



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(21) BR 112019026622-4 A2



* B R 1 1 2 0 1 9 0 2 6 6 2 2 A 2 *

(22) Data do Depósito: 15/06/2018

(43) Data da Publicação Nacional: 30/06/2020

(54) Título: DISRUPÇÃO DIRECIONADA DE RECEPTORES DE CÉLULA T E/OU HLA

(51) Int. Cl.: C12N 15/09; C12N 15/85; C12N 15/113; C12N 9/22; C12N 15/62.

(30) Prioridade Unionista: 07/08/2017 US 62/542,052; 16/06/2017 US 62/521,132; 18/10/2017 US 62/573,956.

(71) Depositante(es): SANGAMO THERAPEUTICS, INC..

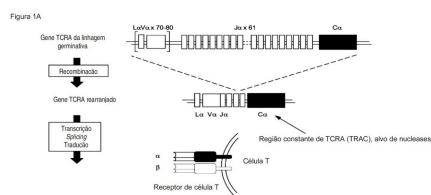
(72) Inventor(es): ANTHONY CONWAY; SUMITI JAIN; GARY K. LEE; DAVID PASCHON; EDWARD J. REBAR; LEI ZHANG.

(86) Pedido PCT: PCT US2018037844 de 15/06/2018

(87) Publicação PCT: WO 2018/232296 de 20/12/2018

(85) Data da Fase Nacional: 13/12/2019

(57) Resumo: A presente invenção refere-se a métodos e composições para inativar os genes de TCR e/ou HLA, utilizando nucleases projetadas que compreendem pelo menos um domínio de ligação ao DNA e um domínio de clivagem ou meio-domínio de clivagem em condições capazes de preservar a viabilidade celular. Além disso, são providos também polinucleotídeos codificadores de nucleases, vetores compreendendo polinucleotídeos codificadores de nucleases e células compreendendo polinucleotídeos codificadores de nucleases e/ou células compreendendo nucleases.



Relatório Descritivo da Patente de Invenção para
"DISRUPÇÃO DIRECIONADA DE RECEPTORES DE CÉLULA T E/OU HLA".

Referência cruzada a pedidos de patente correlatos

[001] O presente pedido de patente reivindica o benefício de prioridade deste ao Pedido de Patente U.S. Provisório Nº 62/521 132, depositado em 16 de junho de 2017; ao Pedido de Patente U.S. Provisório Nº 62/542,052, depositado em 07 de agosto de 2017 e ao Pedido de Patente U.S. Provisório Nº 62/573,956, depositado em 18 de outubro de 2017, cujos conteúdos são aqui incorporados, em sua totalidade, por referência neste pedido de patente.

Campo Técnico

[002] A presente invenção pertence ao campo da modificação genômica de células humanas, incluindo linfócitos e células-tronco.

Antecedente

[003] A terapia gênica encerra um potencial enorme para uma nova era de agentes terapêuticos de uso humano. Essas metodologias permitirão o tratamento de condições que não puderam ainda ser abordadas pela prática médica padrão. A terapia gênica pode incluir as muitas variações de técnicas de edição do genoma tais como disruptão (inativação) ou correção de um lócus gênico e/ou inserção de um transgene que pode ser expresso e controlado por um promotor exógeno específico operacionalmente ligado ao transgene ou pelo promotor endógeno encontrado no sítio de inserção no genoma.

[004] A entrega e inserção do transgene are exemplos de problemas que precisam ser solucionados para alguma real implementação dessa tecnologia. Por exemplo, embora haja uma variedade de métodos de entrega de genes potencialmente disponível para uso terapêuticos, todos envolvem escolhas conflitantes

substanciais entre segurança, durabilidade e nível de expressão. Métodos que fornecem o transgene como um epissoma (por exemplo, sistemas baseados em adenovírus (Ad), vírus adenoassociado (AAV) e plasmídeo) podem produzir elevados níveis de expressão iniciais, contudo, esses métodos carecem de replicação epissomal robusta, o que pode limitar a duração da expressão em tecidos com alta atividade mitótica. Em contraste, métodos de entrega que resultam na integração aleatória do transgene desejado (por exemplo, integração com lentivírus (LV)) propiciam uma expressão mais durável, mas, devido à natureza não direcionada da inserção aleatória, podem provocar o crescimento desregulado nas células receptores, potencialmente levando à malignidade pela ativação de oncogenes nas vizinhanças do cassete transgênico integrado aleatoriamente. Além disso, embora evite a perda da replicação dirigida, a integração do transgene não impede o eventual silenciamento do promotor exógeno fundido ao transgene. Com o tempo, tal silenciamento resulta em expressão reduzida do transgene para a maioria dos eventos de inserção não específica. Além disso, a integração de um transgene raramente ocorre em toda célula alvo, o que fazer com que seja difícil alcançar um nível de expressão do transgene de interesse suficientemente alto para conseguir o efeito terapêutico desejado.

[005] Nos últimos anos, foi desenvolvida uma nova estratégia para modificação genética (por exemplo, inativação, correção e/ou integração do transgene) que faz uso de clivagem com nucleases sítio-específicas (por exemplo, nucleases dedo de zinco (ZFNs, do inglês *zinc finger nucleases*), nucleases com domínio efetor do tipo ativador da transcrição (TALENs, do inglês *transcription activator-like effector domain nucleases*), sistema CRISPR/Cas com um RNA crRNA/tracr projetado (*engineered*) ("RNA guia único") para guiar a clivagem

específica, etc.) visando influenciar a edição em um lócus genômico escolhido. Ver, por exemplo, as Patentes U.S. N°s 9 937 207; 9 255 250; 9 045 763; 9 005 973; 8 956 828; 8 945 868; 8 703 489; 8 586 526; 6 534 261; 6 599 692; 6 503 717; 6 689 558; 7 067 317; 7 262 054; 7 888 121; 7 972 854; 7 914 796; 7 951 925; 8 110 379; 8 409 861; as Publicações de Patentes U.S. N°s 2017/0211075; 2003/0232410; 2005/0208489; 2005/0026157; 2005/0064474; 2006/0063231; 2008/0159996; 2010/00218264; 2012/0017290; 2011/0265198; 2013/0137104; 2013/0122591; 2013/0177983, 2013/0177960 e 2015/0056705. Além disso, estão sendo desenvolvidas nucleases direcionadas tendo por base o sistema Argonaute (por exemplo, de *T. thermophilus*, conhecido como "TtAgo", ver Swarts, et al. (2014) *Nature* 507(7491): 258-261), as quais podem igualmente ter o potencial para usos em edição genômica e terapia genética. Essa abordagem mediada por nucleases para modificação genética oferece o prospecto de expressão melhorada do transgene, segurança aumentada e durabilidade da expressão, em comparação às abordagens clássicas de integração, uma vez que permite o posicionamento exato do transgene com risco mínimo de silenciamento de genes ou ativação de oncogenes próximos.

[006] O receptor de célula T (TCR) é uma parte essencial da ativação seletiva de células T. Tendo alguma semelhança com um anticorpo, a parte de reconhecimento do antígeno do TCR é constituída tipicamente por duas cadeias, α e β, que podem se unir e formar um heterodímero. A semelhança com anticorpos reside na maneira pela qual é montado um único gene codificador de um complexo TCR alfa-beta. A cadeia alfa (TCR α) e a beta (TCR β) do TCR são compostas cada uma por duas regiões, uma região constante C-terminal e uma região variável N-terminal. Os *loci* genômicos que codificam a cadeia alfa e a beta do TCR assemelham-

se aos *loci* que codificam anticorpos pelo fato de que o gene codificador da cadeia α do TCR comprehende os segmentos V e J, enquanto o lócus da cadeia β comprehende segmentos D além de segmentos V e J. Para o lócus da cadeia β do TCR, existem adicionalmente duas regiões constantes diferentes que são selecionadas durante o processo de seleção. Durante o desenvolvimento de células T, os vários segmentos recombinam de tal modo que cada célula T comprehende uma porção variável única do TCR nas cadeias alfa e beta, denominadas a região determinante de complementaridade (CDR), e o corpo possui um grande repertório de células T, devido às suas CDRs únicas, que são capazes de interagir com antígenos únicos exibidos pelas células apresentadoras de antígenos. Uma vez tenha ocorrido um rearranjo do gene da cadeia α ou β do TCR, a expressão do segundo TCR α ou TCR β correspondente é reprimida de tal modo que cada célula T expressa somente uma estrutura única do TCR em um processo denominado "exclusão alélica de receptor do antígeno" (ver, Brady, et al. (2010) *J Immunol* 185:3801-3808).

[007] Durante a ativação de células T, o TCR interage com antígenos exibidos como peptídeos no principal complexo de histocompatibilidade (MHC) de uma célula apresentadora de antígeno. O reconhecimento do complexo antígeno-MHC pelo TCR leva à estimulação de células T, o que, por sua vez, leva à diferenciação de células T *helper* (CD4+) e linfócitos T citotóxicos (CD8+) em linfócitos de memória e efetores. Essas células podem então expandir de maneira clonal, fornecendo uma subpopulação ativada dentro da população total de células T, capaz de reagir a um antígeno em particular.

[008] As proteínas do MHC são de duas classes, I e II. As proteínas do MHC classe I MHC são heterodímeros de duas proteínas,

a cadeia α, que é uma proteína transmembrana codificada pelos genes de MHC1 da classe I, e a cadeia β2-microglobulina (às vezes designada como B2M), que é uma pequena proteína extracelular codificada por um gene que não está situado dentro do *cluster* (agrupamento) de genes do MHC. A cadeia α dobra-se em três domínios globulares e, quando a cadeia β2-microglobulina está associada ao complexo com estrutura globular é funcional e expresso na superfície da célula. Os peptídeos são apresentados nos dois domínios mais N-terminais que são também os mais variáveis. As proteínas do MHC Classe II são igualmente heterodímeros, mas os heterodímeros compreendem duas proteínas transmembrana codificadas por genes dentro do complexo do MHC. O complexo MHC classe I: antígeno interage com células T citotóxicas enquanto o MHC classe II apresenta antígenos às células T *helper*. Além disso, as proteínas do MHC classe I tendem a ser expressas em quase todas as células nucleadas e em plaquetas (e eritrócitos em camundongos) enquanto as proteínas do MHC classe II são expressas de modo mais seletivo. Tipicamente, as proteínas do MHC classe II são expressas em células B, alguns macrófagos e monócitos, células de Langerhans e em células dendríticas.

[009] Nos seres humanos, o complexo principal de histocompatibilidade (MHC) é conhecido comumente como o antígeno leucocitário humano (HLA). O *cluster* gênico do HLA classe I nos humanos compreende três *loci* principais, B, C e A, bem como diversos *loci* menores (incluindo E, G e F, todos encontrados na região do HLA no cromossomo 6). O *cluster* do HLA classe II também compreende três *loci* principais, DP, DQ e DR, e ambos os *clusters* de genes da classe I e classe II são polimórficos, pelo fato de que existem diversos alelos diferentes de ambos os genes da classe I e II dentro da população. Existem igualmente diversas proteínas acessórias que

participam no funcionamento do HLA também. β -2 microglobulina atua como uma chaperona (codificada por B2M, localizado no cromossomo 15) e estabiliza a proteína HLA A, B ou C expressa na superfície celular, além de estabilizar também o sulco de exibição do antígeno na estrutura da classe I. É encontrada no soro e na urina em quantidades baixas normalmente.

[0010] HLA desempenha um papel importante na rejeição ao transplante. A fase aguda da rejeição ao transplante pode ocorrer dentro de aproximadamente 1-3 semanas e normalmente envolve a ação de linfócitos T do hospedeiro sobre tecidos do doador devido à sensibilização do sistema do hospedeiro às moléculas de HLA classe I e classe II do doador. Na maioria dos casos, os抗ígenos desencadeantes são HLAs classe I. Para obter melhor sucesso, é efetuada a tipagem HLA do doador, cuja compatibilidade com a tipagem do paciente receptor deve ser a mais completa possível. No entanto, mesmo a doação entre familiares, que podem compartilhar um alto percentual de identidade HLA, ainda assim frequentemente não é bem-sucedida. Desse modo, a fim de preservar o tecido enxertado no receptor, o paciente frequentemente precisa ser submetido à terapia imunossupressora profunda para prevenir a rejeição. Tal terapia pode levar a complicações e morbidades significantes devido a infecções oportunistas que o paciente pode ter dificuldade em superar. A regulação dos genes classe I ou II pode ser alterada na presença de alguns tumores, e tal disruptão pode ter consequências sobre o prognóstico dos pacientes. Por exemplo, foi verificada redução da expressão de B2M em cânceres colorretais metastáticos (Shrout, et al. (2008) *Br J Canc* 98:1999). Considerando que B2M desempenha um papel fundamental na estabilização do complexo do MHC classe I, foi levantada a hipótese de que a perda de B2M em certos cânceres sólidos poderia ser um mecanismo de

escape à vigilância imunológica dirigida por células T. A expressão deprimida de B2M demonstrou resultar de supressão da regulação da expressão normal de BM2 para IFN gama e/ou de mutações específicas na sequência codificadora de B2M que resultam em *knock-out* (perda de função) do gene (Shrout, et al., *ibid*). Como fator de confundimento, B2M aumentado também está associado com alguns tipos de câncer. Níveis aumentados de B2M na urina serve como indicador prognóstico para diversos cânceres incluindo câncer de próstata, leucemia linfocítica crônica (LLC) e linfomas não Hodgkin.

[0011] A terapia celular adotiva (ACT) é uma forma de terapia antineoplásica em desenvolvimento que tem por base a entrega de células imunes tumor-específicas a um paciente para que as células entregues ataquem e eliminem o câncer do paciente. A ACT pode envolver o uso de linfócitos infiltrantes de tumor (TILs) que são células T isoladas de massas tumorais do próprio paciente e expandidas ex vivo para voltarem a ser infundidas no paciente. Essa abordagem tem sido promissora no tratamento de melanoma metastático, onde, em um estudo, uma taxa de resposta de longo prazo >50% foi observada (ver, por exemplo, Rosenberg, et al. (2011) *Clin Canc Res* 17(13): 4550). TILs representam uma fonte promissora de células porque são um conjunto misto das próprias células do paciente que possuem receptores de células T (TCRs) específicos para os抗ígenos associados a tumores (TAAs) presentes no tumor (Wu, et al. (2012) *Cancer J*, 18(2):160). Outras abordagens envolvem a edição de células T isoladas do sangue de um paciente, de tal modo que são projetadas para que respondam ao tumor de alguma maneira (Kalos, et al. (2011) *Sci Transl Med* 3(95):95ra73).

[0012] Receptores quiméricos de抗ígenos (CARs) são moléculas desenhadas para direcionar as células imunes para alvos moleculares específicos expressos nas superfícies celulares. Na sua forma mais

básica, são receptores introduzidos em uma célula que acoplam um domínio de especificidade expresso no exterior da célula para sinalização de vias no interior da célula, de tal modo que, quando o domínio de especificidade interage com seu alvo, a célula torna-se ativada. Frequentemente, os CARs são produzidos por simulação dos domínios funcionais de receptores de células T (TCRs), em que um domínio específico para um antígeno, como um scFv ou algum tipo de receptor, é fundido ao domínio de sinalização, como ITAMs e outros domínios coestimulatórios. Essas construções são então introduzidas em uma célula T *ex vivo* permitindo à célula T tornar-se ativada na presença de uma célula que expressa o antígeno alvo, resultando no ataque sobre a célula visada pela célula T ativada de maneira não dependente de MHC (ver Chicaybam, *et al.* (2011) *Int Rev Immunol* 30:294-311) quando a célula T é reintroduzida no paciente. Assim, a terapia celular adotiva utilizando células T alteradas *ex vivo* com um TCR ou CAR projetado é uma abordagem clínica muito promissora para diversos tipos de doenças. Por exemplo, os cânceres e seus抗ígenos que constituem alvos incluem linfoma folicular (CD20 ou GD2), neuroblastoma (CD171), linfoma não Hodgkin (CD19 e CD20), linfoma (CD19), glioblastoma (IL13Ra2), leucemia linfocítica crônica ou LLC e leucemia linfocítica aguda ou LLA (ambas CD19). Foram também desenvolvidos CARs vírus-específicos para atacar células que abrigam vírus como HIV. Por exemplo, foi iniciado um estudo clínico utilizando um CAR específico para Gp100 para o tratamento de HIV (Chicaybam, *ibid*).

[0013] ACTRs (Receptores de células T acopladas a anticorpo) são componentes projetados de células T que são capazes de se ligar a um anticorpo exógeno. A ligação do anticorpo ao componente ACTR arma a célula T para interagir com o antígeno reconhecido pelo anticorpo e, quando aquele antígeno é encontrado, a célula T

compreendendo o ACTR é acionada para interagir com o antígeno (ver Publicação de Patente U.S. N° 2015/0139943).

[0014] Uma das desvantagens da terapia celular adotiva, contudo, é que a fonte do produto celular precisa ser específica do paciente (autóloga) para evitar a rejeição potencial das células transplantadas. Isso levou os pesquisadores a desenvolverem métodos para edição das próprias células T de um paciente para evitar essa rejeição. Por exemplo, células T ou células-tronco hematopoiética de um paciente podem ser manipuladas *ex vivo* com a adição de um CAR, ACTR e/ou receptor de célula T (TCR) projetados e, a seguir, tratadas ainda com nucleases projetadas para desabilitar inibidores de *checkpoints* de células T como PD1 e/ou CTLA4 (ver Publicação da Patente Internacional N° WO 2014/059173). Para aplicação dessa tecnologia a uma população maior de pacientes, seria vantajoso desenvolver uma população universal de células (alogênicas). Além disso, o *knockout* do TCR resultará em células que são incapazes de montar uma resposta de doença do enxerto contra hospedeiro (DECH) introduzida em um paciente.

[0015] Assim, ainda há necessidade de métodos e composições que possam ser utilizados para modificar (por exemplo, *knockout*) a expressão de TCR e/ou HLA em células T efetoras, células T reguladoras, células B, células NK ou células-tronco (por exemplo, células-tronco hematopoiéticas, células-tronco pluripotentes induzidas e células-tronco embrionárias).

[0016] Sumário

[0017] São aqui descritas composições e métodos para inativação parcial ou completa ou disruptão de um gene TCR e/ou B2M e composições e métodos para introduzir e expressar níveis desejados de transgenes exógenos em linfócitos T, depois ou simultaneamente com a disruptão do TCR e/ou B2M endógenos. Além disso, são

igualmente providos métodos e composições para deletar (inativar) ou reprimir um gene TCR e/ou B2M para produzir célula T com TCR nulo ou célula T com TCR e HLA classe I nulos, célula B, célula NK, célula-tronco, tecido ou organismo inteiro, por exemplo, uma célula que não expressa um ou mais receptores de células T e/ou um ou mais receptores de HLA classe I sem sua superfície. Modificações genômicas adicionais podem estar presentes nas células com TCR e/ou HLA classe I nulos aqui descritas, mas não se limitam a modificações genômicas a um gene diferente (por exemplo, um gene da proteína 1 de morte celular programada (PD1), um gene do antígeno 4 do linfócito T citotóxico (CTLA-4), um gene CISH, um gene tet2, um gene do antígeno leucocitário humano (HLA) A, um gene HLA B, um gene HLA C, um gene HLA-DPA, um gene HLA-DQ, um gene HLA-DRA, um gene LMP7, um gene do transportador associado ao processamento de antígeno (TAP) 1, um gene TAP2, um gene da tapasina (TAPBP), um gene do transativador do complexo principal de histocompatibilidade classe II (CIITA), um gene do receptor glicocorticoide (GR), um gene IL2RG, um gene RFX5), à inserção de transgene (por exemplo, CAR) em um ou mais desses genes ou de outros (por exemplo, genes em *loci safe harbor*) e qualquer combinação de tais modificações genômicas. Em certas modalidades, as células com TCR nulo e/ou as células com HLA classe I nulo ou tecidos são células ou tecidos humanos cujo uso é vantajoso em transplantes. Em modalidades preferidas, as células com TCR nulo e/ou as células com HLA classe I nulo são preparadas para uso em terapia com células T adotivas.

[0018] Em um aspecto, é descrita uma nuclease de dedo de zinco compreendendo: uma proteína de dedo de zinco (ZFP) de uma ZFN designada 68957, 72678, 72732 ou 72748; um domínio de clivagem FokI projetado; e um ligante (elemento de ligação) entre o domínio de

clivagem FokI e a ZFP. Em certas modalidades, a ZFN comprehende uma primeira e uma segunda ZFNs como segue (sequências de aminoácidos e polinucleotídeos descritas nos Exemplos): uma ZFN comprehendendo uma ZFP da ZFN designada 72678 e uma ZFN comprehendendo uma ZFP da ZFN designada 72732. Em certas modalidades, a ZFN comprehende ZFNs direita e esquerda (primeira e segunda) como segue: uma ZFN designada 57531 e uma ZFN designada 72732; uma ZFN designada 57531 e uma ZFN designada 72748; uma ZFN designada 68957 e uma ZFN designada 57071; uma ZFN designada 68957 e uma ZFN designada 72732; uma ZFN designada 68957 e uma ZFN designada 72748; uma ZFN designada 72678 e uma ZFN designada 57071; uma ZFN designada 72678 e uma ZFN designada 72732; e uma ZFN comprehendendo uma ZFP designada 72678 e uma ZFN designada 727482. Uma nuclease de dedo de zinco (ZFN) comprehendendo ZFNs direita e esquerda (primeira e segunda): uma ZFN designada 68796 e uma ZFN designada 68813; uma ZFN designada 68796 e uma ZFN designada 68861; uma ZFN designada 68812 e uma ZFN designada 68813; uma ZFN designada 68876 e uma ZFN designada 68877; uma ZFN designada 68815 e uma ZFN designada 55266; uma ZFN designada 68879 e uma ZFN designada 55266; uma ZFN designada 68798 e uma ZFN designada 68815; ou uma ZFN designada 68846 e uma ZFN designada 53853. Além disso, são igualmente aqui providos polinucleotídeos (por exemplo, mRNA, plasmídeos, vetores virais, etc.) que codificam uma ZFN (incluindo um par), conforme aqui revelada, inclusive um polinucleotídeo comprehendendo uma sequência 2A entre as sequências que codificam a ZFN esquerda e as ZFNs. Além disso, são descritas células geneticamente modificadas (por exemplo, células-tronco, células precursoras, células T (efetoras e reguladoras), etc.) comprehendendo um ou mais das ZFNs e/ou dos polinucleotídeos.

aqui descritos e células descendentes dessas células (por exemplo, células geneticamente modificadas que não compreendem a ZFN, mas que incluem a modificação genética). As modificações genéticas incluem inserções, deleções e combinações das mesmas no gene visado pela ZFN. Modificações genômicas adicionais, por exemplo, modificação de um gene do receptor de célula T (TCR), modificação de um gene HLA-A, modificação de um gene HLA-B, modificação de um gene HLA-C, modificação de um gene TAP, modificação de um gene CTLA-4, modificação de um gene PD1, modificação de um gene CISH, modificação de um gene tet-2 e/ou a inserção de um transgene (por exemplo, CAR) pode estar presente no alvo e/ou em um ou mais *loci* diferentes. Composições farmacêuticas compreendendo qualquer uma das nucleases de dedo de zinco, polinucleotídeos e/ou células, como aqui descritos, são também providas no presente. Métodos para modificar um gene endógeno de beta-3-microglobulina (B2M) e/ou gene TCR em uma célula são igualmente providos, o método compreendendo administrar um polinucleotídeo ou composição farmacêutica, conforme aqui descritos, à célula de tal modo que o gene endógeno seja modificado (por exemplo, deleção, inserção de uma sequência exógena como um transgene). São também descritos métodos de utilização das ZFNs, polinucleotídeos, células e/ou composições farmacêuticas, conforme aqui descritos, para o tratamento e/ou a prevenção de um câncer, uma doença autoimune ou doença do enxerto contra hospedeiro. Kits compreendendo qualquer uma das ZFNs, polinucleotídeos, células e/ou composições farmacêuticas, conforme aqui descritos, são igualmente providos.

[0019] Em outros aspectos, é aqui descrita uma célula isolada (por exemplo, uma célula eucariótica, como uma célula de mamífero, incluindo uma célula linfoide, uma célula-tronco (por exemplo, iPSC, célula-tronco embrionária, MSC ou HSC) ou uma célula

progenitora/precursora) na qual a expressão de um gene do TCR é modulada por modificação de sequências exônicas do gene do TCR. Em certas modalidades, a modificação é a uma sequência compreendendo uma sequência de 9-25 (inclusive sítios alvo de 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25) ou mais nucleotídeos (contíguos ou não) de uma sequência como mostrada nos sítios alvo no presente) de um sítio alvo mostrado em uma ou mais das Tabelas 1, 2 ou 6 (SEQ ID Nº: 8-21 e/ou 92-103); dentro de 1-5, dentro de 1-10 ou dentro 1-20 pares de base em qualquer dos lados (a sequência genômica flanqueadora) dos sítios alvo mostrados nas Tabelas 1, 2 ou 6 (SEQ ID Nº:8-21 e/ou 92-103); ou dentro de AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e AATCCTC ou um sítio alvo compreendendo AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e AATCCTC. Alternativamente ou além de, as modificações podem também ser efetuadas a sequências (por exemplo, sequências genômicas) entre sítios alvo pareados conforme aqui descritos (por exemplo, sítios alvo para os pares de nucleases mostrados na Tabela 3, incluindo entre os sítios alvo para 55204 e 53759 (entre SEQ ID Nº:8 e SEQ ID Nº:9); entre os sítios alvo para 55229 e 53785 (entre SEQ ID Nº:10 e SEQ ID Nº:11); entre os sítios alvo para 53810 e 55255 (entre SEQ ID Nº:12 e SEQ ID Nº:13); entre os sítios alvo mostrados para 55248 e 55254/55260 (entre SEQ ID Nº:14 e SEQ ID Nº:13); entre os sítios alvo para 55266 e 53853 (entre SEQ ID Nº:15 e SEQ ID Nº:16); entre os sítios alvo para 53860 e 53863 (entre SEQ ID Nº:17 e SEQ ID Nº:18); entre os sítios alvo para 53856 e 55287 (entre SEQ ID Nº:21 e SEQ ID Nº:18); ou entre os sítios alvo para 53885 ou 52774 e 53909 ou 52742 (entre SEQ ID Nº:19 e SEQ ID Nº:20). A modificação pode ser por uma molécula de fusão exógena compreendendo um domínio funcional (por exemplo, domínio regulador da transcrição, domínio nuclease incluindo qualquer

domínio de clivagem FokI com uma ou mais mutações em comparação ao tipo selvagem) e um domínio de ligação ao DNA, incluindo, entre outros: (i) uma célula compreendendo um fator de transcrição exógeno incluindo um domínio de ligação ao DNA que se liga a um sítio alvo mostrado em qualquer uma dentre SEQ ID Nº:8-21 e/ou 92-103 e um domínio regulador da transcrição no qual o fator de transcrição modifica a expressão do gene TRAC e/ou (ii) uma célula compreendendo uma inserção e/ou uma deleção dentro de um ou mais dos sítios alvo aqui mostrados, incluindo SEQ ID Nº:8-21 e/ou 92-103; dentro de 1-5, dentro de 1-10 ou dentro de 1-20 pares de base em qualquer dos lados (a sequência genômica flanqueadora) dos sítios alvo mostrados nas Tabelas 1 e 2 (SEQ ID Nº: 8-21 e/ou 92-103); dentro de AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e AATCCTC; e/ou entre sítios alvo pareados conforme aqui descritos (por exemplo, sítios alvo para os pares de nucleases mostrados na Tabela 3). São igualmente descritas células que compreendem essas modificações ao(s) genes do TCR e modificações genéticas adicionais (por exemplo, modificação no gene B2M, modificações nos genes CTLA, CISH, PD1 e/ou tet2, CAR, um TCR antígeno-específico (cadeias alfa e beta), inserções nesses ou em outros *loci* incluindo um transgene codificador de um receptor de Célula T acoplada a um Anticorpo (ACTR) e/ou um transgene codificador de um anticorpo, etc.).

[0020] Em outro aspecto, é descrita uma célula isolada (por exemplo, uma célula eucariótica como uma célula de mamífero, incluindo uma célula linfoide, uma célula-tronco (por exemplo, iPSC, célula-tronco embrionária, MSC ou HSC ou uma célula progenitora/precursora) na qual a expressão de um gene B2M gene é modulada por modificação no gene B2M. Em certas modalidades, a modificação é a uma sequência compreendendo uma sequência de 9-

25 (incluindo sítios alvo de 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25) ou mais nucleotídeos (contíguos ou não) de uma sequência como mostrada nos sítios alvo no presente) de um sítio alvo mostrado em uma ou mais das Tabelas 5 e 8 (SEQ ID Nº:117, 123, 126 e/ou 127); dentro de 1-5, dentro de 1-10 ou dentro de 1-20 pares de base em qualquer dos lados (a sequência genômica flanqueadora) dos sítios alvo mostrados nas Tabelas 5 e 8 (SEQ ID Nº:117, 123, 126 e/ou 127). Alternativamente ou além de, as modificações podem também ser efetuadas a sequências (por exemplo, sequências genômicas) entre sítios alvo pareados conforme aqui descritos (por exemplo, sítios alvo para os pares de nucleases mostrados nas Tabelas 5 e 8, incluindo entre os sítios alvo como mostrado na Tabela 8 (SEQ ID Nº:126 e 127). A modificação pode ser por uma molécula de fusão exógena compreendendo um domínio funcional (por exemplo, domínio regulador da transcrição, domínio nuclease incluindo qualquer domínio de clivagem FokI com uma ou mais mutações em comparação ao tipo selvagem) e um domínio de ligação ao DNA (por exemplo, uma ZFP como mostrada na Tabela 8 (o componente ZFP (desenhos) das ZFNs designadas 72732; 72748; 68957; ou 72678), incluindo, entre outros: (i) uma célula compreendendo um fator de transcrição exógeno incluindo um domínio de ligação ao DNA que se liga a um sítio alvo como mostrado em qualquer uma das Tabelas 5 ou 8 (por exemplo, SEQ ID Nº:126 ou 127) e um domínio regulador da transcrição em que o fator de transcrição modifica a expressão do gene B2M e/ou (ii) uma célula compreendendo uma inserção e/ou uma deleção dentro de um ou mais dos sítios alvo aqui mostrados, incluindo as Tabelas 5 e 8; dentro de 1-5, dentro de 1-10 ou dentro de 1-20 pares de base em qualquer dos lados (a sequência genômica flanqueadora); e/ou entre sítios alvo pareados conforme aqui descritos (por exemplo, sítios alvo para os pares de nucleases mostrados na

Tabela 8). São igualmente descritas células que compreendem essas modificações a genes B2M e modificações genéticas adicionais (por exemplo, modificação ao gene TCR, modificações aos genes CTLA, CISH, PD1 e/ou tet2, modificação a PD1, uma de CAR, um TCR antígeno-específico (cadeias alfa e beta), inserções nesses ou em outros *loci* incluindo um transgene codificador de um receptor de célula T acoplada a anticorpo (ACTR) e/ou um transgene codificador de um anticorpo, etc.).

