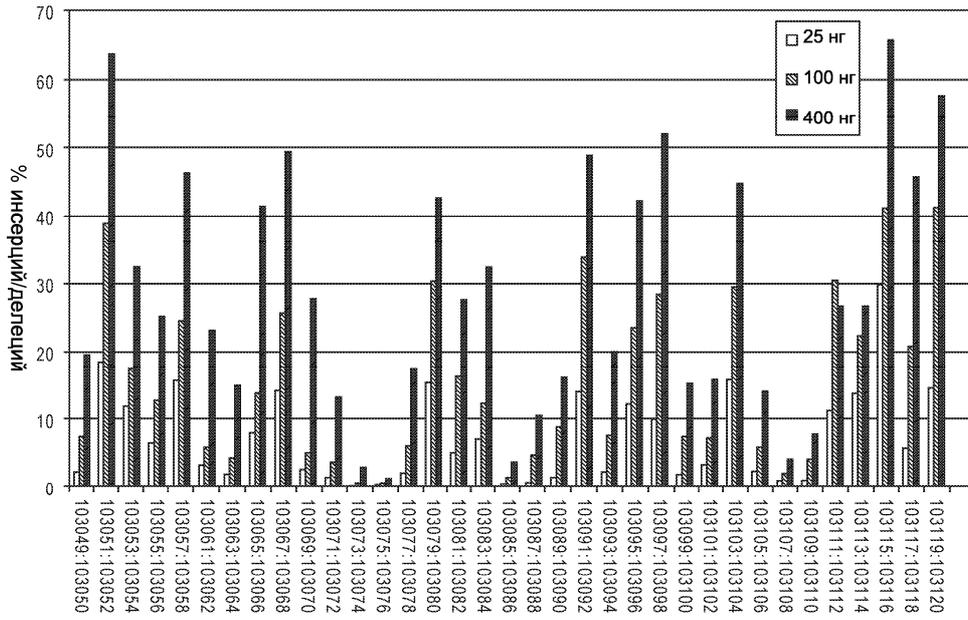
1. Выделенная клетка, в которой экспрессия гена бета-2-микроглобулина (B2M) частично или полностью инактивирована путем инсерции и/или делеции в пределах TCAAAT, GGCCTTA, TCAAATT, TTAAGTGA и/или AATTGAA 1 экзона или 2 экзона гена B2M.

2. Клетка по п.1, характеризующаяся тем, что модификация выполнена посредством экзогенной гибридной молекулы, содержащей ДНК-связывающий домен, связывающийся с сайтом-мишенью, показанным в любой из последовательностей SEQ ID NO: 16-48, и функциональный домен.

3. Клетка по п.1, характеризующаяся тем, что модификация выполнена посредством экзогенной гибридной молекулы, содержащей ДНК-связывающий домен, связывающийся с сайтом-мишенью, показанным в любой из последовательностей SEQ ID NO: 6-15, и функциональный домен.

4. Клетка по п.1, характеризующаяся тем, что модификация выполнена посредством экзогенной гибридной молекулы, содержащей ДНК-связывающий домен, связывающийся с сайтом-мишенью, показанным в любой из последовательностей SEQ ID NO: 137-205, и функциональный домен.

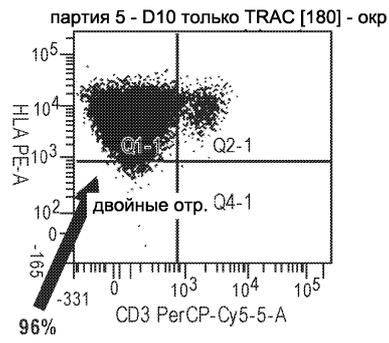
5. Клетка по любому из пп.1-4, дополнительно содержащая инактивированный ген Т-клеточного рецептора, ген PD1 и/или CTLA4.



Контроль без ZFN

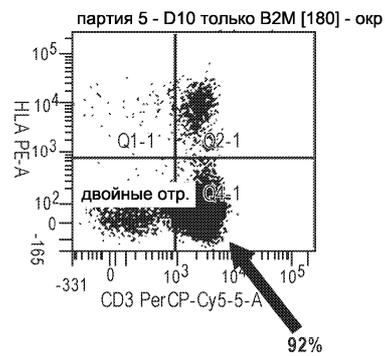


только ZFN TRAC



Фиг. 4В

только ZFN B2M

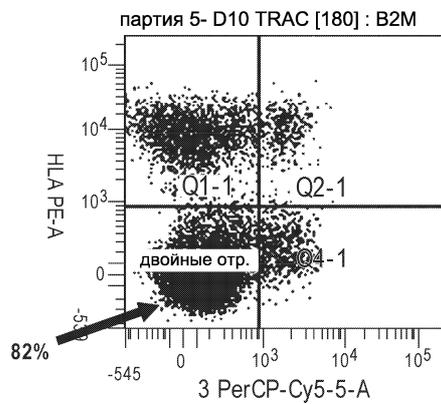


Фиг. 4С



Фиг. 4D

TRAC ZFN + B2M ZFN



Фиг. 4Е

FACS					
D10					
мкг/мл	TRAC-	B2M-	ДВОЙНЫЕ-	Общие GFP+	ДВОЙНЫЕ GFP+
фиктивная группа	0,5	0,1	0,0	0,0	0,0
Только TRAC/B2M КО	85,0	83,6	80,0	0,0	0,0
AAV-донор TRAC/B2M КО + 1E5vg/клетку, локус TRAC	91,9	92,7	89,3	80,8	83,0
AAV-донор TRAC/B2M КО + 3E4vg/клетку, локус TRAC	91,2	93,4	89,1	71,9	74,3
AAV-донор TRAC/B2M КО + 1E5vg/клетку, локус B2M	88,2	90,5	86,4	54,9	59,6
AAV-донор TRAC/B2M КО + 3E4vg/клетку, локус B2M	89,8	92,2	87,9	43,2	46,7

Фиг. 5