[0021] As células com TCR e/ou B2M modificados aqui descritas podem incluir mais modificações, por exemplo, um ou mais genes inativados do receptor de célula T em células com B2M modificado, genes TCR inativados adicionais, gene PD1 e/ou CTLA4 e/ou um transgene codificador de um receptor quimérico de antígeno (CAR), um transgene codificador de um receptor de célula T acoplada a anticorpo (ACTR) e/ou um transgene codificador de um anticorpo. Composições farmacêuticas que compreendem qualquer célula como aqua descrita são também providas, bem como métodos de utilização das células e composições farmacêuticas em terapias *ex vivo* para o tratamento de um transtorno (por exemplo, um câncer) em um indivíduo. Em certas modalidades, é provida uma população de células compreendendo uma ou mais modificações (TCR editado, B2M editado, PD1 editado, CISH, tet2 e/ou CTLA4 editados, gene HLA classe I editado e/ou inserções de transgene (por exemplo, CAR) nesses ou em outros genes, etc.) conforme aqua descrita, incluindo uma população de células na qual menos de 5% (por exemplo, 0-5% ou qualquer valor intermediário), de preferência menos de 3%, ainda mais preferivelmente menos de 2% das células incluem quaisquer outras modificações (por exemplo, modificações fora dos sítios alvo). Em certas modalidades, a população de células inclui modificações fora de sítios alvo em níveis naturais (*background*) (por exemplo, 2-10 vezes menos (ou qualquer

valor intermediário)) quando comparadas a células modificadas com ZFNs que não são modificadas como aqui descrito (essas ZFNs não modificadas são referidas também como ZFNs "progenitoras" ou "parentais"). As modificações efetuadas pelas ZFNs são hereditárias pelo fato de que, *in vivo ou* em cultura, as células que descendem de (incluindo células diferenciadas) células compreendendo as ZFNs (e modificações) incluem as modificações aqui descritas.

[0022] Assim, em um aspecto, são descritas células nas quais a expressão de um gene TCR é modulada (por exemplo, ativada, reprimida ou inativada). Em modalidades preferidas, sequências exônicas de um gene TCR são moduladas. A modulação pode ser por uma molécula exógena (por exemplo, fator de transcrição projetado compreendendo um domínio de ligação ao DNA e um domínio de ativação ou repressão transcrional) que se liga ao gene TCR e regula a expressão de TCR e/ou através de modificação na sequência do gene TCR (por exemplo, utilizando uma nuclease que cliva o gene TCR e modifica a sequência do gene por inserções e/ou deleções), incluindo por exemplo, uma ZFN (por exemplo, par de ZFNs de ZFNs esquerda e direita) como mostrado na Tabela 6. Em algumas modalidades, são descritas células que compreendem uma nuclease projetada para provocar *knockout* de um gene TCR. Em outras modalidades, são descritas células que compreendem um fator de transcrição (TF) projetado de tal modo que a expressão de um gene TCR é modulada. Em algumas modalidades, as células são células T. São descritas ainda células em que a expressão de um gene TCR é modulada e em que as células são projetadas adicionalmente para que compreendam pelo menos um transgene exógeno e/ou *knockout* adicional de pelo menos um gene endógeno (por exemplo, gene beta 2 microglobulina (B2M) e/ou de *checkpoint* imunológico como PD1 e/ou CTLA4) ou combinações dos mesmos.

[0023] Em outro aspecto, são aqui descritas células nas quais a expressão de um gene B2M é modulada (por exemplo, ativada, reprimida ou inativada). A modulação pode ser por uma molécula exógena (por exemplo, fator de transcrição projetado compreendendo um domínio de ligação ao DNA e um domínio de ativação ou repressão transcricional) que se liga ao gene B2M e regula a expressão de B2M e/ou através de modificação na sequência do gene B2M (por exemplo, utilizando uma nuclease que cliva o gene B2M e modifica a sequência do gene por inserções e/ou deleções), incluindo por exemplo, uma ZFN (por exemplo, par de ZFNs de ZFNs esquerda e direita) como mostrado na Tabela 8 ou uma ZFN compreendendo uma ZFP tendo o desenho (região da hélice de reconhecimento e esqueleto (*backbone*) de ZFPs nas ZFNs designadas 72732; 72748; 68957; ou 72678) aqui descrito (por exemplo, Tabela 8) em combinação com qualquer domínio FokI (tipo selvagem ou projetado) e opcionalmente qualquer ligante entre o domínio FokI e a ZFP (por exemplo, L0, N7a, N7c, etc.). Em algumas modalidades, são descritas células que compreendem uma nuclease projetada para provocar *knockout* de um gene B2M. Em outras modalidades, são descritas células que compreendem um fator de transcrição (TF) projetado de tal modo que a expressão de um gene B2M é modulada. Em algumas modalidades, as células são células T, incluindo células T efetoras e células T reguladoras. São ainda descritas células em que a expressão de um gene B2M é modulada e em que as células são ainda projetadas para compreender pelo menos um transgene exógeno e/ou *knockout* adicional de pelo menos um gene endógeno (por exemplo, um ou mais genes TCR e/ou gene de *checkpoint* imunológico como PD1 e/ou CTLA4) ou combinações dos mesmos.

[0024] Em qualquer uma das células aqui descritas que compreendem um transgene exógeno, o transgene exógeno pode ser

integrado em um gene TCR e/ou B2M (por exemplo, quando o gene TCR e/ou B2M é desabilitado) e/ou pode ser integrado em um gene como um gene *safe harbor*. Em alguns casos, o transgene exógeno codifica um ACTR, um TCR antígeno-específico e/ou um CAR. O transgene projetado pode ser inserido por processos orientado por HDR ou NHEJ. Em alguns aspectos, as células com expressão modulada de TCR e/ou B2M compreendem pelo menos um ACTR exógeno, um TCR exógeno e um CAR exógeno. Algumas células compreendendo um modulador de TCR incluem ainda *knockout* de um ou mais genes de inibidores de pontos de controle (*checkpoint*). Em algumas modalidades, o inibidor de *checkpoint* é PD1. Em outras modalidades, o inibidor de *checkpoint* é CTLA4. Em aspectos adicionais, a célula com TCR e/ou B2M modulado compreende *knockout* (silenciamento) de PD1 e *knockout* de CTLA4. Em algumas modalidades, o gene TCR modulado é um gene codificador de TCR β (TCRB). Em algumas modalidades, isso é alcançado pela clivagem direcionada da região constante desse gene (região constante de TCR β ou TRBC). Em certas modalidades, o gene TCR modulado é um gene codificador de TCR α (TCRA). Em modalidades adicionais, a inserção é conseguida pela clivagem direcionada da região constante de um gene TCR, incluindo clivagem direcionada da região constante de um gene TCR α (aqui designada como sequências "TRAC"). Em algumas modalidades, as células com gene TCR modificado são modificadas ainda no gene B2M, os genes HLA-A, B, C ou o gene TAP, ou qualquer combinação dos mesmos. Em outras modalidades, o regulador do HLA classe II, CIITA, é modificado também.

[0025] Em certas modalidades, as células aqui descritas compreendem uma modificação (por exemplo, deleção e/ou inserção, ligação de um TF projetado para reprimir a expressão de TCR) a um gene TCRA (por exemplo, modificação de exons). Em certas

modalidades, a modificação é dentro de qualquer um dos sítios alvo mostrados nas Tabelas 1, 2 ou 6 (SEQ ID Nº:8-21 e/ou 92-103) e/ou entre sítios alvo pareados (por exemplo, sítios alvo de pares de nucleases mostrados na Tabela 3), incluindo modificação por ligação a, clivagem, inserção e/ou deleção de um ou mais nucleotídeos dentro de qualquer uma dessas sequências e/ou dentro de 1-50 pares de base (incluindo qualquer valor intermediário como 1-5, 1-10 ou 1-20 pares de base) das sequências do gene (genômicas) flanqueando essas sequências no gene TCRA. Em certas modalidades, as modificações são efetuadas utilizando uma ZFN (por exemplo, um ou mais pares de ZFNs) como mostrada na Tabela 6. Em certas modalidades, as células compreendem uma modificação (ligação a, clivagem, inserções e/ou deleções) dentro de uma ou mais das sequências seguintes: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e AATCCTC dentro de um gene TCRA (por exemplo, exons, ver Figura 1B). Em certas modalidades, a modificação compreende a ligação de um TF projetado aqui descrito de tal modo que a expressão de um gene TCRA é modulada, por exemplo, reprimida ou ativada.

[0026] Em certas modalidades, as células aqui descritas compreendem uma modificação (por exemplo, deleção e/ou inserção, ligação de um TF projetado para reprimir a expressão de B2M) a um gene B2M. Em certas modalidades, a modificação é dentro de qualquer um dos sítios alvo mostrados nas Tabelas 5 ou 8 e/ou entre sítios alvo pareados (por exemplo, sítios alvo de pares de nucleases mostrados na Tabela 8), incluindo a modificação por ligação a, clivagem, inserção e/ou deleção de um ou mais nucleotídeos dentro de qualquer uma dessas sequências e/ou dentro de 1-50 pares de base (incluindo qualquer valor intermediário como 1-5, 1-10 ou 1-20 pares de base) das sequências do gene (genômicas) flanqueando essas

sequências no gene B2M. Em certas modalidades, as modificações são efetuadas utilizando uma ZFN que comprehende uma ZFP incluindo as regiões de reconhecimento de hélice e esqueleto dos desenhos de ZFP das ZFNs mostradas na Tabela 8, um domínio FokI (qualquer domínio FokI do tipo selvagem ou projetado) e opcionalmente um ligante (qualquer ligante entre o N- ou C-terminal do domínio FokI e o N- ou C-terminal dos desenhos de ZFP mostrados, incluindo, entre outros, L0, N7a, N7c, etc.). Em certas modalidades, a ZFN comprehende uma ZFN (por exemplo, um par de primeira e segunda ZFNs) como mostrada na Tabela 8. Em certas modalidades, as células comprehendem uma modificação (ligação a, clivagem, inserções e/ou deleções) dentro de uma ou mais das sequências seguintes: SEQ ID Nº:126 e 127. Em certas modalidades, a modificação comprehende a ligação de um TF projetado conforme aqui descrito de tal modo que a expressão do gene B2M é modulada, por exemplo, reprimida ou ativada.

[0027] Em outras modalidades, a modificação é uma modificação genética (alteração da sequência de nucleotídeos) em ou próximo do(s) sítio(s) de ligação a nuclease(s) (alvo) e/ou de clivagem, incluindo, entre outras, modificações a sequências dentro de 1-300 (ou qualquer número intermediário de pares de base) pares de base a montante, a jusante e/ou incluindo 1 ou mais pares de base do(s) sítio(s) de clivagem e/ou sítio de ligação; modificações dentro de 1-100 pares de base (ou qualquer número intermediário de pares de base) incluindo e/ou em qualquer dos lados do(s) sítio(s) de ligação e/ou clivagem; modificações dentro de 1 a 50 pares de base (ou qualquer número intermediário de pares de base) incluindo e/ou em qualquer dos lados (por exemplo, 1 a 5, 1 a 10, 1 a 20 ou mais pares de base) do(s) sítio(s) de ligação e/ou clivagem; e/ou modificações a um ou mais pares de base dentro do sítio de ligação e/ou sítio de clivagem da

nuclease. Em certas modalidades, a modificação é em ou próximo de (por exemplo, 1-300 pares de base, 1-50, 1-20, 1-10 ou 1-5 ou qualquer número intermediário de pares de base) e/ou entre sítios alvo pareados (por exemplo, Tabela 3 ou 8) da sequência do gene, em volta ou entre qualquer um dos sítios alvo aqui revelados. Em certas modalidades, a modificação inclui modificações de um gene TCRA e/ou B2M dentro de uma ou mais das sequências mostradas nos sítios alvo das Tabelas 1, 2 e 6 (TCRA) e/ou Tabelas 5 e 8 (B2M), por exemplo, uma modificação de 1 ou mais pares de base a uma ou mais dessas sequências. Em certas modalidades, as modificações genéticas mediadas pela nuclease são entre sítios alvo pareados (quando um dímero é utilizado para clivar o alvo). As modificações genéticas mediadas pela nuclease podem incluir inserções e/ou deleções de qualquer número de pares de base, incluindo inserções de sequências não codificantes de qualquer comprimento e/ou transgenes de qualquer comprimento e/ou deleções de 1 par de bases até mais de 1000 kb (ou qualquer valor intermediário incluindo, entre outros, 1-100 pares de base, 1-50 pares de base, 1-30 pares de base, 1-20 pares de base, 1-10 pares de base ou 1-5 pares de base).

[0028] As células modificadas da invenção podem ser uma célula eucariótica, incluindo uma célula de mamífero não humano e uma célula humana como célula linfoide (por exemplo, uma célula T (incluindo uma célula T efetora (Teff) e uma célula T reguladora (Treg)), uma célula B ou uma célula NK), uma célula-tronco/progenitora (por exemplo, uma célula-tronco pluripotente induzida (iPSC), uma célula-tronco embrionária (por exemplo, ES humana), uma célula-tronco mesenquimal (MSC) ou uma célula-tronco hematopoiética (HSC). As células-tronco podem ser totipotentes ou pluripotentes (por exemplo, parcialmente diferenciadas como uma HSC que é uma célula-tronco mieloide ou linfoide pluripotente). Em

outras modalidades, a invenção provê métodos de produção de células com um genótipo nulo para expressão de TCR e/ou de HLA. Qualquer uma das células-tronco modificadas aqui descritas (modificadas no *loci* de TCRA e/ou B2M) pode então ser diferenciada para gerar uma célula diferenciada (*in vivo* ou *in vitro* (cultura)) descendente de uma célula-tronco conforme aqui descrita com as modificações aqui descritas, incluindo expressão modificada do gene TCRA e/ou B2M.

[0029] Em outro aspecto, as composições (células modificadas) e métodos aqui descritos podem ser utilizados, por exemplo, no tratamento ou na prevenção ou melhora de um transtorno. Os métodos tipicamente compreendem (a) clivagem ou regulação negativa de um gene TCR e/ou B2M endógeno em uma célula isolada (por exemplo, célula T ou outros linfócitos) utilizando uma nuclease (por exemplo, ZFN ou TALEN) ou sistema de nucleases como CRISPR/Cas com um RNA projetado de crRNA/tracr, ou utilizando um fator de transcrição projetado (por exemplo, ZFP-TF, TALE-TF, Cfp1-TF ou Cas9-TF), de tal modo que o gene TCR e/ou B2M seja inativado ou modulado negativamente; e (b) introdução da célula no indivíduo, pelo qual tratando ou prevenindo o transtorno. Em algumas modalidades, o gene codificador de TCR β (TCRB) é inativado ou modulado negativamente. Em algumas modalidades, o gene codificador de B2M é inativado ou modulado negativamente. Em algumas modalidades, a inativação é realizada por clivagem direcionada da região constante desse gene (Região constante de TCR β ou TRBC). Em modalidades preferidas, o gene codificador de TCR α (TCRA) e/ou B2M é inativado ou modulado negativamente. Em modalidades mais preferidas, o transtorno é um câncer, uma doença infecciosa ou uma doença autoimune. Em algumas modalidades, as modificações são efetuadas para induzir tolerância imune. Em modalidades mais preferidas, a inativação é

realizada por clivagem direcionada da região constante desse gene (Região constante de TCR α, ou abreviado como TRAC). Em algumas modalidades, um gene B2M é clivado. Em modalidades adicionais, os genes adicionais (além de TCR e/ou B2M) são modulados (*knocked-out*), por exemplo, *knockouts* duplos de TCR/B2M, genes TCR adicionais, PD1 e/ou CTLA4 e/ou um ou mais transgenes terapêuticos estão presentes na célula (epissomais, integrados aleatoriamente ou integrados via integração direcionada como integração mediada por nuclease). As células modificadas podem incluir uma ou mais ZFNs (por exemplo, pares de ZFN) como aqui descritas, incluindo, entre outras, uma nuclease de dedo de zinco (ZFN) compreendendo uma primeira e uma segunda ZFN, cada ZFN compreendendo um domínio de clivagem (por exemplo, qualquer domínio de clivagem FokI do tipo selvagem ou projetado) e um domínio de ligação ao DNA de ZFP. Em certas modalidades, as modificações são efetuadas utilizando uma ZFN compreendendo uma ZFP (regiões de reconhecimento de hélice e esqueleto) dos "desenhos" aqui descritos (por exemplo, Tabela 6 ou Tabela 8 incluindo as ZFPs das ZFNs designadas 68846, 53853, 72732; 72748; 68957; 55266, 68798, 68879, 68815, 68799 ou 72678), um domínio FokI (qualquer domínio FokI do tipo selvagem ou projetado) e opcionalmente um ligante (qualquer ligante entre o N- ou C-terminal do domínio FokI e o N- ou C-terminal dos desenhos de ZFP aqui descritos). Em algumas modalidades, a ZFN compreende um par de ZFNs, no qual uma ZFN compreende a ZFP de 68846 (SEQ ID Nº:177) operacionalmente ligada a um domínio FokI e a outra ZFN do par compreende a ZFP de 53853 (SEQ ID Nº:178) operacionalmente ligada a um domínio FokI. Em algumas modalidades, a ZFN compreende um par de ZFNs, no qual uma ZFN compreende a ZFP de 72732 (SEQ ID Nº:175) operacionalmente ligada a um domínio FokI e a outra ZFN do par compreende a ZFP de 72678 (SEQ ID Nº:176)

operacionalmente ligada a um domínio FokI. Em certas modalidades, a ZFN comprehende uma ZFN (por exemplo, um par de primeira e segunda ZFNs parceiras (também designada esquerda e direita)) aqui descritas como segue: uma ZFN designada 68796 e uma ZFN designada 68813; uma ZFN designada 68796 e uma ZFN designada 68861; uma ZFN designada 68812 e uma ZFN designada 68813; uma ZFN designada 68876 e uma ZFN designada 68877; uma ZFN designada 68815 e uma ZFN designada 55266; uma ZFN designada 68879 e uma ZFN designada 55266; uma ZFN designada 68798 e uma ZFN designada 68815; ou uma ZFN designada 68846 e uma ZFN designada 53853; uma ZFN designada 57531 e uma ZFN designada 72732; uma ZFN designada 57531 e uma ZFN designada 72748; uma ZFN designada 68957 e uma ZFN designada 57071; uma ZFN designada 68957 e uma ZFN designada 72732; uma ZFN designada 68957 e uma ZFN designada 72748; uma ZFN designada 72678 e uma ZFN designada 57071; uma ZFN designada 72678 e uma ZFN designada 72732; e a compreendendo uma ZPN de ZFP designada 72678 e uma ZFN designada 72748. Assim, uma ZFN (por exemplo, cada ZFN parceira de uma ZFN pareada) comprehende as regiões de reconhecimento de hélice e podem compreender modificações adicionais na ZFP (por exemplo, às regiões do esqueleto) descritas abaixo (por exemplo, desenhos mostrados nas Tabelas 1, 2, 5, 6 e 8) e comprehende ainda qualquer domínio de clivagem FoKI do tipo selvagem ou projetado (incluindo qualquer combinação dos mutantes de FokI por substituição, adição e/ou deleção). Por exemplo, uma ZFN parceira pode compreender um domínio dedo de zinco de ligação ao DNA específico fundido a qualquer domínio de clivagem FokI, incluindo o domínio de clivagem (SEQ ID N°:139) da proteína do tipo selvagem ou proveniente de uma sequência mutada (como mostrada nos Exemplos, SEQ ID N°:140-174). Uma ZFN parceira B2M-

específica pode compreender um domínio dedo de zinco de ligação ao DNA B2M-específico (por exemplo, 72732) fundido com um domínio de clivagem FokI selecionado dentre as SEQ ID N^{os}:139-174. Além disso, a ZFN parceira B2M-específica pode compreender um domínio dedo de zinco de ligação ao DNA B2M-específico (por exemplo, 72678) fundido a um domínio de clivagem FokI selecionado dentre as SEQ ID N^{os}:139-174. Do mesmo modo, uma ZFN parceira TRAC-específica pode compreender um domínio dedo de zinco de ligação ao DNA TRAC-específico (por exemplo, 68846) fundido a um domínio de clivagem FokI selecionado dentre as SEQ ID N^{os}:139-174, e o domínio dedo de zinco de ligação ao DNA TRAC-específico 53853 pode ser fundido a um domínio de clivagem FokI selecionado dentre qualquer um domínio de clivagem FokI do tipo selvagem ou projetado mostrado, por exemplo, nos Exemplos anexados (SEQ ID N^{os}:139-174). Em algumas modalidades, o domínio FokI é fundido à extremidade N-terminal do domínio de ligação ao DNA da ZFP, enquanto em outras, é fundido à extremidade C-terminal do domínio de ligação ao DNA da ZFP. Além disso, qualquer ligante pode ser utilizado para ligar o domínio de ligação ao DNA ao domínio de clivagem FokI.

[0030] São igualmente providas células descendentes de células modificadas como aqui descrito (por exemplo, células compreendendo as ZFNs aqui descritas), incluindo, entre outras, parciais ou totalmente diferenciadas a partir de células-tronco modificadas conforme aqui descritas. Essas células tipicamente não incluem as ZFNs, mas incluem as modificações genéticas feitas pelas mesmas.

[0031] O(s) fator(es) de transcrição e/ou nuclease(s) podem ser introduzidos em uma célula ou no meio de cultura em volta, como mRNA, em forma de proteína e/ou como uma sequência de DNA codificadora da(s) nuclease(s). Em certas modalidades, a célula isolada introduzida no indivíduo compreende ainda modificação

genômica adicional, por exemplo, uma sequência exógena integrada (no gene TCR e/ou B2M clivado ou em um gene diferente, por exemplo, um gene ou lócus *safe harbor*) e/ou inativação (por exemplo, mediada por nuclease) de genes adicionais, por exemplo, um ou mais dentre os genes HLA ou genes CTLA-4, CISH, PD1 ou tet2. A sequência (por exemplo, um CAR ou TCR exógeno) ou proteína exógena pode ser introduzida através de um vetor (por exemplo, Ad, AAV, LV), ou utilizando uma técnica como eletroporação ou transfecção transitória. Em algumas modalidades, as proteínas são introduzidas na célula induzindo estresse mecânico como por *Cell Squeeze* (ver Kollmannsperger, et al. (2016) *Nat Comm* 7, 10372 doi:10.1038/ncomms10372). Em alguns aspectos, a composição pode compreender fragmentos celulares isolados e/ou células (parciais ou totalmente) diferenciadas.

[0032] Em alguns aspectos, as células modificadas podem ser utilizadas para terapia celular, por exemplo, para transferência de células adotivas. Em outras modalidades, as células para uso no transplante de células T contêm outra modificação gênica de interesse. Em um aspecto, as células T contêm um receptor quimérico de antígeno (CAR) inserido, específico para um marcador encontrado em células cancerosas. Em um aspecto adicional, o CAR inserido é específico para o marcador CD19, característico de células B, incluindo neoplasias malignas de células B. Tais células seriam úteis em uma composição terapêutica para tratar pacientes sem ter que igualar HLA e, assim, poderiam ser utilizadas como um agente terapêutico "off-the-shelf" (pronto para uso) por qualquer paciente com necessidade do mesmo. Em outros casos, células-tronco ou precursoras, por exemplo, células-tronco hematopoiéticas ou células precursoras (HSC/PC) ou células-tronco pluripotentes induzidas (iPSC) contendo as modificações aqui descritas são expandidas antes

da introdução. Em outros aspectos, as HSC/PCs geneticamente modificadas são fornecidas ao indivíduo em um transplante de medula óssea, em que as HSC/PC enxertadas, diferenciam-se e maturam *in vivo*. Em algumas modalidades, as HSC/PC são isoladas do indivíduo após mobilização induzida por G-CSF. Mobilização induzida pelo plerixafor e combinações de mobilização induzida por G-CSF e pelo plerixafor e, em outros, as células são isoladas de medula óssea humana ou cordões umbilicais humanos. Em outras modalidades, iPSC são derivadas de células do paciente ou de doador sadio. Em alguns aspectos, o indivíduo é tratado para um procedimento mieloablativo leve antes da introdução do enxerto compreendendo as HSC/PC modificadas ou células modificadas derivadas de iPSC enquanto, em outros aspectos, o indivíduo é tratado com um regime vigoroso de condicionamento mieloablativo. Em algumas modalidades, os métodos e composições da invenção são utilizados para tratar ou prevenir um câncer.

[0033] Em outro aspecto, as células T com TCR e/ou B2M-modulados (modificadas) contêm uma sequência doadora inserida de receptor de célula T acoplado a um anticorpo (ACTR). Em algumas modalidades, a sequência doadora de ACTR é inserida em um gene TCR para prejudicar a expressão daquele gene TCR após a clivagem induzida por nuclease. Em outras modalidades, a sequência doadora é inserida em um lócus "safe harbor", como os genes AAVS1, HPRT, albumina e CCR5. Em algumas modalidades, a sequência de ACTR é inserida através de integração direcionada, em que a sequência doadora de ACTR comprehende braços de homologia flanqueando que possuem homologia com a sequência que flanqueia o sítio de clivagem da nuclease projetada. Em algumas modalidades, a sequência doadora de ACTR comprehende ainda uma sequência promotora e/ou outras reguladoras da transcrição. Em outras

modalidades, a sequência doadora de ACTR é desprovida de um promotor. Em algumas modalidades, a sequência doadora de ACTR é inserida em um gene que codifica TCR β (TCRB). Em algumas modalidades, a inserção é realizada através de clivagem direcionada da região constante desse gene (região constante de TCR β , ou TRBC). Em modalidades preferidas, a sequência doadora de ACTR é inserida em um gene que codifica TCR α (TCRA). Em modalidades mais preferidas, a inserção é realizada através de clivagem direcionada da região constante desse gene (Região constante de TCR α , TRAC abreviado). Em algumas modalidades, a sequência doadora é inserida em uma sequência de exons em TCRA enquanto, em outras, a sequência doadora é inserida em uma sequência intrônica em TCRA. Em ainda outras modalidades, a sequência doadora de ACTR é inserida em um gene B2M. Em algumas modalidades, as células com B2M e/ou TCR modulados compreendem ainda um CAR. Em ainda outras modalidades, as células com B2M e/ou TCR modulados são moduladas adicionalmente em um gene HLA ou um gene de inibidor de *checkpoint*.

[0034] Além disso, são providas também composições farmacêuticas compreendendo as células modificadas aqui descritas (por exemplo, células T ou células-tronco com o gene TCR inativado), ou composições farmacêuticas compreendendo uma ou mais das moléculas de ligação aos genes TCR e/ou B2M (por exemplo, fator de transcrição e/ou nucleases projetados) conforme aqui descritas. Em certas modalidades, as composições farmacêuticas compreendem ainda um ou mais excipientes farmaceuticamente aceitáveis. As células modificadas, as moléculas de ligação aos genes TCR e/ou B2M (ou polinucleotídeos que codificam essas moléculas) e/ou as composições farmacêuticas compreendendo essas células ou moléculas são introduzidas no indivíduo através de métodos

conhecidos na técnica, por exemplo, através de infusão intravenosa, infusão em um vaso específico como a artéria hepática, ou através de injeção direta no tecido (por exemplo, músculo). Em algumas modalidades, o indivíduo é um humano adulto com uma doença ou condição que pode ser tratada ou melhorada com a composição. Em outras modalidades, o indivíduo é pediátrico, em que a composição é administrada para prevenir, tratar ou melhorar a doença ou condição (por exemplo, câncer, doença do enxerto contra hospedeiro, etc.).

[0035] Em alguns aspectos, a composição (células com TCR e/ou B2M modulados compreendendo um ACTR) compreende ainda um anticorpo exógeno. Ver, também, a Publicação de Patente U.S. Nº 2017/0196992. Em alguns aspectos, o anticorpo é útil em armar uma célula T compreendendo ACTR para prevenir ou tratar uma condição. Em algumas modalidades, o anticorpo reconhece um antígeno associado com uma célula tumoral ou com processos associados ao câncer como EpCAM, CEA, gpA33, mucinas, TAG-72, CAIX, PSMA, anticorpos que se ligam ao folato, CD19, EGFR, ERBB2, ERBB3, MET, IGF1R, EPHA3, TRAILR1, TRAILR2, RANKL, FAP, VEGF, VEGFR, integrinas α V β 3 e α 5 β 1, CD20, CD30, CD33, CD52, CTLA4 e enascina (Scott, *et al.* (2012) *Nat Rev Cancer* 12:278). Em outras modalidades, o anticorpo reconhece um antígeno associado com uma doença infecciosa como HIV, HCV e os semelhantes.

[0036] Em outro aspecto, são providos domínios de ligação ao DNA do gene TCR (por exemplo, ZFPs, TALEs e sgRNAs) que se ligam a um sítio alvo em um gene TCR. Em certas modalidades, o domínio de ligação ao DNA compreende uma ZFP com as regiões de reconhecimento de hélice na ordem mostrada em uma linha única da Tabela 1; uma proteína de ligação ao DNA do domínio efetor de TAL com as RVDs (repetições variáveis di-resíduos) de ligação a um sítio alvo conforme mostradas na primeira coluna da Tabela 1 ou na

terceira coluna da Tabela 2; e/ou um sgRNA como mostrado em uma linha única da Tabela 2. Essas proteínas de ligação ao DNA podem ser associadas com domínios reguladores da transcrição para formar fatores de transcrição projetados que modulam a expressão de TCR. Alternativamente, essas proteínas de ligação ao DNA podem ser associadas com um ou mais domínios de nucleases para formar nucleases de dedo de zinco (ZFNs) projetadas, TALENs e/ou sistemas CRISPR/Cas que podem se ligar e clivar um gene TCR. Em certas modalidades, as ZFNs, TALENs ou RNAs guias únicos (sgRNA) de um sistema CRISPR/Cas ligam-se a sítios alvo em um gene TCR humano. O domínio de ligação ao DNA do fator de transcrição ou nuclease (por exemplo, ZFP, TALE, sgRNA) pode ligar-se a um sítio alvo em um gene TCRA compreendendo 9, 10, 11, 12 ou mais (por exemplo, 13, 14, 15, 16, 17, 18, 19, 20 ou mais) nucleotídeos de qualquer um dos sítios alvo aqui mostrados (por exemplo, sítios alvo da Tabela 1 ou 2 como mostrados na SEQ ID N^{os}:8-21 e/ou 92-103). As proteínas dedo de zinco podem incluir 1, 2, 3, 4, 5, 6 ou mais dedos de zinco, cada dedo de zinco tendo uma hélice de reconhecimento para contato específico com um subsítio alvo no gene alvo. Em certas modalidades, as proteínas dedo de zinco compreendem 4 ou 5 ou 6 dedos (designados F1, F2, F3, F4, F5 e F6 e na ordem de F1 para F4 ou F5 ou F6 do N-terminal ao C-terminal), por exemplo, como mostradas na Tabela 1. As ZFPs aqui descritos podem também incluir uma ou mais mutações para contato do fosfato com resíduos da proteína dedo de zinco, por exemplo, o mutante nR-5Qabc descrito na Publicação de Patente U.S. Nº 2018/0087072. Em outras modalidades, os RNAs guias únicos ou domínios de ligação ao DNA do efetor TAL podem ligar-se a um sítio alvo conforme aqui descrito (por exemplo, sítios alvo da Tabela 1 ou Tabela 2 ou Tabela 6 como mostrados em qualquer uma dentre SEQ ID N^{os}:8-21 e/ou 92-103) ou 12 ou mais pares de

base dentro de qualquer um desses sítios alvo ou entre sítios alvo pareados. Sítio alvos exemplares para sgRNA como mostrados na Tabela 2 (SEQ ID N^{os}:92-103). sgRNAs que se ligam a 12 ou mais nucleotídeos dos sítios alvo mostrados na Tabela 1 ou Tabela 2 são também providos. Desenhos de TALENs podem ser desenhados para atingir sítios conforme aqui descritos (sítios alvo da Tabela 1 ou Tabela 2 ou Tabela 6) utilizando RVDs canônicas ou não como descrito nas Patentes U.S. N^{os} 8 586 526 e 9 458 205. As nucleases aqui descritas (compreendendo um domínio de ligação ao DNA de ZFP, de TALE ou de sgRNA) são capazes de efetuar modificações genéticas dentro de um gene TCRA compreendendo qualquer uma dentre SEQ ID N^º:8-21 e/ou 92-103, incluindo modificações (inserções e/ou deleções) dentro de qualquer uma dessas sequências (SEQ ID N^º:8-21 e/ou 92-103) e/ou modificações às sequências do gene TCRA que flanqueiam as sequências do sítio alvo mostradas na SEQ ID N^º:8-21 e/ou 92-103, por exemplo, modificações dentro de sequências exônicas de um gene TCR dentro de uma ou mais das sequências seguintes: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e AATCCTC.

[0037] Em outro aspecto, são providos domínios de ligação ao DNA do gene B2M (por exemplo, ZFPs, TALEs e sgRNAs) que se ligam a um sítio alvo em um gene B2M. Em certas modalidades, o domínio de ligação ao DNA compreende uma ZFP com as regiões de reconhecimento de hélice na ordem mostrada em uma linha única da Tabela 5 ou Tabela 8 (colunas identificadas "desenhos", incluindo as ZFPs das ZFNs designadas 72732; 72748; 68957; ou 72678); uma proteína de ligação ao DNA do efeito TAL com as RVDs que se ligam um sítio alvo mostradas na primeira coluna da Tabela 5 ou Tabela 8; e/ou um sgRNA que se liga a um sítio alvo em B2M conforme aqui descrito (Tabela 5 ou Tabela 8). Essas proteínas de ligação ao DNA podem ser associadas com domínios reguladores da transcrição para

formar fatores de transcrição projetados que modulam a expressão de B2M. Alternativamente, essas proteínas de ligação ao DNA podem ser associadas com um ou mais domínios de nuclease (clivagem) para formar nucleases de dedo de zinco (ZFNs) projetadas, TALENs e/ou sistemas CRISPR/Cas que se ligam e clivam um gene B2M. Em certas modalidades, as ZFNs, TALENs ou RNAs guias únicos (sgRNA) de um sistema CRISPR/Cas se ligam a sítios alvo em um gene B2M humano. O domínio de ligação ao DNA do fator de transcrição ou nuclease (por exemplo, ZFP, TALE, sgRNA) pode ligar-se a um sítio alvo em um gene B2M compreendendo 9, 10, 11, 12 ou mais (por exemplo, 13, 14, 15, 16, 17, 18, 19, 20 ou mais) nucleotídeos de qualquer um dos sítios alvo aqui mostrados (por exemplo, Tabela 5 ou Tabela 8 como mostrados na SEQ ID N^{os}:117, 123, 126 ou 127). As proteínas dedo de zinco podem incluir 1, 2, 3, 4, 5, 6 ou mais dedos de zinco, cada dedo de zinco tendo uma hélice de reconhecimento para contato específico com um subsítio alvo no gene alvo. Em certas modalidades, as proteínas dedo de zinco compreende 4 ou 5 ou 6 dedos (designados F1, F2, F3, F4, F5 e F6 e na ordem de F1 para F4 ou F5 ou F6 do N-terminal ao C-terminal), por exemplo, como mostrado na Tabela 5 ou Tabela 8. As ZFPs aqui descritas podem também incluir uma ou mais mutações para contato do fosfato com resíduos da proteína dedo de zinco, por exemplo, o mutante nR-5Qabc descrito na Publicação de Patente U.S. Nº 2018/0087072, incluindo os desenhos de ZFP (mutantes de regiões de reconhecimento de hélice e esqueleto) da Tabela 8. Em outras modalidades, os RNAs guias únicos ou domínios de ligação ao DNA do efetor TAL podem ligar-se a um sítio alvo conforme aqui descrito (por exemplo, sítios alvo das Tabelas 5 ou 8) ou 12 ou mais pares de base dentro de qualquer um desses sítios alvo ou entre sítios alvo pareados. Domínios TALE podem ser desenhados para atingir sítios conforme aqui descritos (sítios alvo das Tabelas 5 ou

8) utilizando RVDs canônicas ou não conforme descritas nas Patentes U.S. N^{os} 8 586 526 e 9 458 205. As nucleases aqui descritas (compreendendo um domínio de ligação ao DNA de ZFP, de TALE ou de sgRNA) são capazes de efetuar modificações genéticas dentro de um gene B2M compreendendo qualquer um dos sítios alvos de B2M aqui revelados, incluindo modificações (inserções e/ou deleções) dentro de qualquer uma dessas sequências e/ou modificações a sequências do gene B2M que flanqueiam as sequências do sítio alvo mostradas nas Tabelas 5 e 8 (SEQ ID Nº:117, 123, 126 ou 127).

[0038] Qualquer uma das nucleases aqui descritas pode compreender um domínio de ligação ao DNA (por exemplo, desenhos de ZFP da Tabela 6 ou 8, TALE ou sgRNA) conforme aqui descrito e um domínio de clivagem e/ou meio-domínio de clivagem (por exemplo, meio-domínio de clivagem FokI do tipo selvagem ou projetado). Assim, em qualquer uma das nucleases (por exemplo, ZFNs, TALENs, sistemas CRISPR/Cas) aqui descritas, o domínio nuclease pode compreender um domínio nuclease do tipo selvagem ou meio-domínio nuclease (por exemplo, um meio-domínio de clivagem FokI). Em outras modalidades, as nucleases (por exemplo, ZFNs, TALENs, nucleases CRISPR/Cas) compreendem domínios ou meios-domínios de nuclease, por exemplo, meios-domínios de clivagem FokI projetados que formam heterodímeros obrigados. Ver, por exemplo, as Patente U.S. N^{os} 7 914 796 e 8 034 598. Em certas modalidades, um ou mais domínios endonuclease de FokI das nucleases aqui descritas podem também compreender mutantes para contato com fosfato (por exemplo, R416S e/ou K525S) como descritos na Publicação de Patente U.S. N° 2018/0087072. Assim, o domínio FokI das nucleases aqui descritas (por exemplo, ZFNs compreendendo: (i) desenhos de ZFP como mostrados na Tabela 8, incluindo ZFPs das ZFNs designadas 72732; 72748; 68957; ou 72678 e (ii) um domínio FokI)

podem incluir qualquer combinação de mutações ao domínio FokI (posições numeradas em relação ao domínio FokI completo), incluindo a sequência do domínio catalítico FokI do tipo selvagem e também, entre outros, os domínios FokI indicados na Tabela 8, FokI-Sharkey (S418P+K441E); FokI ELD (Q->E na posição 486, I->L em 499, N->D na posição 496); FokI ELD, Sharkey (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E); FokI ELD, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, R416E); FokI ELD, Sharkey, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, R416E); FokI ELD, R416Y (Q->E na posição 486, I->L na posição 499, N->D na posição 496, R416Y); FokI ELD, Sharkey, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, R416E); FokI ELD, S418E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418E); FokI ELD, Sharkey parcial, S418E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, K441E, S418E); FokI ELD, K525S (Q->E na posição 486, I->L na posição 499, N->D na posição 496, K525S); FokI ELD, Sharkey K525S (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, K525S); FokI ELD, I479T (Q->E na posição 486, I->L na posição 499, N->D na posição 496, I479T); FokI ELD, Sharkey, I479T (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, I479T); FokI ELD, P478D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, P478D); FokI ELD, Sharkey, P478D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, P478D); FokI ELD, Q481D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, Q481D); FokI ELD, Sharkey, Q481D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, Q481D); FokI KKR (E->K na posição 490, I->K na posição 538, H->R na posição 537); FokI KKR

Sharkey, (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E); FokI KKR, Q481E (E->K na posição 490, I->K na posição 538, H->R na posição 537, Q481E); FokI KKR, Sharkey Q481E (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, Q481E); FokI KKR, R416E (E->K na posição 490, I->K na posição 538, H->R na posição 537, R416E); FokI KKR, Sharkey, R416E (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, R416E); FokI KKR, K525S (E->K na posição 490, I->K na posição 538, H->R na posição 537, K525S); FokI KKR, Sharkey, K525S (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, K525S); FokI KKR, R416Y (E->K na posição 490, I->K na posição 538, H->R na posição 537, R416Y); FokI KKR, Sharkey, R416Y (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, R416Y); FokI, KKR I479T (E->K na posição 490, I->K na posição 538, H->R na posição 537, I479T); FokI, KKR Sharkey I479T (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, I479T); FokI, KKR P478D(E->K na posição 490, I->K na posição 538, H->R na posição 537, P478D), FokI KKR Sharkey P478D(E->K na posição 490, I->K na posição 538, H->R na posição 537, P478D); FokI DAD (R->D na posição 487, N->D na posição 496, I->A na posição 499); FokI DAD Sharkey (R->D na posição 487, N->D na posição 496, I->A na posição 499, S418P+K441E); FokI RVR (D->R na posição 483, H->R na posição 537, I->V na posição 538); FokI RVR Sharkey (D->R na posição 483, H->R na posição 537, I->V na posição 538, S418P+K441E). As ZFNs aqui descritas podem também incluir qualquer sequência de ligação, incluindo, entre outras, as sequências reveladas na Patente U.S. Nº 7 888 121; 7 914 796; 8 034 598; 8 623 618; 9 567 609; e Publicação U.S. Nº 2017/0218349, as quais podem ser utilizadas entre o N- ou C-terminal do domínio de ligação ao DNA

(por exemplo, ZFP) e N- ou C-terminal do domínio de clivagem FokI.

[0039] Em outro aspecto, a invenção provê um polinucleotídeo que codifica qualquer uma das proteínas, moléculas de fusão e/ou componentes das mesmas (por exemplo, sgRNA ou outro domínio de ligação ao DNA) aqui descritos. O polinucleotídeo pode ser parte de um vetor viral, um vetor não viral (por exemplo, plasmídeo) ou estar em forma de mRNA. Qualquer um dos polinucleotídeos aqui descritos pode também compreender sequências (sequências doadoras, braços de homologia ou retalhos) para inserção direcionada no gene TCR α e/ou o TCR β. Em ainda outro aspecto, é provido um vetor para a entrega de genes que compreende qualquer um dos polinucleotídeos aqui descrito. Em certas modalidades, o vetor é um vetor adenoviral (por exemplo, um vetor Ad5/F35) ou um vetor lentiviral (LV) incluindo o componente competente para integração ou vetores lentivirais com integração defeituosa ou um vetor adenoassociado (AAV). Assim, são também providos vetores virais compreendendo uma sequência que codifica uma nuclease (por exemplo, ZFN ou TALEN) e/ou um sistema de nucleases (CRISPR/Cas ou Ttago) e/ou uma sequência doadora para integração direcionada em um gene alvo. Em algumas modalidades, a sequência doadora e as sequências que codificam nuclease estão em vetores diferentes. Em outras modalidades, as nucleases são fornecidas como polipeptídeos. Em modalidades preferidas, os polinucleotídeos são mRNAs. Em alguns aspectos, o mRNA pode ser quimicamente modificado (Ver, por exemplo, Kormann, *et al.* (2011) *Nature Biotechnology* 29(2):154-157). Em outros aspectos, o mRNA pode compreender um cap ARCA (ver Patentes U.S. N°s 7 074 596 e 8 153 773). Em alguns aspectos, o mRNA pode compreender um cap introduzido por modificação enzimática. O cap introduzido enzimaticamente pode compreender Cap0, Cap1 ou Cap2 (ver, por exemplo, Smietanski, *et al.* (2014)

Nature Communications 5:3004). Em aspectos adicionais, o mRNA pode ser delimitado com cap por modificação química. Em modalidades adicionais, o mRNA pode compreender uma mistura de nucleotídeos não modificados e modificados (ver Publicação de Patente U.S. Nº 2012/0195936). Em ainda outras modalidades, o mRNA pode compreender um elemento WPRE (ver Publicação de Patente U.S. Nº 2016/0326548). Em algumas modalidades, o mRNA é dupla-fita (ver, por exemplo, Kariko, et al. (2011) *Nucl Acid Res* 39:e142).

[0040] Em ainda outro aspecto, a invenção provê uma célula isolada compreendendo qualquer uma das proteínas, polinucleotídeos e/ou vetores aqui descritos. Em certas modalidades, a célula é selecionada a partir do grupo que consiste em célula-tronco/progenitora ou célula T (por exemplo, célula T efetora ou reguladora). Em ainda mais um aspecto, a invenção provê uma célula ou linhagem celular que descende de uma célula ou linhagem compreendendo qualquer uma das nucleases, fatores de transcrição, polinucleotídeos e/ou vetores aqui descritos, a saber, uma célula ou linhagem celular descendente (por exemplo, em cultura) de uma célula na qual TCR e/ou B2M foram inativados por uma ou mais ZFNs e/ou na qual um polinucleotídeo doador (por exemplo, ACTR e/ou CAR) foi integrado estavelmente no genoma da célula. Assim, descendentes de células conforme aqui descritas podem não compreender por si mesmos a molécula, os polinucleotídeos e/ou vetores aqui descritos, mas, nessas células, um gene TCR e/ou B2M está inativado e/ou um polinucleotídeo doador está integrado no genoma e/ou é expresso.

[0041] Em outro aspecto, são descritos métodos para inativação de um gene TCR e/ou B2M em uma célula, introduzindo uma ou mais proteínas, polinucleotídeos e/ou vetores na célula conforme aqui descrito. Em certas modalidades, um ou mais polinucleotídeos

codificadores de uma ZFN (por exemplo, par de ZFNs), como mostrada na Tabela 6, são utilizados para modificar o gene TCR na célula, e células descendentes dessas células (incluindo células diferenciadas) compreendem a(s) modificaçõe(s). Em outras modalidades, um ou mais polinucleotídeos codificadores de uma ZFN (por exemplo, par de ZFNs), como mostrada na Tabela 6, são utilizados para modificar o gene B2M na célula, e células descendentes dessas células (incluindo células diferenciadas) compreendem a modificação. Em qualquer um dos métodos aqui descritos, as nucleases podem induzir mutagênese direcionada, deleções de sequências do DNA celular e/ou facilitar a recombinação direcionada em lócus predeterminado em cromossomos. Assim, em certas modalidades, as nucleases deletam e/ou inserem um ou mais nucleotídeos de ou no gene alvo. Em algumas modalidades, um gene TCR e/ou B2M é inativado pela clivagem por nucleases após junção de extremidades não homólogas. Em outras modalidades, uma sequência genômica no gene alvo (por exemplo, TCR ou B2M) é substituída, por exemplo, usando uma nuclease (ou vetor codificador da dita nuclease) aqui descrita e uma sequência "doadora" que é inserida no gene após clivagem direcionada com a nuclease. A sequência doadora pode estar presente no vetor da nuclease, presente em um vetor separado (por exemplo, plasmídeo, DNA linear fita simples ou fita dupla, vetor AAV, Ad ou LV) ou, alternativamente, pode ser introduzida na célula empregando um mecanismo diferente para entrega de ácidos nucleicos. Em algumas modalidades, os métodos compreendem ainda inativar um ou mais genes adicionais (por exemplo, B2M) e/ou integrar um ou mais transgenes no genoma da célula, incluindo, entre outros, a integração de um ou mais transgenes no gene TCR e/ou B2M inativados e/ou em um ou mais genes *safe harbor*. Em certas modalidades, os métodos aqui descritos

resultam em uma população de células na qual pelo menos 80-100% (ou qualquer valor intermediário), incluindo pelo menos 90-100% (ou qualquer valor intermediário) das células incluem o(s) knockout(s) e/ou o(s) transgene(s) integrado(s).

[0042] Além disso, qualquer um dos métodos aqui descritos pode ser praticado *in vitro*, *in vivo* e/ou *ex vivo*. Em certas modalidades, os métodos são praticados *ex vivo*, por exemplo, para modificar células T (efetoras ou reguladoras), a fim de torná-las úteis como agentes terapêuticos em um contexto alogênico para tratar um indivíduo (por exemplo, um indivíduo com câncer ou doença autoimune). Exemplos não limitantes de cânceres que podem ser tratados e/ou prevenidos incluem carcinomas pulmonares, cânceres pancreáticos, cânceres hepáticos, cânceres ósseos, cânceres de mama, cânceres colorretais, leucemias, cânceres ovarianos, linfomas, cânceres cerebrais e os semelhantes. Exemplos não limitantes de doença autoimune incluem rejeição ao transplante, doença/transtorno do intestino irritável, esclerose múltipla, lúpus, esclerodermia, artrite reumatoide e semelhantes. As células podem também ser utilizadas para induzir tolerância imune.

[0043] Em outro aspecto, é descrito um método para integração de um ou mais transgenes em um genoma de uma célula isolada, o método compreendendo: introduzir, na célula, (a) um ou mais vetores doadores (por exemplo, plasmídeo, DNA linear fita simples ou fita-dupla, AAVs, plasmídeos, Ads, mRNAs, etc.) compreendendo um ou mais transgenes e (b) pelo menos uma nuclease não natural em forma de mRNA, em que pelo menos uma nuclease cliva o genoma da célula de tal modo que um ou mais dos transgenes sejam integrados no genoma da célula (por exemplo, em um receptor TCR), em que o vetor doador é introduzido no tampão de eletroporação compreendendo a célula isolada e o mRNA imediatamente antes ou imediatamente

depois da eletroporação da nuclease na célula. Em certas modalidades, o vetor doador é introduzido no tampão de eletroporação depois da eletroporação e antes de transferir as células para um meio de cultura. Ver, por exemplo, as Publicações de Patentes U.S. N°s 2015/0174169 e 2015/0110762. Os métodos podem ser utilizados para introduzir o(s) transgene(s) em qualquer localização genômica, incluindo, entre outros, um gene TCR, um gene B2M e/ou um gene *safe harbor* (por exemplo, AAVS1, Rosa, albumina, CCR5, CXCR4, etc.).

Breve Descrição dos Desenhos

[0044] As Figuras 1A e 1B são uma representação do gene TCRA mostrando a localização dos sítios visados pelas nucleases. A Figura 1A é uma ilustração do processamento do gene TCRA desde a forma na linhagem germinativa até aquela de uma célula T madura e indica o alvo geral das nucleases. A Figura 1B (SEQ ID N°s:116 (exon c1), 117 (exon c2) e 118 (exon c3)) mostra as regiões entre os sítios alvo na sequência da região constante. A sequência mostrada com letras maiúsculas em preto é a sequência da sequência indicada do exon, enquanto a sequência com letras minúsculas em cor cinza é a sequência do ítron adjacente.

[0045] As Figuras 2A e 2B são gráficos representando o percentual de cada sítio modificado em células T tratadas com ZFNs específicas para os sítios sites A, B e D (Figura 2A) e sites E, F e G (Figura 2B) de TCRA. Muitos dos pares forneceram taxas de modificação iguais ou superiores a 80%.

[0046] A Figura 3 representa o percentual de células T CD3 negativas após o tratamento com os pares de ZFNs TCRA-específicos conforme analisado por FACS.

[0047] A Figura 4 é um gráfico mostrando o alto grau de correlação em células T entre níveis de modificação na sequência de

TCRA, conforme medido por sequenciamento em alta escala (*high-throughput sequencing*), e perda de expressão de CD3 conforme medida por separação de células ativadas por fluorescência.

[0048] As Figuras 5A a 5D são gráficos representando o crescimento de células T após o tratamento com a ZFN TCRA-específica, agrupado de acordo com o sítio alvo no gene TCRA.

[0049] A Figura 6 mostra resultados do *knockout* duplo de TRAC (TCRA) e B2M e da integração direcionada de um doador no lócus de TRAC (TCRA) ou no de B2M.

[0050] Figura 7 mostra resultados por FACS do *knockout* duplo de TRAC (TCRA) e B2M e da integração direcionada de um doador no lócus de TRAC (TCRA) ou no de B2M. Os resultados por FACS são mostrados para as condições indicadas (da esquerda para a direita dos painéis superiores: controle simulado (*sham*); ZFNs para TRAC e B2M sem um doador; ZFNs para TRAC e B2M com doador direcionado para B2M; e ZFNs para TRAC e B2M com doador direcionado para TRAC). O quadrante esquerdo inferior da linha superior de gráficos de FACS mostra células com duplo *knockout* (TRAC/B2M) e o do meio direito à direita da linha inferior de gráficos de FACS mostra células com duplo *knockout* e integração direcionada. A porcentagem de células é também indicada por setas apontando em direção à seção apropriada do gráfico de FACS. Como indicado pelas setas, 85-90% ou mais das células apresentava duplo KO e eram também positivas para integração direcionada.

Descrição Detalhada

[0051] São descritas composições e métodos para geração de células nas quais a expressão de um gene TCR é modulada de tal modo que as células não compreendem mais um TCR nas suas superfícies celulares e/ou nas quais a expressão de um gene B2M é modulada de tal modo que as células não expressam mais B2M.

Células modificadas dessa maneira podem ser utilizadas como agente terapêutico, por exemplo, transplantes, pois a ausência de um complexo TCR impede ou reduz uma resposta imune à base de HLA. Além disso, outros genes de interesse (por exemplo, transgenes) podem ser inseridos em células nas quais os genes TCR e/ou B2M foram manipulados. Um ou mais genes adicionais (não TCR e/ou B2M) (por exemplo, outros genes TCR, B2M, PD1, CTLA4, HLA, genes *safe harbor*, etc.) podem ser modificados por *knockout* (desabilitação) e/ou inserção direcionada de sequências exógenas. As sequências exógenas podem incluir receptores quiméricos de抗ígenos para integração nas células modificadas, as quais podem ser utilizadas para tratar câncer e transtornos autoimunes.

Geral

[0052] A prática dos métodos, bem como a preparação e uso das composições aqui reveladas empregam, a menos que indicado o contrário, técnicas convencionais de biologia molecular, bioquímica, estrutura e análise de cromatina, química computacional, cultura celular, DNA recombinante e campos relacionados abrangidos pela competência no assunto. Essas técnicas estão totalmente explicadas na literatura. Ver, por exemplo, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, segunda edição, Cold Spring Harbor Laboratory Press, 1989, e terceira edição, 2001; Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1987 e atualizações periódicas; a série *Methods in Enzymology*, Academic Press, San Diego; Wolffe, *Chromatin Structure and Function*, terceira edição, Academic Press, San Diego, 1998; *Methods in Enzymology*, Vol. 304, "Chromatin" (P.M. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; e *Methods in Molecular Biology*, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definições

[0053] Os termos "ácido nucleico", "polinucleotídeo" e "oligonucleotídeo" são usados alternadamente e referem-se a um polímero de desorribonucleotídeos ou ribonucleotídeos, em conformação linear ou circular e em forma de fita simples ou de dupla. Para fins da presente invenção, esses termos não deverão ser interpretados como limitação no que diz respeito ao comprimento de um polímero. Os termos podem abranger análogos conhecidos de nucleotídeos naturais, bem como nucleotídeos que são modificados nas porções de base, açúcar e/ou fosfato (por exemplo, cadeias principais de fosforotioato). Em geral, um análogo de um nucleotídeo em particular possui a mesma especificidade de pareamento de bases; ou seja, um análogo de A fará par de base com T.

[0054] Os termos "polipeptídeo", "peptídeo" e "proteína" são usados alternadamente para se referir a um polímero de resíduos de aminoácidos. O termo também se aplica a polímeros de aminoácidos nos quais um ou mais aminoácidos são análogos químicos ou derivados modificados de aminoácidos naturais correspondentes.

[0055] "Ligaçāo" refere-se a uma interação não covalente sequência-específica entre macromoléculas (por exemplo, entre uma proteína e um ácido nucleico). Nem todos os componentes de uma interação por ligação precisam ser sequência-específicos (por exemplo, contatos com resíduos de fosfato em um esqueleto de DNA), desde que a interação como um todo seja sequência-específica. Tais interações são geralmente caracterizadas por uma constante de (K_d) igual ou inferior a 10^{-6} M^{-1} . "Afinidade" refere-se à força de ligação: sendo a afinidade de ligação aumentada correlacionada com K_d menor. "Ligaçāo não específica" refere-se a interações não covalentes que ocorrem entre qualquer molécula de interesse (por exemplo, uma nuclease projetada) e uma

macromolécula (por exemplo, DNA) que não são dependentes da sequência alvo.

[0056] Uma "molécula de ligação ao DNA" é uma molécula que consegue ligar-se ao DNA. Tal molécula de ligação ao DNA pode ser um polipeptídeo, um domínio de uma proteína, um domínio dentro de uma proteína maior ou um polinucleotídeo. Em algumas modalidades, o polinucleotídeo é DNA enquanto, em outras modalidades, o polinucleotídeo é RNA. Em algumas modalidades, a molécula de ligação ao DNA é um domínio proteico de uma nuclease (por exemplo, o domínio FokI) enquanto, em outras modalidades, a molécula de ligação ao DNA é um componente RNA guia de uma nuclease guiada por RNA (por exemplo, Cas9 ou Cfp1).

[0057] Uma "proteína de ligação" é uma proteína que é capaz de ligação não covalente com outra molécula. Uma proteína de ligação pode ligar-se a, por exemplo, uma molécula de DNA (uma proteína de ligação ao DNA), uma molécula de RNA (uma proteína de ligação ao RNA) e/ou uma molécula de proteína (uma proteína de ligação à proteína). No caso de proteína de ligação à proteína, esta pode ligar-se a si mesma (para formar homodímeros, homotrimers, etc.) e/ou pode ligar-se a uma ou mais moléculas de uma proteína ou proteínas diferentes. Uma proteína de ligação pode ter mais de um tipo de atividade de ligação. Por exemplo, proteínas dedo de zinco têm atividade de ligação ao DNA, ligação ao RNA e ligação à proteína.

[0058] Uma "proteína dedo de zinco de ligação ao DNA" (ou domínio de ligação) é uma proteína, ou um domínio dentro de uma proteína maior, que o DNA de maneira sequência-específica através de um ou mais dedos de zinco, os quais são regiões de sequência de aminoácidos dentro do domínio de ligação cuja estrutura é estabilizada através de coordenação de um íon de zinco. Assim, cada dedo de zinco de uma ZFP multi-dedo inclui uma região de reconhecimento de

hélice para ligação ao DNA dentro de um esqueleto. O termo proteína dedo de zinco de ligação ao DNA é frequentemente abreviado como proteína dedo de zinco ou ZFP. O termo "nuclease de dedo de zinco" inclui uma ZFN bem como um par de ZFNs (os membros do par são referidos como "esquerda e direita" ou "primeira e segunda" ou "par") que dimerizam para clivar o gene alvo.

[0059] Um "domínio TALE de ligação ao DNA" ou "TALE" é um polipeptídeo compreendendo um ou mais domínios/unidades TALE repetidos. Os domínios repetidos, cada qual compreendendo uma repetição variável di-resíduo (RVD), estão envolvidos na ligação do TALE a sua sequência cognata de DNA alvo. Uma "unidade repetida" (também designada como "repetição") possui tipicamente 33-35 aminoácidos de comprimento e exibe pelo menos alguma homologia de sequência com outras sequências TALE repetidas dentro de uma proteína TALE natural. As proteínas TALE podem ser desenhadas para se ligarem a um sítio alvo utilizando RVDs canônicas ou não canônicas dentro das unidades repetidas. Ver, por exemplo, as Patentes U.S. N°s 8 586 526 e 9 458 205. Os domínios dedo de zinco e TALE de ligação ao DNA podem ser "projetadas" para se ligarem a uma sequência nucleotídica predeterminada, por exemplo, por engenharia (alterando um ou mais aminoácidos) da região da hélice de reconhecimento de uma proteína dedo de zinco natural ou por engenharia dos aminoácidos envolvidos em ligação ao DNA (a repetição variável di-resíduo ou região RVD). Portanto, proteínas dedo de zinco ou proteínas TALE projetadas são proteínas que não são naturais. Exemplos não limitantes de métodos para engenharia de proteínas dedo de zinco e TALEs são desenho e seleção. Uma proteína desenhada é uma proteína que não ocorre na natureza e cujo desenho/composição resulta principalmente de critérios racionais. Os critérios racionais para o desenho incluem a aplicação de regras de

substituição e algoritmos computadorizados para o processar as informações em um banco de dados que armazena informações de desenhos existentes de ZFP ou TALE (RVDs canônicas e não canônicas) e dados de ligação. Ver, por exemplo, as Patentes U.S. N°s 9 458 205; 8 586 526; 6 140 081; 6 453 242; e 6 534 261; ver também as Publicações de Patente Internacional N°s WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536; e WO 03/016496. O termo "TALEN" inclui uma TALEN bem como um par de TALENs (os membros do par são referidos como "esquerda e direita" ou "primeira e segunda" ou "par") que dimerizam para clivar o gene alvo.

[0060] Uma proteína dedo de zinco, proteína TALE ou sistema CRISPR/Cas "selecionado" não é encontrado na natureza e cuja produção resulta primariamente de um processo empírico como expressão em fagos (*phage display*), mapeamento de interações (*interaction trap*) ou seleção de híbridos. Ver, por exemplo, as Patentes U.S. N°s 5 789 538; 5 925 523; 6 007 988; 6 013 453; 6 200 759 e Publicações de Patente Internacional N°s WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; e WO 02/099084.

[0061] "TtAgo" é uma proteína Argonaute procariótica a qual se acredita esteja envolvida no silenciamento de genes. TtAgo é derivada da bactéria *Thermus thermophilus*. Ver, por exemplo, Swarts, *et al.*, *ibid*, G. Sheng, *et al.* (2013) *Proc. Natl. Acad. Sci. U.S.A.* 111, 652). Um "sistema TtAgo" inclui todos os componentes necessários, entre os quais, por exemplo, DNAs guia para a clivagem de uma enzima TtAgo.

[0062] "Recombinação" refere-se a um processo para troca de informações genéticas entre dois polinucleotídeos. Para fins desta invenção, "recombinação homóloga (HR)" refere-se à forma especializada de tal troca que ocorre, por exemplo, durante o reparo de quebras na dupla fita em células através de mecanismos de reparo

direcionado por homologia. Esse processo requer homologia das sequências nucleotídicas, faz uso de uma molécula "doadora" como molde para o reparo de uma molécula "alvo" (ou seja, aquela que sofreu a quebra na dupla fita) e é diversamente conhecido como "conversão gênica sem crossover" ou "conversão gênica de trecho curto", pois leva à transferência de informações genéticas do doador para o alvo. Sem a intenção de ficar preso a qualquer teoria em particular, tal transferência pode envolver a correção de erro de pareamento (*mismatch*) de DNA heteroduplex que se forma o alvo quebrado e o doador e/ou o "anelamento de fitas dependente de síntese", no qual o doador é utilizado para voltar a sintetizar informações genéticas que se tornarão parte do alvo e/ou processos relacionados. Tal HR especializada frequentemente resulta em uma alteração da sequência da molécula alvo de tal modo que parte ou toda a sequência do polinucleotídeo doador é incorporada no polinucleotídeo alvo.

[0063] Nos métodos da invenção, uma ou mais nucleases direcionadas, conforme aqui descritas, criam uma quebra da fita dupla (DSB) na sequência alvo (por exemplo, cromatina celular) em um sítio predeterminado (por exemplo, um gene ou lócus de interesse), e um polinucleotídeo "doador", tendo homologia à sequência nucleotídica na região da quebra, pode ser introduzido na célula. A presença da DSB demonstrou facilitar a integração da sequência doadora. Opcionalmente, a construção possui homologia com a sequência nucleotídica na região da quebra. A sequência doadora pode ser fisicamente integrada ou, alternativamente, o polinucleotídeo doador é utilizado como molde para o reparo da quebra via recombinação homóloga, resultando na introdução de toda ou de parte da sequência nucleotídica, como no doador, na cromatina da célula. Assim, uma primeira sequência na cromatina celular pode ser alterada e, em certas

modalidades, pode ser convertida em uma sequência presente em um polinucleotídeo doador. Desse modo, pode-se entender que o uso dos termos "substitui" ou "substituição" representam a substituição de uma sequência nucleotídica por outra (ou seja, a substituição de uma sequência no sentido informativo), e não requer necessariamente a substituição física ou química de um polinucleotídeo por outro.

[0064] Em qualquer um dos métodos aqui descritos, é possível utilizar pares adicionais de proteínas dedo de zinco para clivagem dupla-fita adicional de sítios alvo adicionais dentro da célula.

[0065] Em certas modalidades dos métodos para recombinação direcionada e/ou substituição e/ou alteração de uma sequência em uma região de interesse na cromatina celular, uma sequência cromossômica é alterada por recombinação homóloga com uma sequência nucleotídica exógena "doadora". Tal recombinação homóloga é estimulada pela presença de uma quebra de dupla fita na cromatina celular, se sequências homólogas à região da quebra estiverem presentes.

[0066] Em qualquer um dos métodos aqui descritos, a primeira sequência nucleotídica (a "sequência doadora") pode conter sequências que são homólogas, mas não idênticas, às sequências genômicas na região de interesse, estimulando, com isso, a inserção por recombinação homóloga de uma sequência não idêntica na região de interesse. Assim, em certas modalidades, partes da sequência doadora que são homólogas às sequências na região de interesse exibem entre cerca de 80 e 99% (ou qualquer número inteiro intermediário) de identidade de sequência com a sequência genômica que é substituída. Em outras modalidades, a homologia entre a sequência doadora e a genômica é maior que 99%, por exemplo, se somente 1 nucleotídeo diferir nas sequências doadoras e genômicas de mais de 100 pares de bases contíguos. Em certos casos, uma parte

não homóloga da sequência doadora pode conter sequências não presentes na região de interesse, de tal modo que sequências novas são introduzidas na região de interesse. Nesses casos, a sequência não homóloga é geralmente flanqueada por sequências de 50-1.000 pares de base (ou qualquer valor inteiro intermediário) ou qualquer número de pares de base maior que 1.000, que são homólogas ou idênticas às sequências na região de interesse. Em outras modalidades, a sequência doadora não é homóloga à primeira sequência e é inserida no genoma por mecanismos de recombinação não homóloga.

[0067] Qualquer um dos métodos aqui descritos pode ser utilizado para inativação parcial ou completa de uma ou mais sequências alvo em uma célula por integração direcionada da sequência doadora que perturba a expressão do(s) gene(s) de interesse. Linhagens celulares com genes parciais ou completamente inativados são providas também.

[0068] Além disso, os métodos de integração direcionada conforme aqui descritos podem também ser utilizados para integrar uma ou mais sequências exógenas. A sequência exógena de ácido nucleico pode compreender, por exemplo, um ou mais genes ou moléculas de cDNA, ou qualquer tipo de sequência codificadora ou não codificação, bem como um ou mais elementos de controle (por exemplo, promotores). Adicionalmente, a sequência exógena de ácido nucleico pode produzir uma ou mais moléculas de RNA (por exemplo, pequenos RNAs em forma de grampo de cabelo (shRNAs), RNAs inibitórios (RNAis), microRNAs (miRNAs), etc.).

[0069] "Clivagem" refere-se à quebra do esqueleto covalente de uma molécula de DNA. A clivagem pode ser iniciada por uma variedade de métodos incluindo, entre outros, hidrólise enzimática ou química de uma ligação fosfodiéster. Tanto a clivagem da fita simples

como a clivagem da dupla fita são possíveis e, a clivagem da fita dupla pode ocorrer como resultado de dois eventos distintos de clivagem de fita simples. A clivagem do DNA pode resultar na produção de extremidades abruptas (*blunt*) ou extremidades escalonadas (*staggered*). Em certas modalidades, polipeptídeos de fusão são utilizados para clivagem direcionada de DNA dupla-fita.

[0070] Um "meio-domínio de clivagem" é uma sequência polipeptídica que, em conjunto com um segundo polipeptídeo (quer idêntico ou diferente) forma um complexo com atividade de clivagem (de preferência, atividade de clivagem de fita dupla). Os termos "primeiro e segundo meio-domínio de clivagem"; "+ e - meios-domínios de clivagem" e "meio-domínio de clivagem direito e esquerdo" são usados alternadamente para se referir a pares de meios-domínios de clivagem que dimerizam.

[0071] Um "meio-domínio de clivagem projetado" é um meio-domínio de clivagem que foi modificado para que forme obrigatoriamente heterodímeros com outro meio-domínio de clivagem (por exemplo, outro meio-domínio de clivagem projetado). Ver, também, as Patentes U.S. Nós 7 888 121; 7 914 796; 8 034 598; 8 623 618 e Publicação de Patente U.S. N° 2011/0201055, cujos conteúdos são aqui incorporados, em sua totalidade, por referência neste pedido de patente.

[0072] O termo "sequência" refere-se a uma sequência nucleotídica de qualquer comprimento, que pode ser de DNA ou RNA; pode ser linear, circular ou ramificada e pode ser fita simples ou dupla fita. O termo "sequência doadora" refere-se a uma sequência nucleotídica que é inserida em um genoma. Uma sequência doadora pode ter qualquer comprimento, por exemplo, entre 2 e 10.000 nucleotídeos de comprimento (ou qualquer valor inteiro intermediário ou superior), de preferência entre cerca de 100 e 1.000 nucleotídeos

de comprimento (ou qualquer número inteiro intermediário), mais preferivelmente entre cerca de 200 e 500 nucleotídeos de comprimento.

[0073] "Cromatina" é a estrutura nucleoproteica compreendendo o genoma celular. A cromatina celular compreende ácido nucleico, primariamente DNA, e proteína, incluindo proteínas cromossômicas com histonas e sem histonas. Na maioria dos eucarióticos, a cromatina celular existe na forma de nucleossomos, em que um núcleo de nucleossoma compreende aproximadamente 150 pares de base de DNA associados com um octâmero compreendendo duas de cada das histonas H2A, H2B, H3 e H4; e o DNA de ligação (de comprimento variável dependendo do organismo) estende-se entre os núcleos dos nucleossomos. Uma molécula de histona H1 é geralmente associada com o DNA de ligação. Para fins da presente invenção, o termo "cromatina" visa abranger todos os tipos de nucleoproteína celular, tanto procariótica como eucariótica. A cromatina celular inclui a cromatina de cromossomos e episomos.

[0074] Um "cromossomo" é um complexo de cromatina compreendendo todo ou uma parte do genoma de uma célula. O genoma de uma célula é frequentemente caracterizado por seu cariotipo, o qual é a coleção de todos os cromossomos que compreendem o genoma da célula. O genoma de uma célula pode compreender um ou mais cromossomos.

[0075] Um "epissomo" é um ácido nucleico em replicação, um complexo nucleoproteico ou outra estrutura compreendendo um ácido nucleico que não faz parte do cariotipo cromossômico de uma célula. Exemplos de episomas incluem plasmídeos e certos genomas virais.

[0076] Um "sítio alvo" ou "sequência alvo" é uma sequência de ácido nucleico que define uma parte de um ácido nucleico ao qual uma molécula de ligação se ligará, desde que existam condições

suficientes para a ligação. Por exemplo, a sequência 5' GAATTC 3' é um sítio alvo para a endonuclease de restrição Eco RI.

[0077] Uma molécula "exógena" é uma molécula que não está normalmente presente em uma célula, mas que pode ser introduzida em uma célula por um ou mais métodos genéticos, bioquímicos ou outros. A "presença normal na célula" é determinada em relação ao estágio de desenvolvimento em particular e às condições no ambiente da célula. Assim, por exemplo, uma molécula que está presente apenas durante o desenvolvimento embrionário do músculo é uma molécula exógena com respeito a uma célula muscular adulta. Do mesmo modo, uma molécula induzida por choque térmico é uma molécula exógena com respeito a uma célula não submetida a choque térmico. Uma molécula exógena pode compreender, por exemplo, uma versão funcionante de uma molécula endógena disfuncionante ou uma versão disfuncionante de uma molécula endógena que funciona normalmente.

[0078] Uma molécula exógena pode ser, entre outros, uma pequena molécula, como é quando gerada por um processo químico combinatório, ou uma macromolécula, como uma proteína, ácido nucleico, carboidrato, lipídio, glicoproteína, lipoproteína, polissacarídeo, qualquer derivado modificado das moléculas acima ou qualquer complexo compreendendo uma ou mais das moléculas acima. Os ácidos nucleicos incluem DNA e RNA, podem ser fita simples ou dupla fita; podem ser lineares, ramificados ou circulares; e podem ser de qualquer comprimento. Ver, por exemplo, as Patentes U.S. N°s 8 703 489 e 9 255 259. Os ácidos nucleicos incluem aqueles capazes de formar duplexos, bem como ácidos nucleicos que formam triplex. Ver, por exemplo, as Patentes U.S. N°s 5 176 996 e 5 422 251. As proteínas incluem, entre outras, proteínas de ligação ao DNA, fatores de transcrição, fatores remodeladores de cromatina, proteínas

de ligação a DNA metilado, polimerases, metilases, desmetilases, acetilases, desacetilases, quinases, fosfatases, integrases, recombinases, ligases, topoisomerases, girases e helicases.

[0079] Uma molécula exógena pode ser do mesmo tipo que uma molécula endógena, por exemplo, uma proteína ou ácido nucleico exógeno. Por exemplo, um ácido nucleico exógeno pode compreender um genoma viral infeccioso, um plasmídeo ou episomo introduzido em uma célula, ou um cromossomo que não está normalmente presente na célula. Métodos para a introdução de moléculas exógenas em células são conhecidos pelos técnicos no assunto e incluem, entre outros, transferência mediada por lipídios (ou seja, lipossomas, incluindo lipídios neutros e catiônicos), eletroporação, injeção direta, fusão celular, bombardeio de partículas, co-precipitação mediada com fosfato de cálcio, transferência mediada por DEAE-dextrano e transferência mediada por vetor viral. Uma molécula exógena pode também ser do mesmo tipo de molécula que uma molécula endógena, mas derivada de uma espécie diferente da célula da qual é derivada. Por exemplo, uma sequência de ácido nucleico humano pode ser introduzida em uma linhagem celular originalmente derivada de um camundongo ou hamster.

[0080] Por outro lado, uma molécula "endógena" é aquela que está normalmente presente em uma célula em particular em um estágio de desenvolvimento em particular sob condições ambientais em particular. Por exemplo, um ácido nucleico endógeno pode compreender um cromossomo, o genoma de uma mitocôndria, cloroplasto ou outra organela, ou um ácido nucleico episomal natural. Moléculas endógenas adicionais podem incluir proteínas, por exemplo, fatores de transcrição e enzimas.

[0081] Uma molécula de "fusão" é uma molécula na qual duas ou mais moléculas subunitárias são ligadas, de preferência

covalentemente. As moléculas subunitárias podem ser o mesmo tipo químico de moléculas, ou podem ser tipos químicos diferentes de moléculas. Exemplos do primeiro tipo de molécula de fusão incluem, entre outros, proteínas de fusão (por exemplo, uma fusão entre um domínio de ligação ao DNA de ZFP ou TALE e um ou mais domínios de ativação) e ácidos nucleicos de fusão (por exemplo, um ácido nucleico que codifica a proteína de fusão descrita acima). Exemplos do segundo tipo de molécula de fusão incluem, entre outros, uma fusão entre um ácido nucleico formador de triplex e um polipeptídeo, e uma fusão entre um ligante de sulco menor e um ácido nucleico. O termo também inclui sistemas nos quais um componente polinucleotídico associa-se com um componente polipeptídico para formar uma molécula funcional (por exemplo, um sistema CRISPR/Cas no qual um RNA guia único associa-se com um domínio funcional para modular a expressão gênica).

[0082] A expressão de uma proteína de fusão em uma célula pode resultar da entrega da proteína de fusão à célula ou pela entrega de um polinucleotídeo codificador da proteína de fusão a uma célula, em que o polinucleotídeo é transcrito, e o transcrito é traduzido, para gerar a proteína de fusão. *Trans-splicing*, a clivagem de polipeptídeos e a ligação de polipeptídeos podem também estar envolvidos na expressão de uma proteína em uma célula. Métodos para entrega de polinucleotídeos e polipeptídeos a células são apresentados em outro lugar neste relatório descriptivo.

[0083] Um "gene", para fins da presente invenção, inclui uma região do DNA que codifica um produto do gene (ver *infra*), bem como todas as regiões do DNA que regulam a produção do produto gênico, estejam ou não tais sequências reguladoras adjacentes às sequências de codificação e/ou transcritas. Desse modo, um gene inclui, mas não necessariamente limita-se aos mesmos, sequências promotoras,

terminadoras, sequências reguladoras da tradução como sítios de ligação a ribossomos e sítios internos de entrada do ribossomo, intensificadores, silenciadores, isoladoras, elementos de limites, origens de replicação, sítios de ligação à matriz e regiões de controle de lócus.

[0084] Um lócus "safe harbor" é um lócus dentro do genoma em onde um gene pode ser inserido sem quaisquer efeitos prejudiciais à célula hospedeira. O mais benéfico é um lócus *safe harbor* no qual a expressão da sequência gênica inserida não seja perturbada por qualquer expressão *read-through* (transleitura) de genes vizinhos. Exemplos não limitantes de *loci safe harbor* que são alvos de nuclease(s) incluem CCR5, CCR5, HPRT, AAVS1, Rosa e albumina. Ver, por exemplo, as Patentes U.S. N°s 8 771 985; 8 110 379; 7 951 925; as Publicações de Patentes U.S. N°s 2010/0218264; 2011/0265198; 2013/0137104; 2013/0122591; 2013/0177983; 2013/0177960; 2015/0056705; e 2015/0159172).

[0085] "Expressão gênica" refere-se à conversão das informações, contidas em um gene, em um produto gênico. Um produto gênico pode ser o produto direto da transcrição de um gene (por exemplo, mRNA, tRNA, rRNA, RNA *antisense*, ribozima, RNA estrutural ou qualquer outro tipo de RNA) ou uma proteína produzida pela tradução de um mRNA. Os produtos gênicos também incluem RNAs que são modificados por processos como capeamento (*capping*), poliadenilação, metilação e edição, e proteínas modificadas, por exemplo, por metilação, acetilação, fosforilação, ubiquitinação, ADP-ribosilação, miristilação e glicosilação.

[0086] "Modulação" ou "modificação" da expressão gênica refere-se a uma alteração na atividade de um gene. A modulação da expressão pode incluir, entre outros, ativação gênica e repressão gênica, inclusive por modificação do gene através da ligação de uma

molécula exógena (por exemplo, fator de transcrição projetado). A modulação pode também ser alcançada por modificação da sequência do gene via edição genômica (por exemplo, clivagem, alteração, inativação, mutação aleatória). Inativação gênica refere-se a qualquer redução na expressão do gene em comparação a uma célula que não foi modificada como aqui descrito. Assim, a inativação do gene pode ser parcial ou completa.

[0087] Uma "região de interesse" é qualquer região da cromatina celular, como, por exemplo, um gene ou uma sequência não codificante dentro ou adjacente a um gene, na qual se deseja ligar uma molécula exógena. A ligação pode ter por objetivo a clivagem direcionada do DNA e/ou uma recombinação direcionada. Uma região de interesse pode estar presente em um cromossomo, um epissomo, um genoma organelar (por exemplo, de mitocôndria, cloroplasto) ou um genoma viral infeccioso, por exemplo. Uma região de interesse pode ser dentro da região de codificação de um gene, dentro de regiões transcritas não codificantes como, por exemplo, sequências líderes, sequências *trailer* ou íntrons, ou dentro de regiões não transcritas, quer a montante ou a jusante da região de codificação. Uma região de interesse pode ser tão pequena como um único par de nucleotídeos ou com até 2.000 pares de nucleotídeos de comprimento, ou qualquer valor inteiro de pares de nucleotídeos.

[0088] As células "eucarióticas" incluem, entre outras, células de fungos (como leveduras), células vegetais, células animais, células de mamíferos e células humanas (por exemplo, células T).

[0089] Os termos "ligação operativa" e "operativamente ligado" (ou "operacionalmente ligado") são usados alternadamente com referência a uma justaposição de dois ou mais componentes (como elementos da sequência), em que os componentes estão dispostos de tal modo que os dois componentes atuam normalmente e permitem a possibilidade

de que pelo menos um dos componentes possa mediar uma função que é exercida sobre pelo menos um dos outros componentes. A título de ilustração, uma sequência reguladora da transcrição, como um promotor, é operativamente ligada a uma sequência codificadora se a sequência reguladora da transcrição controla o nível de transcrição da sequência codificadora em resposta à presença ou ausência de um ou mais fatores reguladores da transcrição. Uma sequência reguladora da transcrição é em geral operativamente ligada em *cis* com uma sequência codificadora, mas não precisa estar diretamente adjacente à mesma. Por exemplo, um intensificador é uma sequência reguladora da transcrição que está operativamente ligada a uma sequência codificadora, muito embora não sejam contíguos.

[0090] Em relação aos polipeptídeos de fusão, o termo "operativamente ligado" pode referir-se ao fato de que cada um dos componentes executa a mesma função na ligação a outro componente, como o fariam caso não estivessem assim ligados. Por exemplo, em relação a um polipeptídeo de fusão no qual um domínio de ligação ao DNA (por exemplo, ZFP, TALE) é fundido a um domínio de ativação, o domínio de ligação ao DNA e o domínio de ativação estão em ligação operativa se, no polipeptídeo de fusão, a porção do domínio de ligação ao DNA for capaz de se ligar ao seu sítio alvo e/ou seu sítio de ligação, enquanto o domínio de ativação é capaz de regular positivamente a expressão gênica. Quando um polipeptídeo de fusão no qual um domínio de ligação ao DNA é fundido a um domínio de clivagem, o domínio de ligação ao DNA e o domínio de clivagem estão em ligação operativa se, no polipeptídeo de fusão, a porção do domínio de ligação ao DNA for capaz de se ligar ao seu sítio alvo e/ou seu sítio de ligação, enquanto o domínio de clivagem é capaz de clivar o DNA na vizinhança do sítio alvo. Do mesmo modo, com respeito a um polipeptídeo de fusão no qual um domínio de ligação ao DNA é

fundido a um domínio de ativação ou repressão, o domínio de ligação ao DNA e o domínio de ativação ou repressão estão em ligação operativa se, no polipeptídeo de fusão, a porção do domínio de ligação ao DNA for capaz de se ligar ao seu sítio alvo e/ou seu sítio de ligação, enquanto o domínio de ativação é capaz de regular positivamente a expressão gênica ou o domínio de repressão é capaz de regular negativamente a expressão gênica.

[0091] Um "fragmento funcional" de uma proteína, polipeptídeo ou ácido nucleico é uma proteína, polipeptídeo ou ácido nucleico cuja sequência não é idêntica à proteína, polipeptídeo ou ácido nucleico completo, mas que, não obstante, retém a mesma função que a proteína, polipeptídeo ou ácido nucleico completo. Um fragmento funcional pode possuir mais, menos ou o mesmo número de resíduos que a molécula nativa correspondente, e/ou pode conter uma ou mais substituições de aminoácidos ou nucleotídeos. Métodos para determinar a função de um ácido nucleico (por exemplo, função codificadora, capacidade para hibridizar com outro ácido nucleico) são bem conhecidos na técnica. Do mesmo modo, métodos para determinar a função da proteína são bem conhecidos. Por exemplo, a função de ligação ao DNA de um polipeptídeo pode ser determinada, por exemplo, por filtro de ligação, desvio de mobilidade eletroforética ou ensaios de imunoprecipitação. A clivagem do DNA pode ser analisada por eletroforese em gel. Ver Ausubel, et al., *supra*. A capacidade de uma proteína para interagir com outra proteína pode ser determinada, por exemplo, por co-imunoprecipitação, ensaios de dois híbridos ou complementação, ambas, genética e bioquímica. Ver, por exemplo, Fields, et al. (1989) *Nature* 340:245-246; a Patente U.S. N° 5 585 245 e Publicação da Patente Internacional N° WO 98/44350.

[0092] Um "vetor" é capaz de transferir sequências gênicas para células alvo. Tipicamente, "construção de vetor", "vetor de expressão"

e "vetor de transferência gênica" significam qualquer construção de ácido nucleico capaz de direcionar a expressão de um gene de interesse e que pode transferir sequências gênicas a células alvo. Assim, o termo inclui clonagem e veículos de expressão, bem como vetores de integração.

[0093] Um "gene repórter" ou "sequência repórter" refere-se a qualquer sequência que produz um produto proteico que é facilmente medido, de preferência, embora não necessariamente em um ensaio rotineiro. Os genes repórteres adequados incluem, entre outros, sequências codificadoras de proteínas que atuam na mediação de resistência a antibióticos (por exemplo, resistência à ampicilina, resistência à neomicina, resistência a G418, resistência à puromicina), sequências codificadoras de proteínas coloridas ou fluorescentes ou luminescentes (por exemplo, proteína verde fluorescente, proteína verde fluorescente intensificada, proteína vermelha fluorescente, luciferase), e proteínas que atuam na mediação de crescimento celular reforçado e/ou amplificação do gene (por exemplo, dihidrofolato redutase). *Tags* (etiquetas) de epítopos incluem, por exemplo, uma ou mais cópias de FLAG, His, myc, Tap, HA ou qualquer sequência detectável de aminoácidos. "Tags de expressão" incluem sequências que codificam repórteres que podem ser operacionalmente ligados a uma sequência gênica desejada a fim de monitorar a expressão do gene de interesse.

[0094] Os termos "indivíduo" e "paciente" são usados alternadamente e referem-se a mamíferos como pacientes humanos e primatas não humanos, bem como animais experimentais como coelhos, cães, gatos, ratos, camundongos e outros animais. Assim, o termo "indivíduo" ou "paciente", neste relatório descritivo, significa qualquer paciente mamífero ou indivíduo ao qual os cassete de expressão podem ser administrados. Os indivíduos da presente

invenção incluem aqueles com um transtorno ou aqueles em risco de desenvolver um transtorno.

[0095] Os termos "tratar" e "tratamento", neste relatório descritivo, referem-se à redução na gravidade e/ou frequência de sintomas, a eliminação de sintomas e/ou da causa subjacente, a prevenção da ocorrência de sintomas e/ou das suas causas subjacentes e a melhora ou correção de danos. Câncer e doença do enxerto contra hospedeiro são exemplos não limitantes de condições que podem ser tratadas utilizando as composições e métodos aqui descritos. Assim, "tratar" e "tratamento" inclui:

prevenir a ocorrência da doença ou condição em um mamífero, especialmente, quanto tal mamífero é predisposto à condição, mas ainda não foi diagnosticado como portador;

inibir a doença ou condição, ou seja, interromper seu desenvolvimento;

aliviar a doença ou condição, ou seja, provocar a regressão da doença ou condição; ou

aliviar os sintomas resultantes da doença ou condição, ou seja, aliviar a dor sem abordar a doença ou condição subjacente.

[0096] Neste relatório descritivo, os termos "doença" e "condição" podem ser utilizados alternadamente ou podem ser diferentes pelo fato de que a enfermidade ou condição em particular pode não ter um agente causador conhecido (de modo que a etiologia não foi ainda esclarecida) e, portanto, não é ainda reconhecida como doença, mas somente uma condição ou síndrome indesejável, em que um conjunto mais ou menos específico de sintomas foi identificado por clínicos.

[0097] Uma "composição farmacêutica" refere-se a uma formulação de um composto da invenção e um meio em geral aceito na técnica para a entrega do composto biologicamente ativo a mamíferos, por exemplo, humanos. Tal meio inclui todos os veículos,

diluentes ou excipientes farmaceuticamente aceitáveis para o mesmo.

[0098] "Quantidade eficaz" ou "quantidade terapeuticamente eficaz" refere-se àquela quantidade de um composto da invenção que, quando administrada a um mamífero, de preferência um humano, é suficiente para efetuar o tratamento no mamífero, de preferência um humano. A quantidade de uma composição da invenção que constitui uma "quantidade terapeuticamente eficaz" variará dependendo do composto, da condição e sua gravidade, do modo de administração e da idade do mamífero a ser tratado, mas pode ser determinada rotineiramente por qualquer técnico no assunto levando em consideração seu próprio conhecido e esse relatório descritivo.

[0099] Domínios de ligação ao DNA

[00100] São aqui descritas composições compreendendo um domínio de ligação ao DNA que se liga especificamente a um sítio alvo em qualquer gene compreendendo um gene HLA ou um regulador de HLA. Qualquer domínio de ligação ao DNA pode ser utilizado nas composições e métodos aqui revelados, incluindo, entre outros, um domínio de ligação ao DNA de dedo de zinco, um domínio de ligação ao DNA de TALE, a porção de ligação ao DNA (sgRNA) de uma nuclease CRISPR/Cas ou um domínio de ligação ao DNA de uma meganuclease. O domínio de ligação ao DNA pode ligar-se a qualquer sequência alvo dentro do gene, incluindo, entre outras, uma sequência alvo de 12 ou mais nucleotídeos como mostrado em qualquer um dentre os sítios alvo aqui revelados (SEQ ID Nº:8-21 e/ou 92-103).

[00101] Em certas modalidades, o domínio de ligação ao DNA compreende uma proteína dedo de zinco. De preferência, a proteína dedo de zinco não é natural pelo fato de ser projetada para ligar-se a um sítio alvo de escolha. Ver, por exemplo, 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; as Patentes U.S. N^{os} 6 453 242; 6 534 261; 6 599 692; 6 503 717; 6 689 558; 7 030 215; 6 794 136; 7 067 317; 7 262 054; 7 070 934; 7 361 635; e 7 253 273; e Publicações de Patentes U.S. N^{os} 2005/0064474; 2007/0218528; e 2005/0267061, todos aqui incorporados, em sua totalidade, por referência. Em certas modalidades, o domínio de ligação ao DNA compreende uma proteína dedo de zinco revelada na Publicação de Patente U.S. N° 2012/0060230 (por exemplo, Tabela 1), aqui incorporado, em sua totalidade, por referência neste pedido de patente. Em outras modalidades, o domínio de ligação ao DNA compreende o componente ZFP (referido como "desenhos") e incluindo regiões de reconhecimento de hélice e esqueletos conforme apresentados nas ZFNs das Tabelas 1, 2, 4, 5, 6 ou 8, incluindo, entre outras, os domínios de ZFP das ZFNs 72732; 72748; 68957; ou 72678.

[00102] Um domínio de ligação projetado de dedo de zinco pode ter uma nova especificidade de ligação, quando comparado a uma proteína dedo de zinco natural. Os métodos de engenharia incluem, entre outros, desenho racional e vários tipos de seleção. O desenho racional inclui, por exemplo, o uso de banco de dados compreendendo sequências de nucleotídeos triplas (ou quádruplas) e sequências de aminoácidos de dedos de zinco individuais, em que cada sequência nucleotídeo tripla ou quádrupla está associada com uma ou mais sequências de aminoácidos de dedos de zinco que se ligam à sequência tripla ou quádrupla em particular. Ver, por exemplo, as Patentes U.S. N^{os} 6 453 242 e 6 534 261, aqui incorporadas, em sua totalidade, por referência neste pedido de patente.

[00103] Métodos exemplares de seleção, incluindo expressão em fagos e sistemas de dois híbridos, são revelados nas Patentes U.S. N^{os} 5 789 538; 5 925 523; 6 007 988; 6 013 453; 6 410 248; 6 140 466;

6 200 759; e 6 242 568; bem como Publicações de Patente Internacional Nos. WO 98/37186; WO 98/53057; WO 00/27878; e WO 01/88197 e GB 2,338,237. Além disso, o aprimoramento da especificidade de ligação para domínios de ligação dedos de zinco foi descrito, por exemplo, na Patente U.S. Nº 6 794 136.

[00104] Além do mais, como revelado nessas e em outras referências, domínios de dedo de zinco e/ou proteínas dedos de zinco multi-dedo podem ser unidos utilizando quaisquer sequências ligadoras adequadas (ligante), incluindo por exemplo, ligantes de 5 ou mais aminoácidos de comprimento. Ver, também, as Patentes U.S. Nºs 6 479 626; 6 903 185; e 7 153 949 para sequências exemplares de ligação com 6 ou mais aminoácidos de comprimento. As proteínas aqui descritas podem incluir qualquer combinação de ligantes adequados entre os dedos individuais de zinco da proteína. Além disso, o aprimoramento da especificidade de ligação para domínios de ligação dedos de zinco foi descrito, por exemplo, na Patente U.S. Nº 6 794 136.

[00105] A seleção de sítios alvo; ZFPs e métodos para o desenho e construção de proteínas de fusão (e de polinucleotídeos codificadores das mesmas) são conhecidos pelos técnicos no assunto e descritos detalhadamente nas Patentes U.S. Nºs 6 140 081; 5 789 538; 6 453 242; 6 534 261; 5 925 523; 6 007 988; 6 013 453; 6 200 759; e Publicações de Patente Internacional Nºs 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; e WO 03/016496.

[00106] Além disso, como revelado nessas e em outras referências, domínios dedo de zinco e/ou proteínas multi-dedo de dedos de zinco podem ser unidos utilizando quaisquer sequências ligadoras adequadas, incluindo por exemplo, ligantes com 5 ou mais

aminoácidos de comprimento. Ver, também, as Patentes U.S. Nos 6 479 626; 6 903 185; 7 153 949; 7 888 121; 7 914 796; 8 034 598; 8 623 618; 9 567 609; e Publicação de Patente U.S. Nº 2017/0218349 para sequências ligadoras exemplares. As proteínas aqui descritas podem incluir qualquer combinação de ligantes adequados entre os dedos individuais de zinco da proteína.

[00107] Em certas modalidades, o domínio de ligação ao DNA é uma proteína dedo de zinco projetada que se liga (de maneira sequência-específica) a um sítio alvo em um gene TCR ou gene regulador de TCR e modula a expressão de um gene TCR. Em algumas modalidades, a proteína dedo de zinco liga-se a um sítio alvo em TCRA enquanto, em outras modalidades, o dedo de zinco liga-se a um sítio alvo em TRBC. Em outras modalidades, o domínio de ligação ao DNA é uma proteína dedo de zinco projetada que se liga (de maneira sequência-específica) a um sítio alvo em um gene B2M e modula a expressão de um gene B2M. Modalidades exemplares não limitantes desses domínios de ligação ao DNA são mostradas nas Tabelas 1, 2 e 6 (TCR) e Tabelas 5 e 8 (B2M). Em certas modalidades, a ZFP compreende a porção ZFP das ZFNs designadas 72732; 72748; 68957; ou 72678.

[00108] Normalmente, as ZFPs incluem pelo menos três dedos. Algumas das ZFPs incluem quatro, cinco ou seis dedos. As ZFPs que incluem três dedos, tipicamente, reconhecem um sítio alvo que inclui 9 ou 10 nucleotídeos; ZFPs que incluem quatro dedos, tipicamente, reconhecem um sítio alvo que inclui 12 a 14 nucleotídeos; enquanto ZFPs com seis dedos podem reconhecer sítios alvo que incluem 18 a 21 nucleotídeos. As ZFPs podem também ser proteínas de fusão que incluem um ou mais domínios reguladores, e esses domínios podem ser domínios de ativação ou repressão da transcrição.

[00109] Em algumas modalidades, o domínio de ligação ao DNA

pode ser derivado de uma nuclease. Por exemplo, as sequências de reconhecimento de endonucleases teleguiadas (em inglês, *homing*) e meganucleases como as conhecidas I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-Csml, I-PanI, I-SceII, I-Ppol, I-SceIII, I-CreI, I-TevI, I-TevII e I-TevIII. Ver também a Patente U.S. Nº 5 420 032; Patente U.S. Nº 6 833 252; 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 e o catálogo da New England Biolabs. Além disso, a especificidade de ligação ao DNA de endonucleases *homing* e meganucleases pode também ser projetada para ligação a sítios alvo não naturais. Ver, por exemplo, 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; Pâques, *et al.* (2007) *Current Gene Therapy* 7:49-66; Publicação de Patente U.S. Nº 2007/0117128.

[00110] Em outras modalidades, o domínio de ligação ao DNA compreende um domínio projetado de um efetor TAL semelhante àqueles derivados do patógeno de plantas *Xanthomonas* (ver Boch, *et al.* (2009) *Science* 326: 1509-1512 e Moscou e Bogdanove (2009) *Science* 326:1501) e *Ralstonia* (ver Heuer, *et al.* (2007) *Applied and Environmental Microbiology* 73(13): 4379-4384); Publicações de Patentes U.S. Nºs 2011/0301073 e 2011/0145940. As bactérias patogênicas de plantas do gênero *Xanthomonas* são conhecidas por provocar muitas doenças em importantes plantas de cultivo. A patogenicidade de *Xanthomonas* depende de um sistema conversado de secreção do tipo III (T3S) que injeta mais de 25 proteínas efetoras diferentes na célula vegetal. Entre essas proteínas injetadas estão efetores do tipo ativador da transcrição (TALE) que mimetizam

ativadores transpcionais vegetais e manipulam o transcriptoma da planta (ver Kay, et al. (2007) *Science* 318:648-651). Essas proteínas contêm um domínio de ligação ao DNA e um domínio de ativação transicional. Uma das TALEs mais bem caracterizadas é AvrBs3 de *Xanthomonas campestris* pv. *Vesicatoria* (ver Bonas, et al. (1989) *Mol Gen Genet* 218: 127-136 e a Publicação da Patente Internacional Nº WO 2010/079430). TALEs contêm um domínio centralizado de repetições em tandem, cada repetição contendo aproximadamente 34 aminoácidos, que são fundamentais para a especificidade de ligação ao DNA dessas proteínas. Além disso, contêm uma sequência de localização nuclear e um domínio de ativação transicional ácido (para uma revisão, ver Schornack S., et al. (2006) *J Plant Physiol* 163(3): 256-272). Além disso, foi verificado que, nas bactérias fitopatogênicas *Ralstonia solanacearum*, dois genes, designados brg11 e hpx17, são homólogos à família AvrBs3 de *Xanthomonas* na cepa *R. solanacearum* biovar 1 GMI1000 e na cepa biovar 4 RS1000 (ver Heuer, et al. (2007) *Appl and Envir Micro* 73(13):4379-4384). Esses genes são 98,9% idênticos em sequência de nucleotídeo entre si, mas diferem por uma deleção de 1.575 bp no domínio de repetição de hpx17. Contudo, ambos os produtos gênicos têm menos de 40% de identidade de sequência com proteínas da família AvrBs3 de *Xanthomonas*.

[00111] A especificidade desses efetores TAL depende das sequências encontradas nas repetições em tandem. A sequência repetida compreende aproximadamente 102 pares de base, e as repetições são tipicamente 91-100% homólogas entre si (Bonas, et al., *ibid*). O polimorfismo das repetições está normalmente localizado nas posições 12 e 13, e parece ser uma correspondência um para um entre a identidade dos di-resíduos hipervariáveis (a repetição de di-resíduos variáveis ou região RVD) nas posições 12 e 13 com a

identidade dos nucleotídeos contíguos na sequência alvo efetora-TAL (ver Moscou and Bogdanove (2009) *Science* 326:1501 e Boch, *et al.* (2009) *Science* 326:1509-1512). Experimentalmente, o código natural para reconhecimento de DNA desses efetores-TAL foi determinado tal que uma sequência HD nas posições 12 e 13 (Repetição de di-resíduo variável ou RVD) leva a uma ligação a citosina (C), NG liga-se a T, NI a A, C, G ou T, NN liga-se a A ou G e ING liga-se a T. Essas repetições de ligação de DNA foram montadas em proteínas com novas combinações e números de repetições, para produzir fatores de transcrição artificiais que são capazes de interagir com novas sequências e ativar a expressão de um gene repórter não endógeno em células de planta (Boch, *et al.*, *ibid*). As proteínas TAL projetadas foram ligadas a um meio-domínio de clivagem FokI para produzir uma fusão de nuclease de domínio efetor TAL (TALEN), incluindo TALENs com RVDs atípicas. Ver, por exemplo, Patente U.S. Nº 8 586 526.

[00112] Em algumas modalidades, a TALEN compreende um domínio de clivagem meio-domínio de clivagem de endonuclease (por exemplo, FokI). Em outras modalidades, a nuclease TALE é uma mega TAL. Essas mega TAL nucleases são proteínas de fusão compreendendo um domínio de ligação ao DNA de TALE DNA e um domínio de clivagem de meganuclease. O domínio de clivagem de meganuclease é ativo como monômero e não requer dimerização para atividade (ver Boissel, *et al.* (2013) *Nucl Acid Res*: 1-13, doi: 10.1093/nar/gkt1224).

[00113] Em ainda outras modalidades, a nuclease compreende uma TALEN compacta. Essas são proteínas de fusão de cadeia única ligando um domínio de ligação ao DNA de TALE DNA a um domínio nuclease Tevl. A proteína de fusão pode atuar como uma nickase localizada pela região TALE, ou pode criar uma quebra de fita dupla, dependendo de onde o domínio de ligação ao DNA de TALE está

localizado em relação ao domínio nuclease TeV1 (ver Beurdeley, *et al.* (2013) *Nat Comm* 4:1762 DOI: 10.1038/ncomms2782). Além disso, o domínio nuclease pode também exibir funcionalidade de ligação ao DNA. Quaisquer TALENs podem ser utilizadas em combinação com TALENs adicionais (por exemplo, uma ou mais TALENs (cTALENs ou FokI-TALENs) com uma ou mais mega-TALEs.

[00114] Além disso, como revelado nessas e em outras referências, domínios dedo de zinco e/ou proteínas multi-dedo dedo de zinco ou TALEs podem ser unidos utilizando quaisquer sequências ligadoras adequadas, incluindo, por exemplo, ligantes de 5 ou mais aminoácidos de comprimento. Ver, também, as Patentes U.S. N°s 6 479 626; 6 903 185; e 7 153 949 para sequências ligadoras exemplares com 6 ou mais aminoácidos de comprimento. As proteínas aqui descritas podem incluir qualquer combinação de ligantes adequados entre os dedos de zinco individuais da proteína. Além disso, o aprimoramento da especificidade de ligação para domínios de ligação dedo de zinco foi descrito, por exemplo, na Patente U.S. N° 6 794 136.

[00115] Em certas modalidades, o domínio de ligação ao DNA faz parte de um sistema de nuclease CRISPR/Cas, incluindo um RNA guia único (sgRNA) que se liga a DNA. Ver, por exemplo, Patente U.S. N° 8 697 359 e Publicações de Patente U.S. N°s 2015/0056705 e 2015/0159172. O lócus CRISPR (do inglês, *Clustered Regularly Interspaced Short Palindromic Repeats*), ou seja, Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespacadas, que codifica componentes de RNA do sistema, e o lócus de Cas (CRISPR-associado), que codifica proteínas (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) compõem as sequências gênicas do sistema de nuclease CRISPR/Cas. Os *loci* CRISPR em hospedeiros

microbianos contêm uma combinação de genes CRISPR-associados (Cas) bem como elementos não codificantes de RNA capazes de programar a especificidade da clivagem de ácido nucleico mediada por CRISPR.

[00116] O CRISPR Tipo II é um dos sistemas mais bem caracterizados e realiza a quebra direcionada da fita dupla de DNA em quatro etapas sequenciais. Em primeiro lugar, dois RNA não codificantes, o arranjo de pré-crRNA e tracrRNA, são transcritos a partir do lócus de CRISPR. Em segundo lugar, tracrRNA hibridiza as regiões de repetição do pré-crRNA e medeia o processamento de pré-crRNA em crRNAs maduros contendo sequências espaçadoras individuais. Em terceiro lugar, o complexo crRNA maduro:tracrRNA direciona um domínio funcional (por exemplo, nuclease como Cas) ao DNA alvo através do pareamento de bases de Watson-Crick entre o espaçador no crRNA e o protoespaçador no DNA alvo ao lado do motivo protoespaçador adjacente (PAM), um requisito adicional para reconhecimento do alvo. Finalmente, a Cas9 medeia a clivagem do DNA alvo para criar uma quebra na fita dupla dentro do protoespaçador. A atividade do sistema CRISPR/Cas compreende três etapas: (i) inserção de sequências de DNA estranhas no arranjo de CRISPR para prevenir ataques futuros, em um processo denominado "adaptação", (ii) expressão das proteínas relevantes, bem como expressão e processamento do arranjo, seguida por (iii) interferência mediada por RNA com o ácido nucleico estranho. Dessa forma, na célula bacteriana, diversas das chamadas proteínas "Cas" estão envolvidas com a função natural do sistema CRISPR/Cas e desempenham papéis em funções como inserção do DNA estranho etc.

[00117] Em certas modalidades, a proteína Cas pode ser um "derivado funcional" de uma proteína Cas natural. Um "derivado

funcional" de um polipeptídeo de sequência nativa é um composto com uma propriedade biológica qualitativa em comum com um polipeptídeo de sequência nativa. Os "derivados funcionais" incluem, entre outros, fragmentos de uma sequência nativa e derivados de um polipeptídeo de sequência nativa e seus fragmentos, desde que estes tenham uma atividade biológica em comum com um polipeptídeo de sequência nativa correspondente. Uma atividade biológica aqui contemplada é a capacidade do derivado funcional de hidrolisar um substrato de DNA em fragmentos. O termo "derivado" abrange variantes de sequência de aminoácidos do polipeptídeo, modificações covalentes e suas fusões, tais como proteínas derivadas de Cas. Os derivados adequados de um polipeptídeo Cas ou seu fragmento incluem, entre outros, mutantes, fusões, modificações covalentes de proteína Cas ou seu fragmento. Proteína Cas, que inclui proteína Cas ou seu fragmento, bem como derivados da proteína Cas ou seu fragmento, pode ser obtida de uma célula ou sintetizada quimicamente ou uma combinação desses dois procedimentos. A célula pode ser uma célula que produz naturalmente proteína Cas, ou uma célula que produz naturalmente proteína Cas e geneticamente projetada para produzir a proteína Cas endógena em um nível de expressão mais alto ou produzir uma proteína Cas a partir de um ácido nucleico introduzido exogenamente, em que este ácido nucleico codifica uma Cas que é igual ou diferente da Cas endógena. Em alguns casos, a célula não produz naturalmente proteína C e é projetada geneticamente para produzir uma proteína Cas. Em algumas modalidades, a proteína Cas é um ortólogo pequeno Cas9 para entrega através de um vetor AAV (Ran, *et al.* (2015) *Nature* 510:186). [00118] Em algumas modalidades, o domínio de ligação ao DNA faz parte de um sistema TtAgo (ver Swarts, *et al.*, *ibid*; Sheng, *et al.*, *ibid*). Em eucariotos, o silenciamento gênico é mediado pela família Argonaute (Ago) de proteínas. Nesse paradigma, Ago é ligado a

pequenos RNAs (19-31 nt). Esse complexo proteína-RNA de silenciamento reconhece RNAs via pareamento de bases de Watson-Crick entre o pequeno RNA e o alvo e cliva o RNA alvo por meio de endonucleases (Vogel (2014) *Science* 344:972-973). Em contraste, proteínas Ago procarióticas ligam-se a pequenos fragmentos de DNA fita simples e provavelmente atuam na detecção e remoção de DNA exógeno (frequentemente viral) (Yuan, et al. (2005) *Mol. Cell* 19, 405; Olovnikov, et al. (2013) *Mol. Cell* 51:594; Swarts, et al., *ibid*). Proteínas Ago procarióticas exemplares incluem aquelas de *Aquifex aeolicus*, *Rhodobacter sphaeroides* e *Thermus thermophilus*.

[00119] Uma das proteínas Ago procarióticas mais bem caracterizadas é aquela de *T. thermophilus* (TtAgo; Swarts, et al., *ibid*). TtAgo associa-se com fragmentos de DNA fita simples de 15 nt ou 13-25 nt com grupos 5' fosfato. Esse "DNA guia" ligado por TtAgo tem como função direcionar o complexo proteína-DNA para ligar-se a uma sequência de DNA complementar Watson-Crick em uma molécula de DNA de terceiros. Uma vez as informações sobre sequência nesses DNAs guias tenha permitido a identificação do DNA alvo, o complexo TtAgo-DNA guia cliva o DNA alvo. Tal mecanismo é também apoiado pela estrutura do complexo TtAgo-DNA guia enquanto ligado a seu DNA alvo (G. Sheng et al., *ibid*). Ago de *Rhodobacter sphaeroides* (RsAgo) possui propriedades semelhantes (Olovnikov, et al., *ibid*).

[00120] DNAs guias exógenos de sequência arbitrária de DNA podem ser carregados na proteína TtAgo (Swarts, et al., *ibid*). Dado que a especificidade da clivagem por TtAgo é direcionada pelo DNA guia, um complexo TtAgo-DNA formado com um DNA guia exógeno, especificado pelo investigador, direcionará, portanto, a clivagem do DNA alvo por TtAgo para um DNA alvo complementar especificado pelo investigador. Dessa maneira, pode-se criar uma quebra de fita dupla direcionada no DNA. O uso do sistema TtAgo-DNA guia (ou

sistemas Ago-DNA guia ortólogos de outros organismos) permite a clivagem direcionada do DNA genômico dentro das células. Tal clivagem pode ser de fita simples ou dupla-fita. Para clivagem do DNA genômico de mamíferos, seria preferível utilizar uma versão de TtAgo com códons otimizados para expressão em células de mamíferos. Além disso, poderia ser preferível tratar células com complexo TtAgo-DNA formado *in vitro* onde a proteína TtAgo é fundida a um peptídeo que penetra a célula. Além disso, poderia ser preferível utilizar uma versão da proteína TtAgo que tenha sido alterada via mutagênese para ter atividade melhorada a 37°C. A clivagem de DNA mediada por Ago-RNA poderia ser utilizada para afetar uma panóplia de resultados, incluindo knockout (silenciamento) gênico, adição direcionada de gene, correção de gene, deleção direcionada de gene utilizando técnicas padrão no assunto para a exploração de quebras de DNA.

[00121] Assim, pode-se utilizar qualquer domínio de ligação ao DNA.

Moléculas de Fusão

[00122] São também providas moléculas de fusão que compreendem domínios de ligação ao DNA (por exemplo, ZFPs ou TALEs, componentes de CRISPR/Cas como RNAs guias únicos), conforme aqui descritos, associados com um domínio regulador heterólogo (funcional) (ou fragmento funcional do mesmo). Domínios comuns incluem, por exemplo, domínios de fatores de transcrição (ativadores, repressores, coativadores, co-repressores), silenciadores, oncogenes (por exemplo, membros das famílias myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos etc.); enzimas de reparo de DNA e seus fatores e modificadores associados; enzimas de rearranjo de DNA e seus fatores e modificadores associados; proteínas associadas à cromatina e seus modificadores (por exemplo, quinases, acetilases e desacetilases); e enzimas que modificam DNA (por exemplo,

metiltransferases, topoisomerase, helicases, ligases, quinases, fosfatases, polimerases, endonucleases) e seus fatores e modificadores associados. Tais moléculas de fusão incluem fatores de transcrição compreendendo os domínios de ligação ao DNA aqui descritos e um domínio regulador da transcrição bem como nucleases compreendendo os domínios de ligação ao DNA e um ou mais domínios nuclease.

[00123] Os domínios adequados para alcançar ativação (domínios de ativação transcricional) incluem o domínio de ativação VP16 de HSV (ver, por exemplo, Hagmann, *et al.* (1997) *J. Virol.* 71:5952-5962) receptores nucleares hormonais (ver, por exemplo, Torchia, *et al.* (1998) *Curr. Opin. Cell. Biol.* 10:373-383); a subunidade p65 do fator nuclear kappa B (Bitko & Barik (1998) *J. Virol.* 72:5610-5618 e Doyle & Hunt (1997) *Neuroreport* 8:2937-2942); Liu, *et al.* (1998) *Cancer Gene Ther.* 5:3-28) ou domínios funcionais químéricos artificiais como VP64 (Beerli, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:14623-33) e degron (Molinari, *et al.* (1999) *EMBO J.* 18, 6439-6447). Domínios de ativação exemplares adicionais incluem Oct 1, Oct-2A, Sp1, AP-2 e CTF1 (Seipel, *et al.* (1992) *EMBO J.* 11, 4961-4968 bem como p300, CBP, PCAF, SRC1 PvALF, AtHD2A e ERF-2. Ver, por exemplo, Robyr, *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; e Lemon, *et al.* (1999) *Curr. Opin. Genet. Dev.* 9:499-504. Domínios de ativação exemplares adicionais incluem, entre outros, OsGAI, HALF-1, C1, AP1, ARF-5,-6,-7, e -8, CPRF1, CPRF4, MYC-RP/GP e TRAB1. Ver, por exemplo, 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; e Hobo, et al. (1999) *Proc. Natl. Acad. Sci. USA* 96:15.348-15.353.

[00124] Ficará claro para os técnicos no assunto que, na formação de uma proteína de fusão (ou um ácido nucleico codificador da mesma) entre um domínio de ligação ao DNA e um domínio funcional, um domínio de ativação ou uma molécula que interaja com um domínio de ativação é adequado como domínio funcional. Essencialmente qualquer molécula capaz de recrutar um complexo ativador e/ou atividade que ativa (como, por exemplo, acetilação de histona) o gene alvo é útil como domínio ativador de uma proteína de fusão. São descritos domínios isoladores, domínios de localização e proteínas de remodelamento da cromatina como domínios contendo ISWI e/ou proteínas com domínio de ligação a metila para uso como domínios funcionais em moléculas de fusão, por exemplo, na Patente U.S. Nº 7 053 264.

[00125] Domínios de repressão exemplares incluem, entre outros, KRAB A/B, KOX, gene precoce induzível por TGF-beta (TIEG), v-erbA, SID, MBD2, MBD3, membros da família DNMT (por exemplo, DNMT1, DNMT3A, DNMT3B), Rb e MeCP2. Ver, por exemplo, Bird, et al. (1999) *Cell* 99:451-454; Tyler, et al. (1999) *Cell* 99:443-446; Knoepfler, et al. (1999) *Cell* 99:447-450; e Robertson, et al. (2000) *Nature Genet.* 25:338-342. Domínios de repressão exemplares adicionais incluem, entre outros, ROM2 e AtHD2A. Ver, por exemplo, Chem, et al. (1996) *Plant Cell* 8:305-321; e Wu, et al. (2000) *Plant J.* 22:19-27.

[00126] Moléculas de fusão são construídas por métodos de clonagem e conjugação bioquímica que são bem conhecidos pelos técnicos no assunto. As moléculas de fusão compreendem um domínio de ligação ao DNA (por exemplo, ZFP, TALE, sgRNA)

associado com um domínio funcional (por exemplo, um domínio de ativação ou repressão transcrecional). As moléculas de fusão também compreendem opcionalmente sinais de localização nuclear (como, por exemplo, aquele do antígeno T médio de SV40) e etiquetas de epítopo (como, por exemplo, FLAG e hemaglutinina). Proteínas de fusão (e ácidos nucleicos que as codificam) são desenhadas de tal modo que o quadro de leitura da tradução é conservado entre os componentes da fusão.

[00127] As fusões entre um componente polipeptídico de um domínio funcional (ou um fragmento funcional do mesmo), por um lado, e um domínio não proteico de ligação ao DNA (por exemplo, antibiótico, intercalador, ligante de sulco menor, ácido nucleico), pelo outro, são construídos por métodos conjugação bioquímica conhecidos pelos técnicos no assunto. Ver, por exemplo, o catálogo da Pierce Chemical Company (Rockford, IL). Foram descritos métodos e composições para criar fusões entre um ligante de sulco menor e um polipeptídeo. Mapp, *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97:3930-3935. Além do mais, RNAs guias únicos do sistema CRISPR/Cas associam-se com domínios funcionais para formar reguladores transpcionais ativos e nucleases.

[00128] Em certas modalidades, o sítio alvo está presente em uma região acessível de cromatina celular. Regiões acessíveis podem ser determinadas como descrito, por exemplo, nas Patentes U.S. N°s 7 217 509 e 7 923 542. Se o sítio alvo não estiver presente em uma região acessível de cromatina celular, uma ou mais regiões acessíveis podem ser geradas como descrito nas Patentes U.S. N°s 7 785 792 e 8 071 370. Em modalidades adicionais, o domínio de ligação ao DNA de uma molécula de fusão é capaz de ligar-se à cromatina celular independentemente de seu sítio alvo estar ou não em uma região acessível. Por exemplo, tais domínios de ligação ao DNA são capazes

de se ligarem a um DNA de ligação e/ou DNA de nucleossomo. Exemplos desse tipo de domínio de ligação ao DNA "pioneiro" são encontrados em certos receptores esteroides e no fator nuclear 3 de hepatócitos (HNF3) (Cordingley, et al. (1987) *Cell* 48:261-270; Pina, et al. (1990) *Cell* 60:719-731; e Cirillo, et al. (1998) *EMBO J.* 17:244-254).

[00129] A molécula de fusão pode ser formulada com um veículo farmaceuticamente aceitável, como é conhecido por técnicos no assunto. Ver, por exemplo, *Remington's Pharmaceutical Sciences*, 17a ed., 1985; e Patentes U.S. N°s 6 453 242 e 6 534 261.

[00130] O componente/domínio funcional de uma molécula de fusão pode ser selecionado a partir de qualquer uma de uma variedade de componentes diferentes capazes de influenciar a transcrição de um gene uma vez a molécula de fusão ligue-se a uma sequência alvo via seu domínio de ligação ao DNA. Consequentemente, o componente funcional pode incluir, entre outros, vários domínios de fatores de transcrição, como ativadores, repressores, co-ativadores, co-repressores e silenciadores.

[00131] Domínios funcionais exemplares adicionais são descritos, por exemplo, nas Patentes U.S. N°s 6 534 261 e 6 933 113.

[00132] Domínios funcionais que são regulados por pequenas moléculas ou ligantes exógenos podem também ser selecionados. Por exemplo, a tecnologia RheoSwitch® pode ser empregada, em que um domínio funcional somente assume sua conformação ativa na presença do ligante externo RheoChem™ (ver, por exemplo, Publicação de Patente U.S. N° 2009/0136465). Assim, a ZFP pode ser operacionalmente ligada ao domínio funcional regulável, em que a atividade resultante do ZFP-TF é controlada pelo ligante externo.

Nucleases

[00133] Em certas modalidades, a molécula de fusão compreende um domínio de ligação ao DNA associado com um domínio de

clivagem (nuclease). Assim, a modificação do gene pode ser realizada utilizando uma nuclease, por exemplo, uma nuclease projetada. A tecnologia da engenharia de nucleases baseia-se na engenharia de proteínas de ligação ao DNA de ocorrência natural. Por exemplo, foi descrita a engenharia de endonucleases *homing* com especificidades de ligação ao DNA sob medida. Chames, et al. (2005) *Nucleic Acids Res* 33(20):e178; Arnould, et al. (2006) *J. Mol. Biol.* 355:443-458. Além disso, a engenharia de ZFPs foi igualmente descrita. Ver, por exemplo, as Patentes U.S. N°s 6 534 261; 6 607 882; 6 824 978; 6 979 539; 6 933 113; 7 163 824; e 7 013 219.

[00134] Além disso, ZFPs e/ou TALEs podem ser fundidas a domínios nuclease para criar ZFNs e TALENs – uma entidade funcional que é capaz de reconhecer seu ácido nucleico alvo pretendido através de seu domínio de ligação ao DNA projetado (ZFP ou TALE) e fazer com que o DNA seja cortado próximo do DNA sítio de ligação ao DNA através da atividade de nuclease.

[00135] Assim, os métodos e composições aqui descritos são amplamente aplicáveis e podem envolver qualquer nuclease de interesse. Exemplos não limitantes de nucleases incluem meganucleases, TALENs e nucleases de dedo de zinco. A nuclease pode compreender domínios heterólogos de ligação ao DNA e de clivagem (por exemplo, nucleases de dedo de zinco; domínios de ligação ao DNA de meganucleases com domínios heterólogos de clivagem) ou, alternativamente, o domínio de ligação ao DNA de uma nuclease natural pode ser alterado para ligar-se a um sítio alvo selecionado (por exemplo, uma meganuclease que tenha sido projetada para se ligar a um sítio diferente do sítio de ligação cognato).

[00136] Em qualquer uma das nucleases aqui descritas, a nuclease pode compreender um domínio de ligação ao DNA de TALE projetado e um domínio nuclease (por exemplo, domínio endonuclease e/ou

meganuclease), também referidas como TALENs. Métodos e composições para engenharia dessas proteínas TALEN para interação robusta, sítio-específica com a sequência alvo escolhida pelo usuário foram publicados (ver Patente U.S. Nº 8,586,526). Em algumas modalidades, a TALEN compreende um domínio de clivagem ou meio-domínio de clivagem de endonuclease (por exemplo, FokI). Em outras modalidades, a TALE-nuclease é uma mega TAL. Essas mega TAL nucleases são proteínas de fusão compreendendo um domínio de ligação ao DNA de TALE e um domínio de clivagem de meganuclease. O domínio de clivagem de meganuclease é ativo como monômero e não requer dimerização para atividade (ver Boissel, et al. (2013) *Nucl Acid Res*:1-13, doi: 10.1093/nar/gkt1224). Além disso, o domínio nuclease pode também exibir funcionalidade de ligação ao DNA.

[00137] Em ainda outras modalidades, a nuclease compreende uma TALEN compacta (cTALEN). Essas são proteínas de fusão de cadeia única ligando um domínio de ligação ao DNA de TALE a um domínio nuclease TevI. A proteína de fusão pode atuar como uma nickase localizada pela região TALE, ou pode criar uma quebra de fita dupla, dependendo de onde o domínio de ligação ao DNA de TALE está localizado em relação ao domínio nuclease TevI (ver Beurdeley, et al. (2013) *Nat Comm*: 1-8 DOI: 10.1038/ncomms2782). Quaisquer TALENs podem ser utilizadas em combinação com TALENs adicionais (por exemplo, uma ou mais TALENs (cTALENs ou FokI-TALENs) com uma ou mais mega-TALs) ou outras enzimas de clivagem de DNA.

[00138] Em certas modalidades, a nuclease compreende uma meganuclease (endonuclease *homing*) ou uma parte da mesma que exibe atividade de clivagem. Meganucleases naturais reconhecem sítios de clivagem com 14-40 pares de base e são comumente agrupadas em quatro famílias: a família LAGLIDADG ("LAGLIDADG" descrita como SEQ ID Nº:122), a família GIY-YIG, a família His-Cyst

Current Gene Therapy 7:49-66; Publicação de Patente U.S. N°s 2007/0117128; 2006/0206949; 2006/0153826; 2006/0078552; e 2004/0002092). Além disso, domínios de ligação ao DNA naturais ou projetados de meganucleases podem ser operacionalmente ligados com um domínio de clivagem de uma nuclease heteróloga (por exemplo, FokI) e/ou domínios de clivagem de meganucleases podem ser operacionalmente ligados com um domínio de ligação ao DNA heterólogo (por exemplo, ZFP ou TALE).

[00140] Em outras modalidades, a nuclease é uma nuclease de dedo de zinco (ZFN) ou fusão domínio de ligação ao DNA de TALE-nuclease (TALEN). ZFNs e TALENs compreendem um domínio de ligação ao DNA (proteína dedo de zinco ou domínio de ligação ao DNA de TALE) que foi projetado para ligar-se a um sítio alvo em um gene de escolha e um domínio de clivagem ou um meio-domínio de clivagem (por exemplo, de uma nuclease restrição e/ou meganuclease conforme aqui descrita).

[00141] Como descrito detalhadamente acima, os domínios de ligação dedo de zinco e os domínios de ligação ao DNA de TALE podem ser projetados para se ligarem a uma sequência de escolha. Ver, por exemplo, 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. Um domínio de ligação dedo de zinco ou uma proteína TALE projetados podem ter uma nova especificidade de ligação, quando comparados a uma proteína natural. Os métodos de engenharia incluem, entre outros, desenho racional e vários tipos de seleção. O desenho racional inclui, por exemplo, o uso de bancos de dados compreendendo sequências de nucleotídeos triplas (ou quádruplas) e dedo de zinco individual ou sequências de aminoácidos

de TALE, nas quais cada sequência de nucleotídeo tripla ou quádrupla é associada com uma ou mais sequências de aminoácidos de dedos de zinco ou unidades repetidas de TALE que se ligam à sequência tripla ou quádrupla em particular. Ver, por exemplo, as Patentes U.S. N°s 6 453 242 e 6 534 261, aqui incorporadas, em sua totalidade, por referência neste pedido de patente. Em certas modalidades, os domínios de ligação ao DNA compreendem ZFPs derivadas de (por exemplo, o componente ZFP) das ZFNs designadas 68957, 72678, 72732, 72748 (B2M) ou 68846 (TCR).

[00142] A seleção de sítios alvo; e métodos para o desenho e construção de proteínas de fusão (e polinucleotídeos codificadores das mesmas) são conhecidos pelos técnicos no assunto e descritos detalhadamente nas Patentes U.S. N°s 7 888 121 e 8 409 861, aqui incorporadas, em sua totalidade, por referência neste pedido de patente.

[00143] Além disso, como revelado nessas e em outras referências, domínios dedo de zinco, TALEs e/ou proteínas dedo de zinco multidedo podem ser unidos utilizando quaisquer sequências ligadoras adequadas, incluindo, por exemplo, ligantes de 5 ou mais aminoácidos de comprimento. Ver, por exemplo, as Patentes U.S. N°s 6 479 626; 6 903 185; e 7 153 949 para sequências ligadoras exemplares com 6 ou mais aminoácidos de comprimento. As proteínas aqui descritas podem incluir qualquer combinação de ligantes adequados entre os dedos de zinco individuais da proteína. Ver, também, Patente U.S. N° 8 772 453.

[00144] Assim, nucleases como ZFNs, TALENs e/ou meganucleases podem compreender qualquer domínio de ligação ao DNA e qualquer domínio nuclease (clivagem) (domínio de clivagem, meio-domínio de clivagem). Tal como mencionado acima, o domínio de clivagem pode ser heterólogo ao domínio de ligação ao DNA, por exemplo, um domínio de ligação ao DNA de dedo de zinco ou TAL-

efetor e um domínio de clivagem de uma nuclease ou um domínio de ligação ao DNA de meganuclease e domínio de clivagem de uma nuclease diferente. Domínios de clivagem heterólogos podem ser obtidos de qualquer endonuclease ou exonuclease. Endonucleases exemplares das quais um domínio de clivagem pode ser derivado incluem, entre outras, endonucleases de restrição e endonucleases *homing*. Ver, por exemplo, Catálogo de 2002-2003, New England Biolabs, Beverly, MA; e Belfort, *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388. São conhecidas enzimas adicionais que clivam DNA (por exemplo, S1 Nuclease; nuclease do feijão mungo; DNase I pancreática; nuclease do micrococo; endonuclease HO de levedura; ver também Linn, *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Uma ou mais dessas enzimas (ou fragmentos funcionais das mesmas) podem ser utilizadas como fonte de domínios de clivagem e meios-domínios de clivagem).

[00145] Do mesmo modo, um meio-domínio de clivagem pode ser derivado de qualquer nuclease ou porção da mesma, como colocado acima, que requeria dimerização para atividade de clivagem. Em geral, são necessárias duas proteínas de fusão para clivagem se as proteínas de fusão compreenderem meios-domínios de clivagem. Alternativamente, uma única proteína compreendendo dois meios-domínios de clivagem pode ser utilizada. Os dois meios-domínios de clivagem podem ser derivados da mesma endonuclease (ou fragmentos funcionais da mesma), ou cada meio-domínio de clivagem pode ser derivado de uma endonuclease diferente (ou fragmentos funcionais das mesmas). Além disso, os sítios alvo para as duas proteínas de fusão estão de preferência dispostos, em relação um ao outro, de tal modo que a ligação das duas proteínas de fusão aos seus respectivos sítios alvo coloca os meios-domínios de clivagem em uma orientação espacial entre si que permite aos meios-domínios de clivagem formarem um

domínio de clivagem funcional, por exemplo, por dimerização. Assim, em certas modalidades, as bordas próximas dos sítios alvo são separadas por 5-8 nucleotídeos ou por 15-18 nucleotídeos. Contudo, qualquer número inteiro de nucleotídeos ou pares de nucleotídeos pode intervir entre os dois sítios alvo (por exemplo, de 2 a 50 pares de nucleotídeos ou mais). Em geral, o sítio de clivagem situa-se entre os sítios alvo, mas pode estar 1 ou mais quilobases afastado do sítio de clivagem, incluindo entre 1-50 pares de base (ou qualquer valor intermediário incluindo 1-5, 1-10 e 1-20 pares de base), 1-100 pares de base (ou qualquer valor intermediário), 100-500 pares de base (ou qualquer valor intermediário), 500 a 1000 pares de base (ou qualquer valor intermediário) ou mesmo mais de 1 kb do sítio de clivagem.

[00146] As endonucleases de restrição (enzimas de restrição) estão presentes em muitas espécies e são capazes de ligação específica de sequência a DNA (em um sítio de reconhecimento) e clivagem de DNA em ou próximo ao sítio de ligação. Certas enzimas de restrição (por exemplo, Tipo IIS) clivam DNA em sítios removidos do sítio de reconhecimento, e têm domínios separáveis de ligação e clivagem. Por exemplo, a enzima FokI Tipo IIS catalisa a clivagem de fita dupla de DNA, em 9 nucleotídeos de seu sítio de reconhecimento em uma fita e 13 nucleotídeos de seu sítio de reconhecimento na outra. Ver, por exemplo, as Patentes U.S. N°s 5 356 802; 5 436 150 e 5 487 994; bem como Li, *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:4275-4279; Li, *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:2764-2768; Kim, *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* 91:883-887; Kim, *et al.* (1994b) *J. Biol. Chem.* 269:31,978-31,982. Assim, em uma modalidade, as proteínas de fusão compreendem o domínio de clivagem (ou meio-domínio de clivagem) de pelo menos uma enzima de restrição Tipo IIS e um ou mais domínios de ligação dedo de zinco, os quais podem ser projetados ou não.

[00147] Uma enzima de restrição Tipo IIS exemplar, cujo domínio de clivagem é separável do domínio de ligação, é FokI. Essa enzima específica é ativa como dímero. Bitinaite, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:10,570-10,575. A sequência da FokI completa é mostrada abaixo. O domínio de clivagem utilizado nas nucleases aqui descritas é mostrado em itálico e sublinhado (posições 384 a 579 da proteína completa), onde a sequência da proteína holo é descrita abaixo (SEQ ID N°:138):

[00148] MVSKIRTFGVQNPQKFENLKRVVQVFDRNSKVHNEVK
NIKIPTLVKESKIQKELVAIMNQHDLIYTYKELVGTGTSIRSEAPCDAIQ
ATIADQGNKKGYIDNWSSDGFLRWAHALGFIEYINKSDSFVITDVGLA
YSKSADGSAIEKEILIEAISSYPPAIRILTLEDGQHLTKFDLGKNLGFS
GESGFTSLPEGILLDTLANAMPDKGEIRNNWEGSSDKYARMIGGWL
DKLGLVKQGKKEFIIPTLGKPDNKEFISHAFKITGEGLKVLRRAKGSTK
FTRVPKRVYWEMLATNLTDKEYVRTRRALILEILIKAGSLKIEQIQDNL
KKLGFDEVIETIENDIKGLINTGIFIEIKGRFYQLKDHLQFVIPNRGVTK
QLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFF
MKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPI
GQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFK
GNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFN
NGEINF (SEQ ID N°:138)

[00149] Consequentemente, para fins da presente invenção, a porção da enzima FokI utilizada nas proteínas de fusão reveladas é considerada um meio-domínio de clivagem. Assim, para clivagem direcionada de dupla fita e/ou substituição direcionada de sequências celulares usando fusões dedo de zinco-FokI, duas proteínas de fusão, cada uma compreendendo um meio-domínio de clivagem FokI, podem ser utilizadas para reconstituir um domínio de clivagem cataliticamente ativo. Alternativamente, uma única molécula de polipeptídeo contendo um domínio de ligação dedo de zinco e dois meios-domínios de

clivagem FokI pode também ser utilizada. Parâmetros para clivagem direcionada e alteração direcionada da sequência, usando fusões dedo de zinco-FokI, são fornecidos em outro lugar neste relatório descritivo.

[00150] Um domínio de clivagem ou meio-domínio de clivagem pode ser qualquer porção de uma proteína que retenha atividade de clivagem, ou que retenha a capacidade para multimerizar (por exemplo, dimerizar) e formar um domínio de clivagem funcional.

[00151] Enzimas de restrição Tipo IIS exemplares são descritas na Publicação de Patente Internacional Nº WO 07/014275, aqui incorporada em sua totalidade. Enzimas de restrição adicionais também contêm domínios separáveis de ligação e clivagem, e estas são contempladas pela presente invenção. Ver, por exemplo, Roberts, *et al.* (2003) *Nucleic Acids Res.* 31:418-420.

[00152] Em certas modalidades, o domínio de clivagem compreende um ou mais meios-domínios de clivagem projetados (também referidos como mutantes de domínio de dimerização) que minimizam ou impedem a homodimerização, conforme descrito, por exemplo, nas Patentes U.S. Nºs 7 914 796; 8 034 598; e 8 623 618; e Publicação de Patente U.S. Nº 2011/0201055, cujos conteúdos são aqui incorporados, em sua totalidade, por referência neste pedido de patente. Mutações de "Sharkey" (por exemplo, 418 e 441, numerados em relação ao comprimento total) e mutações adicionais, por exemplo, no resíduo 416 (por exemplo, R416S) e/ou resíduo 525 (por exemplo, K525S), conforme descritas na Publicação de Patente U.S. Nº 2018/0087072, podem igualmente ser incluídas. Assim, os domínios de clivagem FokI usados nas nucleases da invenção podem sofrer mutação em uma ou mais das seguintes posições de resíduos de aminoácido (numerados em relação ao comprimento total): 416, 418, 441, 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 525, 531, 534, 537 e/ou 538.

[00153] Os meios-domínios de clivagem projetados exemplares de FokI que foram heterodímeros obrigados incluem um par no qual um primeiro meio-domínio de clivagem inclui mutações em resíduos de aminoácido nas posições 490 e 538 de FokI e um segundo meio-domínio de clivagem inclui mutações nos resíduos de aminoácido 486 e 499.

[00154] Desse modo, em uma modalidade, uma mutação em 490 substitui Glu (E) por Lys (K); a mutação em 538 substitui Iso (I) por Lys (K); a mutação em 486 substitui Gln (Q) por Glu (E); e a mutação na posição 499 substitui Iso (I) por Lys (K). Especificamente, os meios-domínios de clivagem projetados aqui descritos foram preparados por mutação nas posições 490 (E→K) e 538 (I→K) em um meio-domínio de clivagem para produzir um meio-domínio de clivagem projetado designado "E490K:I538K", e por mutação nas posições 486 (Q→E) e 499 (I→L) em outro meio-domínio de clivagem para produzir um meio-domínio de clivagem projetado designado "Q486E:I499L". Os meios-domínios de clivagem projetados aqui descritos são mutantes de heterodímeros obrigados nos quais a clivagem anormal é minimizada ou abolida. Ver, por exemplo, as Patentes U.S. N°s 7 914 796 e 8 034 598, cujos conteúdos são aqui incorporados, em sua totalidade, por referência para todos os efeitos. Em certas modalidades, o meio-domínio de clivagem projetado compreende mutações nas posições 486, 499 e 496 (numeradas em relação a FokI do tipo selvagem), por exemplo, mutações que substituem o resíduo de Gln (Q) do tipo selvagem na posição 486 por um resíduo de Glu (E), o resíduo de Iso (I) do tipo selvagem na posição 499 por um resíduo de Leu (L) e o resíduo de Asn (N) do tipo selvagem na posição 496 por um resíduo de Asp (D) ou Glu (E) (também referidos como domínios "ELD" e "ELE", respectivamente). Em outras modalidades, o meio-domínio de clivagem projetado compreende mutações nas posições 490, 538 e

537 (numeradas em relação a FokI do tipo selvagem), por exemplo, mutações que substituem o resíduo de Glu (E) do tipo selvagem na posição 490 por um resíduo de Lys (K), o resíduo de Iso (I) do tipo selvagem na posição 538 por um resíduo de Lys (K) e o resíduo de His (H) do tipo selvagem na posição 537 por um resíduo de Lys (K) resíduo ou um resíduo de Arg (R) (também referidos como domínios "KKK" e "KKR", respectivamente). Em outras modalidades, o meio-domínio de clivagem projetado compreende mutações nas posições 490 e 537 (numeradas em relação a FokI do tipo selvagem), por exemplo, mutações que substituem o resíduo de Glu (E) do tipo selvagem na posição 490 por um resíduo de Lys (K) e o resíduo de His (H) do tipo selvagem na posição 537 por um resíduo de Lys (K) ou um resíduo de Arg (R) (também referidos como domínios "KIK" e "KIR", respectivamente).

[00155] Em outras modalidades, o meio-domínio de clivagem projetado compreende mutações nas posições 487, 499 e 496 (numeradas em relação a FokI do tipo selvagem), por exemplo, mutações que substituem o resíduo de Arg (R) do tipo selvagem na posição 487 por um resíduo de Asp (D) e o resíduo de Ile (I) do tipo selvagem na posição 499 por um resíduo de Ala (A) e o resíduo de Asn (N) do tipo selvagem na posição 496 por um resíduo de Asp (D) (também referido como "DAD") e/ou mutações nas posições 483, 538 e 537 (numeradas em relação a FokI do tipo selvagem), por exemplo, mutações que substituem o resíduo de Asp (D) do tipo selvagem na posição 483 por um resíduo de Arg (R) e o resíduo de Ile (I) do tipo selvagem na posição 538 por um resíduo de Val (V) e o resíduo de His (H) do tipo selvagem na posição 537 por um resíduo de Arg (R) resíduo (também referido como "RVR"). Ver, por exemplo, as Patentes U.S. N^{os} 8 962 281; 7 914 796; 8 034 598; e 8 623 618, cujos conteúdos são aqui incorporados, em sua totalidade, por referência

para todos os efeitos. Em outras modalidades, o meio-domínio de clivagem projetado compreende a mutação "Sharkey" e/ou mutações "Sharkey" (ver Guo, *et al.* (2010) *J. Mol. Biol.* 400(1):96-107).

[00156] Assim, exemplos não limitantes de domínios FokI que podem ser utilizados nas nucleases aqui descritas incluem: mutantes de Fok mostrados na Tabela 8 (por exemplo, ELD, KKR, etc.), FokI-Sharkey (S418P+K441E), FokI ELD (Q->E na posição 486, I->L em 499, N->D na posição 496), FokI ELD, Sharkey (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E), FokI ELD, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, R416E), FokI ELD, Sharkey, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, R416E), FokI ELD, R416Y (Q->E na posição 486, I->L na posição 499, N->D na posição 496, R416Y), FokI ELD, Sharkey, R416E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, R416E), FokI ELD, S418E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418E), FokI ELD, Sharkey parcial, S418E (Q->E na posição 486, I->L na posição 499, N->D na posição 496, K441E, S418E), FokI ELD, K525S (Q->E na posição 486, I->L na posição 499, N->D na posição 496, K525S), FokI ELD, Sharkey K525S (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, K525S), FokI ELD, I479T (Q->E na posição 486, I->L na posição 499, N->D na posição 496, I479T), FokI ELD, Sharkey, I479T (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, I479T), FokI ELD, P478D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, P478D), FokI ELD, Sharkey, P478D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, S418P+K441E, P478D), FokI ELD, Q481D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, Q481D), FokI ELD, Sharkey, Q481D (Q->E na posição 486, I->L na posição 499, N->D na posição 496, Q481D).

>D na posição 496, S418P+K441E, Q481D), FokI KKR (E->K na posição 490, I->K na posição 538, H->R na posição 537), FokI KKR Sharkey, (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E), FokI KKR, Q481E (E->K na posição 490, I->K na posição 538, H->R na posição 537, Q481E), FokI KKR, Sharkey Q481E (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, Q481E), FokI KKR, R416E (E->K na posição 490, I->K na posição 538, H->R na posição 537, R416E), FokI KKR, Sharkey, R416E (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, R416E), FokI KKR, K525S (E->K na posição 490, I->K na posição 538, H->R na posição 537, K525S), FokI KKR, Sharkey, K525S (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, K525S), FokI KKR, R416Y (E->K na posição 490, I->K na posição 538, H->R na posição 537, R416Y), FokI KKR, Sharkey, R416Y (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, R416Y), FokI, KKR I479T (E->K na posição 490, I->K na posição 538, H->R na posição 537, I479T), FokI, KKR Sharkey I479T (E->K na posição 490, I->K na posição 538, H->R na posição 537, S418P+K441E, I479T, FokI, KKR P478D(E->K na posição 490, I->K nas posições 538, H->R na posição 537, P478D), FokI, KKR Sharkey P478D(E->K na posição 490, I->K na posição 538, H->R na posição 537, P478D), FokI DAD (R->D na posição 487, N->D na posição 496, I->A na posição 499), FokI DAD Sharkey (R->D na posição 487, N->D na posição 496, I->A na posição 499, S418P+K441E), FokI RVR (D->R na posição 483, H->R na posição 537, I->V na posição 538), FokI RVR Sharkey (D->R na posição 483, H->R na posição 537, I->V na posição 538, S418P+K441E).

[00157] As ZFNs aqui descritas podem também incluir qualquer sequência ligadora, incluindo, entre outras, as sequências aqui reveladas (L0, N7a, N7c, etc.) e/ou aquelas reveladas na Patente U.S.

Nº 7 888 121; 7 914 796; 8 034 598; 8 623 618; 9 567 609; e Publicação U.S. Nº 20170218349, que podem ser utilizadas entre o N- ou C-terminal do domínio de ligação ao DNA e N- ou C-terminal do domínio de clivagem FokI.

[00158] ZFPs das ZFNs conforme aqui descritas (incluindo domínios de clivagem projetados e/ou do tipo selvagem) podem também incluir modificações para aumentar a especificidade de uma ZFN, incluindo um par de nucleases, por seu alvo pretendido em relação a outros sítios de clivagem não pretendidos, conhecidos como sítios fora do alvo (ver Publicação de Patente U.S. Nº 20180087072). Assim, as nucleases aqui descritas podem compreender ligantes específicos entre o domínio de ligação ao DNA e o domínio de clivagem; e/ou podem compreender mutações em uma ou mais das suas regiões no esqueleto (*backbone*) de domínio de ligação ao DNA e/ou uma ou mais mutações em seus domínios nuclease de clivagem como descrito acima. As ZFPs dessas nucleases podem incluir mutações em aminoácidos dentro do domínio de ligação ao DNA da ZFP ("esqueleto da ZFP") que podem interagir não especificamente com fosfatos no esqueleto do DNA, porém não compreendem alterações nas hélices de reconhecimento do DNA. Consequentemente, a invenção inclui ZFPs compreendendo mutações de resíduos de aminoácidos catiônicos no esqueleto de ZFP que não são necessários para a especificidade do nucleotídeo pelo alvo. Em algumas modalidades, essas mutações no esqueleto de ZFP compreendem uma mutação em um resíduo de aminoácido catiônico para um resíduo de aminoácido neutro ou aniônico. Em algumas modalidades, essas mutações no esqueleto de ZFP compreendem uma mutação em um resíduo de aminoácido polar para um resíduo de aminoácido neutro ou não polar. Em modalidades preferidas, as mutações são feitas na posição (-5), (-9) e/ou posição (-14) em relação

à hélice de ligação ao DNA. Em algumas modalidades, um dedo de zinco pode compreender uma ou mais mutações em (-5), (-9) e/ou (-14). Em modalidades adicionais, um ou mais dedos de zinco em uma proteína dedo de zinco multi-dedo pode compreender mutações em (-5), (-9) e/ou (-14). Em algumas modalidades, os aminoácidos em (-5), (-9) e/ou (-14) (por exemplo, uma arginina (R) ou lisina (K)) sofrem mutação para alanina (A), leucina (L), Ser (S), Asp (N), Glu (E), Tyr (Y) e/ou glutamina (Q).

[00159] Em certas modalidades, as ZFNs compreendem pelo menos um dos pares seguintes: 68796 e 68813; 68796 e 68861; 68812 e 68813; 68876 e 68877; 68815 e 55266; 68879 e 55266; 68798 e 68815; ou 68846 e 53853 como mostrado na Tabela 6. Em outras modalidades, as ZFNs compreendem pelo menos um dos pares seguintes: 57531 e 72732; 57531 e 72748; 68957 e 57071; 68957 e 72732; 68957 e 72748; 72678 e 57071; 72678 e 72732; ou 72678 e 72748 como mostrado na Tabela 8.

[00160] Alternativamente, as nucleases podem ser montadas *in vivo* no sítio alvo do ácido nucleico utilizando a chamada tecnologia de "split-enzyme" (ver, por exemplo, Publicação de Patente U.S. Nº 2009/0068164). Os componentes de tais enzimas divididas podem ser expressos em construções de expressão separadas ou podem ser ligados em um quadro de leitura aberta onde os componentes individuais são separados, por exemplo, por um peptídeo de autoclivagem 2A ou sequência IRES. Os componentes podem ser domínios de ligação dedo de zinco individuais ou domínios de um domínio de ligação a ácido nucleico de meganuclease.

[00161] As nucleases (por exemplo, ZFNs e/ou TALENs) podem ser rastreadas quanto à atividade antes do uso, por exemplo, em um sistema cromossômico baseado em levedura como a descrito na Patente U.S. Nº 8 563 314.

[00162] Em certas modalidades, a nuclease compreende um sistema CRISPR/Cas. O lócus de CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*), que codifica componentes de RNA do sistema, e o lócus de Cas (CRISPR-associado), que codifica proteínas (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) compõem as sequências gênicas do sistema de nuclease CRISPR/Cas. Os *loci* CRISPR em hospedeiros microbianos contêm uma combinação de genes CRISPR-associados (Cas) bem como elementos não codificantes de RNA capazes de programar a especificidade da clivagem de ácido nucleico mediada por CRISPR.

[00163] O CRISPR Tipo II é um dos sistemas mais bem caracterizados e realiza a quebra direcionada da fita dupla de DNA em quatro etapas sequenciais. Em primeiro lugar, dois RNA não codificantes, o arranjo de pré-crRNA e tracrRNA, são transcritos a partir do lócus de CRISPR. Em segundo lugar, tracrRNA hibridiza as regiões de repetição do pré-crRNA e medeia o processamento de pré-crRNA em crRNAs maduros contendo sequências espaçadoras individuais. Em terceiro lugar, o complexo crRNA maduro:tracrRNA direciona um domínio funcional (por exemplo, nuclease como Cas) ao DNA alvo através do pareamento de bases de Watson-Crick entre o espaçador no crRNA e o protoespaçador no DNA alvo ao lado do motivo protoespaçador adjacente (PAM), um requisito adicional para reconhecimento do alvo. Finalmente, a Cas9 medeia a clivagem do DNA alvo para criar uma quebra na fita dupla dentro do protoespaçador. A atividade do sistema CRISPR/Cas compreende três etapas: (i) inserção de sequências de DNA estranhas no arranjo de CRISPR para prevenir ataques futuros, em um processo denominado "adaptação", (ii) expressão das proteínas relevantes, bem como

expressão e processamento do arranjo, seguida por (iii) interferência mediada por RNA com o ácido nucleico estranho. Dessa forma, na célula bacteriana, diversas das chamadas proteínas "Cas" estão envolvidas com a função natural do sistema CRISPR/Cas e desempenham papéis em funções como inserção do DNA estranho etc.

[00164] Em certas modalidades, a proteína Cas pode ser um "derivado funcional" de uma proteína Cas natural. Um "derivado funcional" de um polipeptídeo de sequência nativa é um composto com uma propriedade biológica qualitativa em comum com um polipeptídeo de sequência nativa. Os "derivados funcionais" incluem, entre outros, fragmentos de uma sequência nativa e derivados de um polipeptídeo de sequência nativa e seus fragmentos, desde que estes tenham uma atividade biológica em comum com um polipeptídeo de sequência nativa correspondente. A atividade biológica aqui contemplada é a capacidade do derivado funcional de hidrolisar um substrato de DNA em fragmentos. O termo "derivado" abrange tanto variantes de sequência de aminoácidos de polipeptídeo, modificações covalentes e suas fusões. Os derivados adequados de um polipeptídeo Cas ou seu fragmento incluem, entre outros, mutantes, fusões, modificações covalentes de proteína Cas ou seu fragmento. A proteína Cas, que inclui a proteína Cas ou seu fragmento, bem como derivados da proteína Cas ou seu fragmento, pode ser obtida de uma célula ou sintetizada quimicamente ou por uma combinação desses dois procedimentos. A célula pode ser uma célula que produz naturalmente proteína Cas, ou uma célula que produz naturalmente proteína Cas e é geneticamente projetada para produzir a proteína Cas endógena em um nível de expressão mais alto ou para produzir uma proteína Cas a partir de um ácido nucleico introduzido exogenamente, ácido nucleico o qual que codifica uma Cas que é igual ou diferente da Cas

endógena. Em alguns casos, a célula não produz naturalmente proteína Cas e é geneticamente projetada para produzir uma proteína Cas.

[00165] Sistemas nuclease CRISPR/Cas exemplares, direcionados para genes TCR e outros genes, são revelados, por exemplo, na Publicação de Patente U.S. Nº 2015/0056705. A(s) nuclease(s) pode(m) fazer um ou mais cortes de fita dupla e/ou fita simples no sítio alvo. Em certas modalidades, a nuclease compreende um domínio de clivagem cataliticamente inativo (por exemplo, FokI e/ou proteína Cas). Ver, por exemplo, as Patentes U.S. Nós 9 200 266 e 8 703 489 e Guillinger, *et al.* (2014) *Nature Biotech.* 32(6):577-582. O domínio de clivagem cataliticamente inativo pode, em combinação com um domínio cataliticamente ativo, atuar como uma nickase para fazer um corte de fita simples. Portanto, duas nickases podem ser utilizadas em combinação para fazer um corte de fita dupla em uma região específica. Nickases adicionais são igualmente conhecidas na técnica, por exemplo, McCaffrey, *et al.* (2016) *Nucleic Acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Epub 2015 Oct 19. Além disso, Cas dead ("dCas") ou uma nickase Cas pode ser fundida a uma enzima modificadora de base (por exemplo, citidina desaminase) para criar um sistema de edição de base (Komor, *et al.* (2016) *Nature* 533:420). Esses sistemas permitem alterar uma base do DNA (modificação) pelo complexo editor de base sem criar uma quebra de fita dupla no DNA. Assim, em algumas modalidades, RNAs guias (Tabela 2) podem ser utilizados para introduzir mutações em um gene TRAC e provocar um knockout.

Entrega

[00166] As proteínas (por exemplo, fatores de transcrição, nucleases, moléculas de TCR e CAR), polinucleotídeos e/ou composições compreendendo as proteínas e/ou os polinucleotídeos

aqui descritos podem ser entregues a uma célula alvo por qualquer meio adequado, incluindo, por exemplo, por injeção da proteína e/ou de componentes de mRNA. Em algumas modalidades, as proteínas são introduzidas na célula por método de *Cell Squeeze* (ver Kollmannsperger, et al. (2016) *Nat Comm* 7, 10372 doi:10.1038/ncomms10372).

[00167] As células adequadas incluem, entre outras, células e/ou linhagens celulares eucarióticas e procarióticas. Exemplos não limitantes de tais células ou linhagens celulares geradas a partir de tais células incluem células T, COS, CHO (por exemplo, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (por exemplo, HEK293-F, HEK293-H, HEK293-T) e células perC6, bem como células de insetos, como *Spodoptera frugiperda* (Sf), ou células de fungos como *Saccharomyces*, *Pichia* e *Schizosaccharomyces*. Em certas modalidades, a linhagem celular é uma linhagem de células CHO-K1, MDCK ou HEK293. As células adequadas também incluem células-tronco como, a título de exemplo, células-tronco embrionárias, células-tronco pluripotentes induzidas (células iPS), células-tronco hematopoiéticas, células-tronco neuronais e células-tronco mesenquimais.

[00168] Métodos para entrega de proteínas compreendendo domínios de ligação ao DNA, conforme aqui descritos, são descritos, por exemplo, nas Patentes U.S. N°s 6 453 242; 6 503 717; 6 534 261; 6 599 692; 6 607 882; 6 689 558; 6 824 978; 6 933 113; 6 979 539; 7 013 219; e 7 163 824, cujos conteúdos são aqui incorporados, em sua totalidade, por referência neste pedido de patente.

[00169] Os domínios de ligação ao DNA e as proteínas de fusão compreendendo esses domínios de ligação ao DNA, conforme aqui descritos, podem igualmente ser entregues utilizando vetores

contendo sequências codificadoras de uma ou mais da(s) proteína(s) de ligação ao DNA. Adicionalmente, outros ácidos nucleicos (por exemplo, doadores) também podem ser entregues através desses vetores. Quaisquer sistemas de vetores podem ser utilizados, incluindo, entre outros, vetores de plasmídeo, vetores retrovirais, vetores lentivirais, vetores adenovirais, vetores de poxvírus; vetores de herpesvírus e vetores de vírus adenoassociados, etc. Ver, também, as Patentes U.S. N^{os} 6 534 261; 6 607 882; 6 824 978; 6 933 113; 6 979 539; 7 013 219; e 7 163 824, aqui incorporadas, em sua totalidade, por referência neste pedido de patente. Além do mais, será evidente que qualquer um desses vetores pode compreender uma ou mais sequências codificadoras de proteínas de ligação ao DNA e/ou ácidos nucleicos adicionais conforme adequado. Desse modo, quando introduzidas na célula, uma ou mais proteínas de ligação ao DNA, conforme aqui descritas, e DNAs adicionais, se adequado, podem ser carregadas no mesmo vetor ou em vetores diferentes. Quando múltiplos vetores são utilizados, cada vetor pode compreender uma sequência codificadora de uma ou de múltiplas proteínas de ligação ao DNA e ácidos nucleicos adicionais, se desejado.

[00170] Métodos convencionais de transferência gênica, baseados em vírus e não virais podem ser utilizados para introduzir ácidos nucleicos codificadores de proteínas projetadas de ligação ao DNA em células (por exemplo, células de mamíferos) e tecidos alvos e para introduzir simultaneamente sequências nucleotídicas adicionais, se desejado. Tais métodos podem também ser utilizados para administrar ácidos nucleicos (por exemplo, que codificam proteínas de ligação ao DNA e/ou doadores) a células *in vitro*. Em certas modalidades, os ácidos nucleicos são administrados para usos *in vivo* ou *ex vivo* de terapia gênica. Sistemas de entrega por vetor não viral incluem plasmídeos de DNA, ácido nucleico nu e ácido nucleico complexado

com um veículo de entrega como um lipossomo, nanopartícula lipídica ou poloxâmero. Os sistemas de entrega por vetor viral incluem vírus DNA e RNA, que tenham genomas episomais ou integrados depois da entrega à célula. Para uma revisão de procedimentos de terapia gênica, ver Anderson (1992) *Science* 256:808-813; Nabel & Felgner (1993) *TIBTECH* 11:211-217; Mitani & Caskey (1993) *TIBTECH* 11:162-166; Dillon (1993) *TIBTECH* 11:167-175; Miller (1992) *Nature* 357:455-460; Van Brunt (1988) *Biotechnology* 6(10):1149-1154; Vigne (1995) *Restorative Neurology and Neuroscience* 8:35-36; Kremer & Perricaudet (1995) *British Medical Bulletin* 51(1):31-44; Haddada, et al. (1995) *Current Topics in Microbiology and Immunology* Doerfler e Böhm (eds.); e Yu, et al. (1994) *Gene Therapy* 1:13-26.

[00171] Os métodos de entrega não viral de ácidos nucleicos incluem eletroporação, lipofecção, microinjeção, biobalística, virossomos, lipossomos, nanopartículas lipídicas, imunolipossomos, policáton ou conjugados lipídio: ácido nucleico, DNA nu, mRNA, vírions artificiais e captação de DNA reforçada por agente. A sonoporação utilizando, por exemplo, o sistema Sonitron 2000 (Rich-Mar) pode também ser empregada para entrega de ácidos nucleicos. Em uma modalidade preferida, um ou mais ácidos nucleicos são entregues como mRNA. Além disso, é preferido também o uso de capeamento do mRNA para aumentar a eficiência da tradução e/ou a estabilidade do mRNA. Os caps especialmente preferidos são ARCA (*anti-reverse cap analog*) ou seus variantes. Ver as Patentes U.S. N°s 7 074 596 e 8 153 773, aqui incorporadas por referência.

[00172] Sistemas de entrega de ácido nucleico adicionais exemplares incluem aqueles fornecidos pela Amaxa Biosystems (Colônia, Alemanha), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) e Copernicus Therapeutics Inc, (ver, por exemplo, a Patente U.S. N° 6 008 336). A lipofecção é

descrita em, por exemplo, as Patentes U.S. N°s 5 049 386; 4 946 787; e 4 897 355) e reagentes para lipofecção são vendidos comercialmente (por exemplo, Transfectam™, Lipofectin™ e Lipofectamine™ RNAiMAX). Os lipídios catiônicos e neutros que são adequados para lipofecção eficiente por reconhecimento do receptor de polinucleotídeos incluem aqueles de Felgner, Publicações de Patente Internacional N°s WO 91/17424 e WO 91/16024. A entrega pode ser a células (administração *ex vivo*) ou a tecidos alvo (administração *in vivo*).

[00173] A preparação de complexos lipídio:ácido nucleico, incluindo lipossomos direcionados como complexos imunolipídicos, é bem conhecida pelo técnico no assunto (ver, por exemplo, Crystal (1995) *Science* 270:404-410; Blaese, *et al.* (1995) *Cancer Gene Ther.* 2:291-297; Behr, *et al.* (1994) *Bioconjugate Chem.* 5:382-389; Remy, *et al.* (1994) *Bioconjugate Chem.* 5:647-654; Gao, *et al.* (1995) *Gene Therapy* 2:710-722; Ahmad, *et al.* (1992) *Cancer Res.* 52:4817-4820; Patentes U.S. N°s 4 .186 .183; 4 .217 .344; 4 .235 .871; 4 .261 .975; 4 .485 .054; 4 .501 .728; 4 .774 .085; 4 .837 .028; e 4 .946 .787).

[00174] Métodos adicionais de entrega incluem o uso de empacotamento dos ácidos nucleicos a serem entregues em veículos de entrega EnGeneIC (EDVs). Esses EDVs são entregues especificamente aos tecidos alvo utilizando anticorpos biespecíficos, em que um braço do anticorpo tem especificidade pelo tecido alvo e o outro tem especificidade pelo EDV. O anticorpo leva os EDVs à superfície da célula alvo e, a seguir, o EDV é levado a entrar na célula por endocitose. Uma vez na célula, o conteúdo é liberado (ver MacDiarmid, *et al.* (2009) *Nature Biotechnology* 27(7):643).

[00175] O uso de sistemas baseados em RNA ou DNA viral para a entrega de ácidos nucleicos codificadores de proteínas projetadas de ligação ao DNA e/ou doadores (por exemplo, CARs ou ACTRs), se

desejado, aproveita-se de processos altamente evoluídos de direcionamento de um vírus para células específicas no corpo e o trânsito da carga viral para o núcleo. Os vetores virais podem ser administrados diretamente aos pacientes (*in vivo*) ou podem ser utilizados para tratar células *in vitro* e as células modificadas são administrados aos pacientes (*ex vivo*). Os sistemas convencionais de base viral para a entrega de ácidos nucleicos incluem, entre outros, vetores retrovirais, lentivirais, adenovirais, adenoassociados do vírus *Vaccinia* e *herpes simplex* para a transferência do gene. A integração no genoma do hospedeiro é possível com os métodos de transferência gênica mediadas por retrovírus, lentivírus e vírus adenoassociado, frequentemente resultando em expressão prolongada do transgene inserido. Além disso, foram observados altos níveis de eficiência de transdução em muitos tipos diferentes de células e tecidos alvo.

[00176] O tropismo de um retrovírus pode ser alterado incorporando proteínas estranhas do envelope, expandindo a população alvo potencial de células alvo. Vetores lentivirais são vetores retrovirais que são capazes de transduzir ou infectar células não se dividindo e tipicamente produzem altos títulos virais. A seleção de um sistema de transferência gênica retroviral depende do tecido alvo. Os vetores retrovirais são compostos por repetições terminais longas atuantes em *cis* (do inglês, *cis-acting*) com capacidade de empacotamento de até 6-10 kb de sequência estranha. Um número mínimo de LTRs atuantes em *cis* é suficiente para replicação e empacotamento dos vetores, os quais são então usados para entregar o gene terapêutico na célula alvo para proporcionar expressão permanente do transgene. Vetores retrovirais amplamente utilizados incluem aqueles baseados no vírus da leucemia murina (MuLV), vírus da leucemia do macaco gibão (GaLV), vírus da imunodeficiência símia (SIV), vírus da imunodeficiência humana (HIV) e combinações dos mesmos (ver, por

exemplo, Buchscher, *et al.* (1992) *J. Virol.* 66:2731-2739; Johann, *et al.* (1992) *J. Virol.* 66:1635-1640; Sommerfelt, *et al.* (1990) *Virol.* 176:58-59; Wilson, *et al.* (1989) *J. Virol.* 63:2374-2378; Miller, *et al.* (1991) *J. Virol.* 65:2220-2224; Publicação da Patente Internacional Nº WO 1994/026877).

[00177] Em aplicações em que a expressão transitória é preferida, sistemas baseados em adenovírus podem ser utilizados. Os vetores baseados em adenovírus são capazes de eficiência de transdução muito alta em muitos tipos de células e não requerem divisão celular. Com tais vetores, foram obtidos alto título e altos níveis de expressão. Esse vetor pode ser produzido em grandes quantidades em um sistema relativamente simples. Os vetores de vírus adenoassociados ("AAV") são também utilizados para transduzir células com ácidos nucleicos alvo, por exemplo, na produção *in vitro* de ácidos nucleicos e peptídeos e para procedimentos de terapia gênica *in vivo* e *ex vivo* (ver, por exemplo, West, *et al.* (1987) *Virology* 160:38-47; Patente U.S. Nº 4 797 368; Publicação da Patente Internacional Nº WO 93/24641; Kotin (1994) *Human Gene Therapy* 5:793-801; Muzyczka (1994) *J. Clin. Invest.* 94:1351. A construção de vetores de AAV recombinantes é descrita em várias publicações, incluindo a Patente U.S. Nº 5 173 414; Tratschin, *et al.* (1985) *Mol. Cell. Biol.* 5:3251-3260; Tratschin, *et al.* (1984) *Mol. Cell. Biol.* 4:2072-2081; Hermonat & Muzyczka (1984) *PNAS USA* 81:6466-6470; e Samulski *et al.* (1989) *J. Virol.* 63:03822-3828.

[00178] Há pelo menos seis abordagens com vetores virais atualmente disponíveis para transferência de genes em estudos clínicos, que utilizam abordagens que envolvem complementação de vetores defeituosos por genes inseridos em linhagens celulares auxiliares (*helper*) para gerar o agente transdutor.

[00179] pLASN e MFG-S são exemplos de vetores retrovirais que

foram utilizados em estudos clínicos (Dunbar, *et al.* (1995) *Blood* 85:3048-305; Kohn, *et al.* (1995) *Nat. Med.* 1:1017-102; Malech, *et al.* (1997) *PNAS USA* 94:22 12133-12138). PA317/pLASN foi o primeiro vetor terapêutico usado em um estudo de terapia gênica. (Blaese, *et al.* (1995) *Science* 270:475-480). Eficiências de transdução iguais ou superiores a 50% foram observados para vetores empacotados com MFG-S. (Ellem, *et al.* (1997) *Immunol Immunother.* 44(1):10-20; Dranoff, *et al.* (1997) *Hum. Gene Ther.* 1:111-2.

[00180] Os vetores de vírus adenoassociados recombinantes (rAAV) constituem uma alternativa promissora para sistema de entrega de genes baseados no vírus adenoassociado tipo 2, parvovírus defeituoso e não patogênico. Todos os vetores são derivados de um plasmídeo que retém somente as repetições terminais invertidas com 145 bp do AAV, flanqueando o cassete de expressão do transgene. A transferência eficiente do gene e a entrega estável do transgene que se devem à integração nos genomas da célula transduzida são características fundamentais desse sistema de vetor (Wagner, *et al.* (1998) *Lancet* 351(9117):1702-3, Kearns, *et al.* (1996) *Gene Ther.* 9:748-55). Outros sorotipos de AAV, incluindo AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV8.2, AAV9 e AAVrh10 e AAV pseudotipado como AAV2/8, AAV2/5 e AAV2/6 podem também ser utilizados de acordo com a presente invenção.

[00181] Vetores adenovirais recombinantes com replicação deficiente (Ad) podem ser produzidos com título elevado e infectarem rapidamente vários tipos diferentes de células. A maioria dos vetores adenovirais é projetada de tal modo que um transgene substitui os genes Ad E1a, E1b e/ou E3; subsequentemente, o vetor com defeito de replicação é propagado em células 293 humanas que suprem a função do gene deletado em *trans*. Vetores Ad podem transduzir múltiplos tipos de tecidos *in vivo*, incluindo células diferenciadas que

não se dividem como aquelas encontradas no fígado, rim e no músculo. Os vetores Ad convencionais possuem uma grande capacidade de carga. Um exemplo do uso de um vetor Ad em um estudo clínico envolveu a terapia com polinucleotídeo para imunização antitumoral com injeção intramuscular (Sterman, *et al.* (1998) *Hum. Gene Ther.* 7:1083-9). Exemplos adicionais do uso de vetores adenovirais para transferência de genes em estudos clínicos incluem Rosenecker, *et al.* (1996) *Infection* 24(1):5-10; Sterman, *et al.* (1998) *Hum. Gene Ther.* 9(7):1083-1089; Welsh, *et al.* (1995) *Hum. Gene Ther.* 2:205-18; Alvarez, *et al.* (1997) *Hum. Gene Ther.* 5:597-613; Topf, *et al.* (1998) *Gene Ther.* 5:507-513; Sterman, *et al.* (1998) *Hum. Gene Ther.* 7:1083-1089.

[00182] Células de empacotamento são usadas para formar partículas virais que são capazes de infectar uma célula hospedeira. Tais células incluem células 293, que empacotam adenovírus, e células ψ2 ou células PA317, que empacotam retrovírus. Os vetores virais usados em terapia gênica são normalmente gerados por uma linhagem celular produtora que empacota um vetor de ácido nucleico em uma partícula viral. Os vetores tipicamente contêm as sequências virais mínimas necessárias para o empacotamento e a integração subsequente em um hospedeiro (se aplicável), sendo as outras sequências virais substituídas por um cassete de expressão que codifica a proteína a ser expressa. As funções virais ausentes são supridas em *trans* pela linhagem celular de empacotamento. Por exemplo, os vetores AAV usados em terapia gênica tipicamente possuem apenas sequências de repetições terminais invertidas (ITR) do genoma do AAV, necessárias para o empacotamento e a integração no genoma do hospedeiro. O DNA viral é empacotado em uma linhagem celular, a qual contém um plasmídeo auxiliar que codifica os outros genes do AAV, a saber, rep e cap, mas desprovido

de sequências ITR. A linhagem celular é também infectada com adenovírus como um auxiliar (*helper*). O vírus auxiliar promove a replicação do vetor AAV e a expressão de genes do AAV a partir do plasmídeo auxiliar. O plasmídeo auxiliar não é empacotado em quantidades suficientes devido a uma falta de sequências ITR. A contaminação com adenovírus pode ser reduzida, por exemplo, por tratamento com calor ao qual o adenovírus é mais sensível que o AAV. Além disso, AAV pode ser produzido utilizando um sistema de baculovírus (ver, por exemplo, as Patentes U.S. N°s 6 723 551 e 7 271 002).

[00183] A purificação de partículas de AAV de um sistema de células 293 ou baculovírus tipicamente envolve o crescimento das células que produzem o vírus, seguido pela coleta das partículas virais no sobrenadante celular ou pela lise das células e coleta do vírus no lisado bruto. AAV é então purificado por métodos conhecidos na técnica, incluindo cromatografia de troca iônica (por exemplo, ver as Patentes U.S. N°s 7 419 817 e 6 989 264), cromatografia de troca iônica e centrifugação por gradiente de densidade com CsCl (por exemplo, Publicação da Patente Internacional N° WO 2011/094198 A10), cromatografia de imunoafinidade (por exemplo, Publicação da Patente Internacional N° WO 2016/128408) ou purificação utilizando AVB Sepharose (por exemplo, GE Healthcare Life Sciences).

[00184] Em muitas aplicações de terapia gênica, é desejável que o vetor da terapia gênica seja entregue com alto grau de especificidade a um determinado tipo de tecido. Assim, um vetor viral pode ser modificado para que tenha especificidade para um dado tipo de célula, expressando um ligante como uma proteína de fusão com uma proteína do revestimento viral na superfície exterior do vírus. O ligante é escolhido pela afinidade por um receptor conhecido como presente no tipo celular de interesse. Por exemplo, Han, et al. (1995) *Proc. Natl.*

Acad. Sci. USA 92:9747-9751, relaram que o vírus da leucemia murina de Moloney pode ser modificado para expressar heregulina humana fundida a gp70, e o vírus recombinante infecta certas células de câncer de mama humano que expressam o receptor do fator de crescimento epidérmico humano. Esse princípio pode ser estendido a outros pares de vírus-célula alvo, em que a célula alvo expressa um receptor e o vírus expressa uma proteína de fusão compreendendo um ligante para o receptor na superfície da célula. Por exemplo, fagos filamentosos podem ser projetados para exibirem fragmentos de anticorpos (por exemplo, FAB ou Fv) com afinidade de ligação específica por virtualmente qualquer receptor celular escolhido. Embora a descrição acima aplique-se primariamente a vetores virais, os mesmos princípios podem ser aplicados a vetores não virais. Tais vetores podem ser projetados para conter sequências de captação específica que favoreçam a captação por células alvo específicas.

[00185] Os vetores de terapia gênica podem ser entregues *in vivo* por administração a um paciente individual, tipicamente por administração sistêmica (por exemplo, infusão intravenosa, intraperitoneal, intramuscular, subdérmica ou intracraniana) ou aplicação tópica, como descrito abaixo. Alternativamente, os vetores podem ser entregues a células *ex vivo*, como células explantadas de um paciente individual (por exemplo, linfócitos, aspirados de medula óssea, biópsia de tecido) ou células-tronco hematopoiéticas doadoras universais, seguido por reimplante das células em um paciente, normalmente após a seleção de células que incorporaram o vetor.

[00186] As células, como aqui descritas, podem também ser utilizadas para terapias celulares, por exemplo, terapia celular adotiva para o tratamento e/ou prevenção de um câncer. Terapia celular é um tipo especializado de transplante em que células de certo tipo (por exemplo, células T reativas a antígeno tumoral ou células B) são

fornecidas a um receptor. A terapia celular pode ser realizada com células que são autólogas (derivadas do receptor) ou alógénicas (derivadas de um doador), e as células podem ser células imaturas como células-tronco, ou células completamente maduras e funcionais como células T. De fato, em algumas doenças como certos cânceres, células T podem ser manipuladas *ex vivo* para aumentar sua avidez por certos抗ígenos tumorais, expandidas e, a seguir, introduzidas no paciente que sofre daquele tipo de câncer em uma tentativa para erradicar o tumor. Isso é especialmente útil quando a resposta de células T autógenas é suprimida pelo próprio tumor.

[00187] A transfecção de células *ex vivo* para diagnóstico, pesquisa, transplante ou para terapia gênica e/ou celular (por exemplo, através da reinfusão das células transfectadas no organismo hospedeiro) é bem conhecida pelos técnicos no assunto. Em uma modalidade preferida, as células são isoladas do organismo em questão, transfectada com um ácido nucleico (gene ou cDNA) de proteínas de ligação ao DNA e voltam a ser infundidas no organismo em questão (por exemplo, paciente). Vários tipos de células adequadas para transfecção *ex vivo* transfecção são bem conhecidos pelos técnicos no assunto (ver, por exemplo, Freshney, *et al.*, *Culture of Animal Cells, A Manual of Basic Technique* (3a ed., 1994)) e as referências ali citadas para uma discussão do modo como isolar e cultivar células de pacientes).

[00188] Em uma modalidade, células-tronco são usadas em procedimentos *ex vivo* para transfecção das células e terapia gênica. A vantagem de se usar células-tronco é que estas podem ser diferenciadas em outros tipos de células *in vitro* ou podem ser introduzidas em um mamífero (como o doador das células) onde serão enxertadas na medula óssea. Métodos para diferenciar células CD34+ *in vitro* em tipos de células imunes clinicamente importantes utilizando

citocinas, como GM-CSF, IFN- γ e TNF- α , são conhecidos (ver Inaba, et al. (1992) *J. Exp. Med.* 176:1693-1702).

[00189] Células-tronco são isoladas para transdução e diferenciação empregando métodos conhecidos. Por exemplo, células-tronco são isoladas de células da medula óssea pela técnica de *panning* das células da medula óssea com anticorpos que ligam células não desejadas, como CD4+ e CD8+ (células T), CD45+ (células panB), GR-1 (granulócitos) e Iad (células apresentadoras de抗ígenos diferenciadas) (ver Inaba, et al. (1992) *J. Exp. Med.* 176:1693-1702).

[00190] Células-tronco que foram modificadas podem também ser usadas em algumas modalidades. Por exemplo, células-tronco neuronais que foram tornadas resistentes à apoptose podem ser utilizadas como composições terapêuticas, em que as células-tronco também podem conter os ZFP-TFs da invenção. A resistência à apoptose pode ser ocasionada, por exemplo, pelo *knockout* (silenciamento) de BAX e/ou BAK, utilizando ZFNs específicas para BAX ou BAK (ver a Patente U.S. N° 8 597 912) nas células-tronco, ou aquelas que sofreram alteração em uma caspase, novamente utilizando ZFNs específicas para caspase-6, por exemplo. Essas células podem ser transfectadas com os ZFP-TFs conhecidos por regular TCR.

[00191] Vetores (por exemplo, retrovírus, adenovírus, lipossomos, etc.) contendo proteínas terapêuticas de ligação ao DNA (ou ácidos nucleicos que codificam essas proteínas) podem também ser administrados diretamente a um organismo para transdução de células *in vivo*. Alternativamente, DNA nu pode ser administrado. A administração é por qualquer uma das vias normalmente utilizadas para introduzir uma molécula em contato final com sangue ou células de tecidos, incluindo, entre outros, injeção, infusão, aplicação tópica e

eletroporação. Métodos para administração de tais ácidos nucleicos estão disponíveis e são bem conhecidos pelos técnicos no assunto e, embora mais de uma via possa ser utilizada para administrar uma composição em particular, uma via específica pode frequentemente propiciar uma reação mais imediata e mais eficaz do que outra via.

[00192] Métodos para introdução de DNA em células-tronco hematopoiéticas são revelados, por exemplo, na Patente U.S. Nº 5 928 638. Vetores úteis para introdução de transgenes em células-tronco hematopoiéticas, por exemplo, células CD34+, incluem adenovírus Tipo 35.

[00193] Os vetores adequados para introdução de transgenes em células imunes (por exemplo, células T) incluem vetores de lentivírus não integrantes. Ver, por exemplo, Ory, *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:11382-11388; Dull, *et al.* (1998) *J. Virol.* 72:8463-8471; Zuffery, *et al.* (1998) *J. Virol.* 72:9873-9880; Follenzi, *et al.* (2000) *Nature Genetics* 25:217-222.

[00194] Veículos farmaceuticamente aceitáveis são determinados, em parte, pela composição em particular sendo administrada, bem como pelo método específico utilizado para administrar a composição. Assim, há uma ampla variedade disponível de formulações adequadas de composições farmacêuticas, como descritas abaixo (ver, por exemplo, *Remington's Pharmaceutical Sciences*, 17a ed., 1989).

[00195] Tal como mencionado acima, os métodos e as composições revelados podem ser utilizados em qualquer tipo de célula incluindo, entre outras, células procarióticas, células de fungos, células de *Archaea*, células de plantas, células de insetos, células animais, células de vertebrados, células de mamíferos e células humanas, incluindo células T e células-tronco de qualquer tipo. Linhagens celulares adequadas para a expressão de proteínas são conhecidas pelos técnicos no assunto e incluem, entre outras COS,

CHO (por exemplo, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (por exemplo, HEK293-F, HEK293-H, HEK293-T), perC6, células de inseto como *Spodoptera frugiperda* (Sf) e células de fungos como *Saccharomyces*, *Pichia* e *Schizosaccharomyces*. A progénie, variantes e derivadas dessas linhagens celulares podem também ser utilizadas.

[00196] Aplicações

[00197] As composições e os métodos revelados podem ser utilizados para qualquer aplicação na qual se deseja modular a expressão e/ou a funcionalidade de TCR e/ou B2M, incluindo, entre outras, aplicações terapêuticas e em pesquisa nas quais a modulação de TCR e/ou B2M é desejável. Por exemplo, as composições reveladas podem ser utilizadas *in vivo* e/ou *ex vivo* (terapias celulares) para afetar a expressão de TCRs e/ou B2M endógenos em células T modificadas para terapia celular adotiva com objetivo de expressar um ou mais CARs exógenos, TCRs exógenos ou outras moléculas receptoras específicas de câncer, pelo qual tratando e/ou prevenindo o câncer. As células T podem ser células T efetoras ou células T reguladoras. Além disso, em tais contextos, a revogação da expressão de TCR em uma célula pode eliminar ou reduzir substancialmente o risco de uma reação cruzada indesejada com tecido sadio, não visado (ou seja, uma resposta do enxerto contra hospedeiro). As células modificadas com aqui descrito podem também ser utilizadas para o tratamento de cânceres, incluindo, entre outros, câncer de próstata, leucemia linfocítica crônica (LLC) e linfomas não Hodgkin.

[00198] Os métodos e composições também incluem composições de células-tronco (por exemplo, iPSC e HSC/HSPC), em que os genes B2M, TCRA e/ou TCRB dentro das células-tronco foram modulados (modificados) e as células compreendem ainda um ACTR e/ou um

CAR e/ou um TCR isolado ou projetado. Por exemplo, células-tronco hematopoiéticas alogênicas moduladas pelo *knockout* (silenciamento) ou *knockdown* (diminuição da expressão) de TCR podem ser introduzidas em um paciente com HLA compatível após ablação da medula óssea. Essas HSC alteradas permitiriam a recolonização do paciente, mas não provocariam uma possível doença do enxerto contra hospedeiro (DECH). As células introduzidas podem também ter outras alterações como ajuda durante a terapia subsequente (por exemplo, resistência à quimioterapia) para tratar a doença subjacente. As células com HLA classe I nula também têm utilidade como uma terapia "off the shelf" (para uso imediato) em situações no pronto-socorro de pacientes com trauma.

[00199] Os métodos e as composições da invenção são igualmente úteis para o desenho e a implementação de modelos *in vitro* e *in vivo*, por exemplo, modelos animais de TCR ou B2M e transtornos associados, o que permite o estudo desses transtornos.

[00200] Todas as patentes, os pedidos de patente e publicações mencionados são aqui incorporados, em suas totalidades, por referência neste pedido de patente.

[00201] Embora a descrição tenha sido fornecida em alguns detalhes a título de ilustração e exemplo por questões de clareza e entendimento, será evidente para os técnicos no assunto que várias alterações e modificações podem ser praticadas sem que, contudo, se desviem do espírito e âmbito da invenção. Assim, a descrição anterior e os exemplos a seguir não devem ser interpretados como limitações.

Exemplos

Exemplo 1: Desenho de nucleases TCR-específicas

[00202] ZFNs TCR-específicas foram construídas para possibilitar a introdução sítio-específica de quebras de fita dupla no gene TCRA (TCRA). As ZFNs foram desenhadas essencialmente como descrito

em Urnov, *et al.* (2005) *Nature* 435(7042):646-651, Lombardo, *et al.* (2007) *Nat Biotechnol.* 25(11):1298-306, e Publicações de Patentes U.S. N^{os} 2008/0131962; 2015/016495; 2014/0120622; e 2014/0301990 e Patente U.S. N° 8 956 828. Os pares de ZFNs visavam sítios diferentes na região constante do gene TCRA (ver Figura 1). As hélices de reconhecimento de pares exemplares de ZFNs, bem como a sequência alvo são mostrados abaixo na Tabela 1. Os sítios alvo dos desenhos de TCRA dedo de zinco são mostrados na primeira coluna. Os nucleotídeos no sítio alvo que são visados pelas hélices de reconhecimento de ZFP estão indicados em letras maiúsculas; os nucleotídeos não visados indicados em letras minúsculas. Os ligantes usados para unir o domínio nuclease FokI e o domínio de ligação ao DNA de ZFP são também mostrados (ver Publicação de Patente U.S. N° 2015/0132269). Por exemplo, a sequência de aminoácidos do ligante L0 do domínio é domínio de ligação ao DNA-QLVKS-domínio nuclease FokI (SEQ ID N°:5). Do mesmo modo, as sequências de aminoácidos para o ligante N7a do domínio é domínio nuclease FokI-SGTPHEVGVYTL- domínio de ligação ao DNA (SEQ ID N°:6), e N7c é domínio nuclease FokI-SGAIRCHDEFWF-domínio de ligação ao DNA (SEQ ID N°:7).

Tabela 1: Desenhos de TCR- α (TCRA) dedo de zinco

Nome de ZFN Sequência alvo	F1	F2	F3	F4	F5	F6	Ligante do domínio
SBS55204 5'ttGCTCTTGAAGTCcATAGACctcatgt (SEQ ID N°:8)	DRSNLSR (SEQ ID N°:22)	QKVTLAA (SEQ ID N°:23)	DRSALSR (SEQ ID N°:24)	TSGNLTR (SEQ ID N°:25)	YRSSLKE (SEQ ID N°:26)	TSGNLTR (SEQ ID N°:25)	L0
SBS53759 5'gtGCTGTGgCCTGGAGCAACAAatcta (SEQ ID N°:9)	QQNVLIN (SEQ ID N°:27)	QNATRTK (SEQ ID N°:28)	QSGHLAR (SEQ ID N°:29)	NRYDLMT (SEQ ID N°:30)	RSDSLLR (SEQ ID N°:31)	QSSDLTR (SEQ ID N°:32)	L0
SBS55229 5'ctGTTGCTCTTGAAGTCcatagaccta	DRSALAR (SEQ ID	QSGNLAR (SEQ ID	HRSTLQG (SEQ ID	QSGDLTR (SEQ ID	TSGSLTR (SEQ ID	NA	L0

Nome de ZFN Sequência alvo	F1	F2	F3	F4	F5	F6	Ligante do domínio
(SEQ ID Nº:10)	Nº:33)	Nº:34)	Nº:35)	Nº:36)	Nº:37)		
SBS53785 5'ctGTGGCCtGGAGCAACAAatctgact (SEQ ID Nº:11)	QHQVLVR (SEQ ID Nº:38)	QNATRTK (SEQ ID Nº:28)	QSGHLSR (SEQ ID Nº:39)	DRSDLSR (SEQ ID Nº:40)	RSDALAR (SEQ ID Nº:41)	NA	L0
SBS53810 5'agGATTCGGAACCCAATCACtg (SEQ ID Nº:12)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	DSSTRKT (SEQ ID Nº:44)	QSGNLAR (SEQ ID Nº:34)	RSDDLSE (SEQ ID Nº:45)	TNSNR KR (SEQ ID Nº:46)	L0
SBS55255 5'ctCCTGAAAGTGGCCGGgttaatctgc (SEQ ID Nº:13)	RSDHLST (SEQ ID Nº:47)	DRSHLAR (SEQ ID Nº:48)	LKQHLNE (SEQ ID Nº:49)	TSGNLTR (SEQ ID Nº:25)	HRTSLTD (SEQ ID Nº:50)	NA	L0
SBS55248 5'agGATTCGGAACCCAATCACtgacagg (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLS R (SEQ ID Nº:53)	L0
SBS55254 5'ctCCTGAAAGTGGCCGGgttaatctgc (SEQ ID Nº:13)	RSDHLST (SEQ ID Nº:47)	DRSHLAR (SEQ ID Nº:48)	LKQHLNE (SEQ ID Nº:49)	QSGNLAR (SEQ ID Nº:34)	HNSSLKD (SEQ ID Nº:54)	NA	L0
SBS55260 5'ctCCTGAAAGTGGCCGGgttaatctgc (SEQ ID Nº:13)	RSDHLST (SEQ ID Nº:47)	DRSHLAR (SEQ ID Nº:48)	LNHHLQQ (SEQ ID Nº:55)	QSGNLAR (SEQ ID Nº:34)	HKTSLKD (SEQ ID Nº:56)	NA	L0
SBS55266 5'tcAAGCTGGTCGAGaAAAGCTtgtaac (SEQ ID Nº:15)	QSSDLSR (SEQ ID Nº:57)	QSGNRTT (SEQ ID Nº:58)	RSANLAR (SEQ ID Nº:59)	DRSALAR (SEQ ID Nº:33)	RSDVLSE (SEQ ID Nº:60)	KHSTR RV (SEQ ID Nº:61)	N7c
SBS53853 5'aaCAGGTAAcGACAGGGGTCTAgcctgg g (SEQ ID Nº:16)	TMHQRVE (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHILTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNL DA (SEQ ID Nº:67)	L0
SBS53860 5'ctGTGCTAGACATGaGGTCTAtggact (SEQ ID Nº:17)	TMHQRVE (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RNDSLKT (SEQ ID Nº:68)	DSSNLSR (SEQ ID Nº:69)	QKATRTT (SEQ ID Nº:70)	RNASR TR (SEQ ID Nº:72)	N7a
SBS53863 5'ttCAAGAGCAACAGtGCTGTGgcctgga (SEQ ID Nº:18)	RSDSLLR (SEQ ID Nº:31)	QSSDLRR (SEQ ID Nº:73)	RSDNLSE (SEQ ID Nº:74)	ERANRNS (SEQ ID Nº:75)	RSDNLAR (SEQ ID Nº:76)	QKVNL MS (SEQ ID Nº:77)	L0
SBS55287 5'ttCAAGAGCAACAGtGCTGTGgcctgga (SEQ ID Nº:18)	RSDSLLR (SEQ ID Nº:31)	QSSDLRR (SEQ ID Nº:73)	RSDNLSE (SEQ ID Nº:74)	ERANRNS (SEQ ID Nº:75)	RSDNLAR (SEQ ID Nº:76)	QKVNL RE (SEQ ID	L0

Nome de ZFN Sequência alvo	F1	F2	F3	F4	F5	F6	Ligante do domínio
						Nº:78)	
SBS53855 5'ctGTGCTAGACATGaGGTCTAtggactt (SEQ ID Nº:17)	TMHQRVE (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDTLSQ (SEQ ID Nº:79)	DRSDLSR (SEQ ID Nº:40)	QKATRTT (SEQ ID Nº:70)	RNASR TR (SEQ ID Nº:72)	N7a
SBS53885 5'ccTGTCAcTGATTGGGTCCGaatctc (SEQ ID Nº:19)	RSDTLSE (SEQ ID Nº:79)	TSGSLTR (SEQ ID Nº:37)	RSDHLST (SEQ ID Nº:47)	TSSNRTK (SEQ ID Nº:71)	RSDNLSE (SEQ ID Nº:74)	WHSSL RV (SEQ ID Nº:83)	N7a
SBS52774 5'ccTGTCAcTGATTGGGTCCGaatctc (SEQ ID Nº:19)	RKQTRTT (SEQ ID Nº:80)	HRSSLRR (SEQ ID Nº:81)	RSDHLST (SEQ ID Nº:47)	TSANLSR (SEQ ID Nº:82)	RSDNLSE (SEQ ID Nº:74)	WHSSL RV (SEQ ID Nº:83)	N7a
SBS53909 5'tcCTCCTGAAAGTGGCCGGGIttaatct (SEQ ID Nº:20)	RSAHLSR (SEQ ID Nº:84)	DRSDLSR (SEQ ID Nº:40)	RSDVLSV (SEQ ID Nº:85)	QNNHRIT (SEQ ID Nº:86)	RSDVLSE (SEQ ID Nº:60)	SPSSR RT (SEQ ID Nº:87)	L0
SBS52742 5'tcCTCCTGAAAGTGGCCGGGIttaatct (SEQ ID Nº:20)	RSAHLSR (SEQ ID Nº:84)	DRSDLSR (SEQ ID Nº:40)	RSDSLSV (SEQ ID Nº:88)	QNANRK (SEQ ID Nº:89)	RSDVLSE (SEQ ID Nº:60)	SPSSR RT (SEQ ID Nº:87)	L0
SBS53856 5'ctGTGCTAGACATGaGGTCTAtg (SEQ ID Nº:21)	TMHQRVE (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDSLST (SEQ ID Nº:90)	DRANRIK (SEQ ID Nº:91)	QKATRTT (SEQ ID Nº:70)	RNASR TR (SEQ ID Nº:72)	N7a

[00203] Todas as ZFNs foram testadas e demonstraram ligar-se aos seus sítios alvo e serem ativas como nucleases.

[00204] As ZFPs aqui descritas podem também incluir uma ou mais mutações em resíduos da proteína dedo de zinco em contato com fosfato e/ou no domínio FokI, por exemplo, o mutante nR-5Qabc (no esqueleto de ZFP) e/ou mutantes R416S e/ou K525S (em FokI), descritas na Publicação de Patente U.S. Nº 20180087072.

[00205] Construíram-se também RNAs guias para o sistema CRISPR/Cas9 de *S. pyogenes* para atingir o gene TCRA. Ver, também, a Publicação de Patente U.S. Nº 2015/00566705 para RNAs guias adicionais direcionados para TCR alfa. As sequências alvo no

gene TCRA estão indicadas, bem como as sequências do RNA guia na Tabela 2 abaixo. Todos os RNAs guias são testados no sistema CRISPR/Cas9 e verificados ser ativos.

Tabela 2: RNAs guias para a região constante de TCRA humano (TRAC)

Nome	Fita	Alvo (5'->3')	gRNA (5' > 3')
TRAC-Gr14	R	GCTGGTACACGGCAGGGTCAGGG (SEQ ID Nº:92)	GCTGGTACACGGCAGGGTCA (SEQ ID Nº:104)
TRAC-Gr25	R	AGAGTCTCTCAGCTGGTACACGG (SEQ ID Nº:93)	gAGAGTCTCTCAGCTGGTACA (SEQ ID Nº:105)
TRAC-Gr71	R	GAGAATCAAATCGGTGAATAGG (SEQ ID Nº:94)	GAGAATCAAATCGGTGAAT (SEQ ID Nº:106)
TRAC-Gf155	F	ACAAAACTGTGCTAGACATGAGG (SEQ ID Nº:95)	gACAAAACTGTGCTAGACATG (SEQ ID Nº:107)
TRAC-Gf191	F	AGAGAACAGTGCTGTGGCTGG (SEQ ID Nº:96)	gAGAGAACAGTGCTGTGGCC (SEQ ID Nº:108)
TRAC-Gf271	F	GACACCTTCTCCCCAGCCCAGG (SEQ ID Nº:97)	GACACCTTCTCCCCAGCCC (SEQ ID Nº:109)
TRAC-Gr2146	R	CTCGACCAGCTTGACATCACAGG (SEQ ID Nº:98)	gCTCGACCAGCTTGACATCAC (SEQ ID Nº:110)
TRAC-Gf2157	F	AAGTTCCCTGTGATGTCAAGCTGG (SEQ ID Nº:99)	gAAGTTCCCTGTGATGTCAAGC (SEQ ID Nº:111)
TRAC-Gf2179	F	GTCGAGAAAGCTTGAACAGG (SEQ ID Nº:100)	GTCGAGAAAGCTTGAAC (SEQ ID Nº:112)
TRAC-Gr3081	R	TTCGGAACCCAATCACTGACAGG (SEQ ID Nº:101)	gTTCGGAACCCAATCACTGAC (SEQ ID Nº:113)
TRAC-Gr3099	R	CCACTTTCAGGAGGAGGATTCTGG (SEQ ID Nº:102)	gCCACTTTCAGGAGGAGGATT (SEQ ID Nº:114)
TRAC-Gr3105	R	ACCCGGCCACTTCAGGAGGAGG (SEQ ID Nº:103)	gACCCGGCCACTTCAGGAGG (SEQ ID Nº:115)

[00206] Assim, as nucleases aqui descritas (por exemplo, nucleases compreendendo um domínio de ligação ao DNA de ZFP, de TALE ou de sgRNA) ligam-se a seus sítios alvo e clivam o gene TCRA, fazendo, desse modo, modificações genéticas dentro de um

gene TCRA compreendendo qualquer uma dentre SEQ ID Nº:6-48 ou 137-205, incluindo modificações (inserções e/ou deleções) dentro de qualquer uma dessas sequências (por exemplo, as sequências alvo mostradas em qualquer uma dentre SEQ ID Nº:8-21 e/ou 92-103; 12-25 nucleotídeos desses sítios alvo; e/ou entre sítios alvo pareados) e/ou modificações dentro das sequências seguintes: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e/ou AATCCTC (ver, Figura 1B). As nucleases TALE direcionadas para esses sítios alvo são também desenhadas e demonstraram ser funcionais em termos de ligação e atividade.

[00207] Além disso, todos os domínios de ligação ao DNA (ZFPs e sgRNAs) ligaram-se aos seus sítios alvo, e os domínios de ligação ao DNA de ZFP, TALE e sRNA que reconhecem esses sítios alvo são também formulados em fatores de transcrição projetados ativos quando associados com um ou mais domínios reguladores da transcrição.

Exemplo 2: Atividade nuclease *in vitro*

[00208] As ZFNs descritas na Tabela 1 foram utilizadas para testar a atividade nuclease em células K562. Para testar a atividade de clivagem, plasmídeos que codificam os pares de ZFNs específicos para TCRA humano, descritos acima, foram transfetados em células K562 com plasmídeo ou mRNAs. As células K562 foram obtidas da *American Type Culture Collection* e cultivadas como recomendado em meio RPMI (Invitrogen) suplementado com soro fetal bovino qualificado 10% (FBS, Cyclone). Para transfeção, ORFs para as nucleases ativas listadas na Tabela 1 foram clonadas em um vetor de expressão otimizado para produção de mRNA com UTRs 5' e 3' e um sinal poliA sintético. Os mRNAs foram gerados com o kit mMessage mMachine T7 Ultra (Ambion) seguindo as instruções do fabricante. A síntese *in vitro* de mRNAs das nucleases utilizou um vetor baseado em

pVAX contendo um promotor T7, a nuclease correta e um motivo poliA para adição enzimática de uma cauda poliA após a reação de transcrição *in vitro*, ou um vetor baseado em pGEM contendo um promotor T7, uma 5'UTR, a nuclease correta, uma 3'UTR e um trecho poliA de 64 bp, ou um amplicon de PCR contendo um promotor T7, uma 5'UTR, a nuclease correta, uma 3'UTR e um trecho poliA de 60 bp. Um milhão de células K562 foram misturadas com 250 ng ou 500 ng do mRNA codificador da ZFN. As células foram transfectadas em um Amaxa Nucleofector IITM, utilizando o programa T-16, e recuperadas em 1,4 mL de meio RPMI morno + FBS 10%. A atividade nuclease foi avaliada por sequenciamento profundo (MiSeq, Illumina) de acordo com protocolos padrão, três dias após a transfecção. Os resultados são apresentados abaixo na Tabela 3.

Tabela 3: Atividade de nuclease dedo de zinco

Nº do par	Par de ZFNs	NHEJ % (250 ng/ZFN)	DP	NHEJ % (500 ng/ZFN)	DP	Sítio
1	55204:53759	76,7	1,3	87,7	1	A2
2	55229:53785	91,4	1,5	93,6	1,7	B
3	53810:55255	81,6	0,6	91,5	1,3	D1
4	55248:55254	95,4	1,8	96,2	1,2	D2
5	55248:55260	87,9	1,3	93,0	1	D3
6	55266:53853	85,3	1,4	88,9	0,4	E
7	53860:53863	77,1	1,7	87,3	1,1	F1
8	53856:55287	53,6	3,2	74,8	1,3	F2
9	53885:53909	90,1	1,6	90,2	1,5	G1
10	52774:52742	76,8	0,8	84,4	2,2	G0
11	GFP	0		0		

[00209] TALENs específicas para TCRA altamente ativas foram também descritas anteriormente (ver Publicação da Patente Internacional N° WO 2014/153470).

[00210] Os sistemas CRISPR/Cas9 específicos para TCRA humano foram também testados. A atividade dos sistemas CRISPR/Cas9 em células K562 humanas foi medida por análise MiSeq. A clivagem da sequência de DNA do TCRA endógeno por Cas9 é analisada por sequenciamento em alta escala (Miseq, Illumina).

[00211] Nesses experimentos, Cas9 foi fornecida em um plasmídeo pVAX, e o sgRNA é fornecido em um plasmídeo sob o controle de um promotor (por exemplo, o promotor U6 ou um promotor CMV). Os plasmídeos foram misturados a 100 ng de cada ou 400 ng de cada e foram misturados com 2^5 células por rodada. As células foram transfetadas utilizando o sistema Amaxa. Resumidamente, um kit de transfecção Amaxa é utilizado e os ácidos nucleicos são transfetados utilizando um protocolo padrão do *shuttle* Amaxa. Após a transfecção, as células são deixadas em repouso por 10 minutos à temperatura ambiente e, a seguir, ressuspensas em RMPI preaquecido. As células foram então cultivadas em condições padrão a 37°C. O DNA genômico foi isolado 7 dias depois da transfecção e submetido à análise MiSeq.

[00212] Resumidamente, os RNAs guias listados na Tabela 2 foram testados quanto à atividade. Os RNAs guias foram testados em três configurações diferentes: G0 é a montagem descrita acima. G1 usou um vetor pVAX compreendendo um promotor de CMV direcionando a expressão do gene Cas9 e um cassete de expressão com U6-RNA guia-rastreador, em que a transcrição de ambas as fases de leitura é na mesma orientação. G2 é semelhante a G1 exceto que o Cas9 e os cassetes de expressão com U6-Guia estão em orientações opostas. Essas três montagens foram testadas utilizando 100 ng ou 400 ng de DNA transfetado, e os resultados são apresentados abaixo na Tabela

4. Os resultados são expressos nos "indels percentuais" ou "NHEJ %", em que "indels" significa pequenas inserções e/ou deleções encontradas como resultado do processo de reparo NHEJ propenso a erro no sítio de uma clivagem de dupla fita induzida por nuclease.

Tabela 4: Atividade de CRISPR/Cas

	% de indels totais					
	GR0		GR1		GR2	
	Guia usado	NHEJ % (100 ng)	NHEJ % (400 ng)	NHEJ % (100 ng)	NHEJ % (400 ng)	NHEJ % (100 ng)
TCRA-Gr14	6,4	25,8	0,6	12,4	0,5	10,2
TCRA-Gr25	14,6	26,9	2,4	21,7	1,1	21,6
TCRA-Gr71	3,7	13,8	0,3	4,2	0,3	7,8
TCRA-Gf155	6,0	19,5	1,2	12,7	0,8	15,9
TCRA-Gf191	1,0	6,9	0,3	2,3	0,4	4,5
TCRA-Gf271	4,7	21,5	0,8	10,3	0,7	15,2
TCRA-Gr2146	1,1	8,8	0,3	1,7	0,2	2,0
TCRA-Gf2157	3,8	22,2	0,6	9,6	0,6	12,0
TCRA-Gf2179	0,8	4,9	0,2	1,8	0,2	1,4
TCRA-Gr3081	5,9	23,6	0,7	11,5	0,8	12,6
TCRA-Gr3099	2,1	21,1	0,4	7,1	0,3	6,2
TCRA-Gr3105	12,1	45,9	2,2	22,0	1,0	7,6
ZFN controles						
55248:55254	24,2	52,4				
55229:53785	6,0	24,5				
55266:53853	12,0	37,0				

[00213] Como mostrado, as nucleases aqui descritas induzem clivagem e modificações genômicas no sítio visado.

[00214] Assim, as nucleases aqui descritas (por exemplo, nucleases compreendendo um domínio de ligação ao DNA de ZFP, de TALE ou de sgRNA) ligam-se aos seus sítios alvo e clivam o gene TCRA, pelo qual efetuando modificações genéticas dentro de um gene

TCRA compreendendo qualquer uma dentre SEQ ID Nº:8-21 ou 92-103, incluindo modificações (inserções e/ou deleções) dentro de qualquer uma dessas sequências (SEQ ID Nº:8-21, 92-103); modificações dentro de 1-50 (por exemplo, 1 a 10) pares de base dessas sequências gênicas; modificações entre sítios alvo de sítios alvo pareados (para dímeros); e/ou modificações dentro de uma ou mais das sequências seguintes: AACAGT, AGTGCT, CTCCT, TTGAAA, TGGACTT e/ou AATCCTC (ver, Figura 1B).

[00215] Além do mais, todos os domínios de ligação ao DNA (ZFPs, TALEs e sgRNAs) ligaram-se aos seus sítios alvo e também são formulados em fatores de transcrição projetados ativos quando associados com um ou mais domínios reguladores da transcrição.

Exemplo 3: Atividade de ZFN específica para TCRA em células T

[00216] Os pares de ZFNs específicos para TCRA foram também testados em células T humanas quanto à atividade de nuclease. mRNAs codificadores das ZFNs foram transfetados em células T purificadas. Resumidamente, as células T foram obtidas de um derivado de leucóferese e purificadas com o sistema Miltenyi CliniMACS (seleção dupla de CD4 e CD8). Essas células foram então ativadas utilizando Dynabeads (ThermoFisher) de acordo com o protocolo do fabricante. 3 dias após a ativação, as células foram transfetados com três doses de mRNA (60, 120 e 250 µg/mL) utilizando um eletroporador Maxcyte (Maxcyte), OC-100, 30⁶ células/mL, volume de 0,1 mL. As células foram analisadas para modificação de TCRA no alvo por meio de sequenciamento profundo (Miseq, Illumina), 10 dias após a transfecção. Mediram-se a viabilidade celular e o crescimento celular (duplicações de células totais) por todos os 13-14 dias de cultura. Além disso, TCR na superfície celular das células tratadas foi medido por análise FACS padrão no 10º dia de cultura com coloração para CD3.

[00217] Todos os pares de ZFNs específicos para TCRA foram ativos em células T e alguns foram capazes de provocar mais de 80% de modificação no alelo de TCRA nessas condições (ver Figuras 2A e 2B). Do mesmo modo, células T tratadas com as ZFNs perderam a expressão de CD3, em que a análise FACS revelou que, em alguns casos, entre 80 e 90% das células T eram negativas para CD3 (Figura 3). Uma comparação entre o percentual de TRAC modificado por ZFN e perda de CD3 nessas células demonstrou um alto grau de correlação (Figura 4). A viabilidade celular foi comparável aos controles de tratamento simulado, e o crescimento celular com *knockout* de TCRA foi também comparável aos controles (ver Figura 5A-5D).

Exemplo 4: Knockout duplo de B2M e TCRA com integração direcionada

[00218] Nucleases como descritas acima e as nucleases direcionadas para B2M descritas na Tabela 5 (ver também Publicação de Patente U.S. Nº 2017/0173080) foram utilizadas para inativar B2M e TCRA e para introduzir, via integração direcionada, um doador (transgene) no lócus de TCRA ou de B2M. As ZFNs específicas para B2M são mostradas abaixo na Tabela 5:

Tabela 5: Desenhos de ZFN B2M-específica

Nome de ZFN Sequência alvo	F1	F2	F3	F4	F5	F6	Ligante do domínio
SBS57327 5' taGCAATTAGGAAaTTTG ACttccat (SEQ ID Nº:123)	DRSNLSR (SEQ ID Nº:22)	ARWYLDK (SEQ ID Nº:125)	QSGNLA R (SEQ ID Nº:34)	AKWNLD A (SEQ ID Nº:67)	QQHVQLN Q (SEQ ID Nº:119)	QNATRTK (SEQ ID Nº:28)	L0
SBS57332 5'tgTCGGATgGATGAAAC CCAGacacata (SEQ ID Nº:117)	RSDNLSE (SEQ ID Nº:74)	ASKTRTN (SEQ ID Nº:120)	QSGNLA R (SEQ ID Nº:34)	TSANLSE (SEQ ID Nº:82)	TSGNLTR (SEQ ID Nº:25)	RTEDRLA (SEQ ID Nº:121)	N6a
SBS57531 5' gaGTAGCGCgAGGCACAG Ctaaggccacg (SEQ ID Nº:126)	AQCCLFH (SEQ ID Nº:128)	DQSNLRA (SEQ ID Nº:42)	RSANLTR (SEQ ID Nº:129)	RSDDLTR (SEQ ID Nº:130)	QSGSLTR (SEQ ID Nº:66)	N/A	N6a
SBS57071 gcCACGGAgCGAGACATC TCGGccccaa (SEQ ID Nº:127)	RSDDLSE (SEQ ID Nº:131)	DSSARKK (SEQ ID Nº:132)	DRSNLSR (SEQ ID Nº:22)	QRTHLRD (SEQ ID Nº:133)	QSGHLAR (SEQ ID Nº:29)	DSSNREA (SEQ ID Nº:134)	L0

[00219] Nesse experimento, o par de ZFNs específico para TCRA era SBS#55266/SBS#53853, compreendendo a sequência TTGAAA entre os sítios alvo de ZFN específicos para TCRA (Tabela 1), e o par de B2M era SBS#57332/SBS#57327 (Tabela 5), compreendendo a sequência TCAAAT entre os sítios alvo de ZFN específicos para B2M.

[00220] Resumidamente, células T (AC-TC-006) foram descongeladas e ativadas com Dynabeads CD3/28 (razão 1:3 células:microesferas) em meio de cultura X-vivo15 de células T (dia 0). Depois de dois dias em cultura (dia 2), um doador AAV (compreendendo um transgene de GFP e braços de homologia ao gene TCRA ou B2M) foi adicionado à cultura celular, exceto grupos controle sem doador que também foram mantidos. No dia seguinte (dia 3), ZFNs para TCRA e B2M foram adicionadas através da entrega de mRNA nos 5 Grupos seguintes:

[00221] Grupo 1 (ZFNs para TCRA e B2M somente, sem doador): TCRA 120 µg/mL; B2M somente 60 µg/mL;

[00222] Grupo 2 (ZFNs para TCRA e B2M e doador com braços de homologia a TCRA): TCRA 120 µg/mL; B2M 60 µg/mL e AAV (TCRA-Sítio E-hPGK-eGFP-Clone E2) 1E5vg/célula;

[00223] Grupo 3 (ZFNs para TCRA e B2M e doador com braços de homologia a TCRA): TCRA 120 µg/mL; B2M 60 µg/mL; e AAV (TCRA-Sítio E-hPGK-eGFP-Clone E2) 3E4vg/célula;

[00224] Grupo 4 (ZFNs para TCRA e B2M e doador com braços de homologia a B2M): TCRA 120 µg/mL; B2M 60 µg/mL e AAV (pAAV B2M-hPGK GFP) 1E5vg/célula

[00225] Grupo 5 (ZFNs para TCRA e B2M e doador com braços de homologia a B2M): TCRA 120 µg/mL; B2M 60 µg/mL e AAV (pAAV B2M-hPGK GFP) 3E4vg/célula.

[00226] Todos os experimentos foram conduzidos com 3e7células/mL de densidade celular utilizando o protocolo como

descrito na Publicação de Patente U.S. Nº 2017/0137845 (choque por frio extremo) e as células foram cultivadas até choque por frio a 30°C durante a noite após a eletroporação.

[00227] No dia seguinte (dia 4), as células foram diluídas para 0,5e6 células/mL e transfectada para culturas a 37°C. Três dias depois (dia 7), as células foram diluídas para 0,5e6 células/mL novamente. Depois de três e sete dias mais em cultura (dias 10 e 14, respectivamente), as células foram colhidas para análise por FACS e MiSeq (diluídas para 0,5e6 células/mL).

[00228] Tal como mostrado na Figura 6, a expressão de GFP indicou que a integração no alvo foi bem-sucedida e que células geneticamente modificadas compreendendo modificações em B2M e TCRA (inserções e/ou deleções) dentro dos sítios alvo da nuclease (ou dentro de 1 a 50, 1-20, 1-10 ou 1-5 pares de base dos sítios alvo da nuclease), incluindo dentro das sequências TTGAAA e TCAAAT (entre os sítios alvo pareados), como aqui revelado, foram obtidas.

[00229] Experimentos adicionais foram realizados para gerar células com *knockouts* duplos de TRAC e B2M e integração direcionada de um vetor doador. Especificamente, o par de ZFNs específico para TRAC SBS#55266/SBS#53853 e o par para B2M SBS#57071/SBS#57531 foram introduzidos em células T. Resumidamente, células T humanas com razão 1:1 de CD4:CD8 foram descongeladas e ativadas com CD3/28 Dynabeads® (razão 1:3 células:microesferas) em meio de cultura X-vivo15 de células T (dia 0).

[00230] Depois de 3 dias em cultura (dia 3), as células foram concentradas até 3e7 células/mL em tampão de eletroporação Maxcyte na presença de mRNA de ZFN, então foram eletroporadas com o dispositivo Maxcyte. As células concentradas, eletroporadas foram colocadas, a seguir, em uma cavidade de cultura de tecido, em seguida, AAV6 codificador para um transgene doador hPGK-GFP-

BGHpoliA foi adicionado às células concentradas, que foram deixadas recuperar e incubadas a 37°C por 20 minutos. Alternativamente, o vetor doador pode ser adicionado ao tampão de eletroporação no dispositivo. As células foram então diluídas em meio de cultura até 3e6 células/mL e cultivadas a 30°C durante a noite. Na manhã seguinte, as células foram diluídas para 0.5e6 células/mL em meio de cultura adicional. A seguir, é apresentada uma descrição dos grupos:

[00231] *Sham*: células eletroporadas sem mRNA de ZFN ou AAV doador adicionado;

[00232] ZFNs para TRAC e B2M somente, sem doador): TRAC 120 µg/mL; B2M somente 30 µg/mL;

[00233] ZFNs para TRAC e B2M e doador com braços de homologia a B2M: TRAC 120 µg/mL; B2M 30 µg/mL e AAV6 (B2M-Sítio A-hPGK-eGFP) 3E4 vg/célula;

[00234] ZFNs para TRAC e B2M e doador com braços de homologia a TRAC: TRAC 120 µg/mL; B2M 30 µg/mL; e AAV6 (TCRA-Sítio E-hPGK-eGFP) 3E4 vg/célula.

[00235] Todos os experimentos foram conduzidos com 3e7 células/mL de densidade celular utilizando o protocolo descrito na Publicação de Patente U.S. N° 2017/0137845 (choque de frio extremo) e as células foram cultivadas até o choque de frio a 30°C durante a noite após a eletroporação. No dia seguinte (dia 4), as células foram diluídas para 0,5e6 células/mL e transferidas para culturas a 37°C. Três dias depois (dia 7), as células foram diluídas para 0,5e6 células/mL novamente. Depois de três e sete dias mais em cultura (dias 10 e 14, respectivamente), as células foram colhidas para análise por FACS e MiSeq (diluídas para 0,5e6 células/mL).

[00236] Como mostrado na Figura 7, a expressão de GFP (doador) indicou que a integração no alvo foi bem-sucedida e que células geneticamente modificadas compreendendo modificações em B2M e

TRAC (inserções e/ou deleções) dentro dos sítios alvo da nuclease (ou dentro de 1 a 50, 1-20, 1-10 ou 1-5 pares de base dos sítios alvo da nuclease, incluindo entre sítios pareados), como aqui revelados, foram obtidas com alta frequência (incluindo taxas de 80-90% de *knockout* e integração direcionada).

[00237] Os experimentos são também realizados, nos quais um transgene de CAR é integrado em *knockouts* duplos de B2M e TCRA, em qualquer lócus de B2M, TCRA ou outro para criar *knockouts* duplos B2M/TCRA que expressam um CAR.

Exemplo 5: Otimização de ZFNs para TCRA e B2M

[00238] Para diminuir a clivagem fora do alvo, uma estratégia para otimização da nuclease, na qual contatos não específicos com fosfato são seletivamente removidos para ocasionar a supressão global de clivagem fora do alvo (Guilinger, *et al.* (2014) *Nat Methods*. 11(4):429-35. doi: 10.1038/nmeth.2845; Kleinstiver, *et al.* (2016) *Nature* 529(7587):490-5. doi: 10.1038/nature16526; Slaymaker, *et al.* (2016) *Science* 351(6268):84-8. doi: 10.1126/science.aad5227) foi adotada (ver Publicação de Patente U.S. Nº 2018/0087072). Substituições de aminoácidos foram efetuadas em uma ou mais posições fundamentais dentro da estrutura dedo de zinco que interage com o esqueleto de fosfato do DNA (Pavletich and Pabo (1991) *Science* 252(5007):809-17; Elrod-Erickson, *et al.* (1996) *Structure* 4(10):1171-80) bem como nas posições no domínio da ZFN FokI direita, também previsto para contato com fosfato.

[00239] Na Tabela 6 abaixo, são mostradas informações que caracterizam cada ZFN. Partindo da esquerda, o número SBS (por exemplo, 55254) é exibido com o DNA alvo ao qual a ZFN liga-se que é exibido abaixo do número SBS. Em seguida, são mostrados os desenhos de hélice de reconhecimento de aminoácidos para os dedos 1-6 ou 1-5 (coluna 2 subdividida da Tabela 6). São também mostradas

na Tabela 6, abaixo dos desenhos adequados para hélice, as mutações feitas às sequências do esqueleto de ZFP do dedo indicado, como descrito no Pedido de Patente U.S. Nº 15/685,580. Na anotação usada na Tabela 6, "Qm5" significa que, na posição menos 5 (em relação à hélice que é numerada -1 a +6) do dedo indicado, a arginina nessa posição foi substituída por uma glutamina (Q), enquanto "Qm14" significa que a arginina (R) normalmente presente na posição menos 14 foi substituída por uma glutamina (Q). A abreviação "n", como em nQm5, significa que a mutação é no dedo N-terminal do módulo de dois dedos usado na construção da proteína com 5 ou 6 dedos. "Nenhum" indica nenhuma alteração fora da região da hélice de reconhecimento. Assim, por exemplo, SBS# 68797 inclui a mutação nQm5 nos dedos 1, 3 e 5 enquanto os dedos 2, 4 e 6 não têm mutações no esqueleto do dedo de zinco (por exemplo, a sequência de dedo de zinco fora da região da hélice de reconhecimento).

[00240] Finalmente, a coluna mais à direita da Tabela 6 mostra o ligante usado para ligar o domínio de ligação ao DNA ao domínio de clivagem FokI, por exemplo, "L0" LRGSQQLVKS (SEQ ID Nº:135), referido como o ligante "padrão" e descrito, por exemplo, na Patente U.S. Nº 9 567 609, é exibido na linha de cima da coluna, com os sítios das mutações em contato de FokI com fosfato e mutações de dimerização mostradas na caixa abaixo da designação do ligante. Outros ligantes incluem N7c (SGAIRCHDEFWF, SEQ ID Nº:136) e N7a (SGTPHEVGVYTL, SEQ ID Nº:137). Especificamente, na linha de cima das caixas de mutantes Fok, é indicado o tipo de mutação encontrada no domínio de dimerização (por exemplo, ELD ou KKR como descritas, por exemplo, na Patente U.S. Nº 8 962 281). Abaixo das designações de mutantes de dimerização, é mostrada qualquer mutação presente no domínio FokI, efetuada para remover um contato não específico com fosfato mostrado no fundo (por exemplo, K52S ou

R416S, onde resíduos de serina nas posições de aminoácido 525 ou 416 foram substituídas por lisina ou arginina, respectivamente, como descrito na Publicação U.S. N° 20180087072). Assim, por exemplo, em SBS# 68796, o ligante é um ligante L0 e o domínio de clivagem FokI inclui os mutantes de dimerização ELD e nenhuma mutação de contato com fosfato. Além disso, para SBS# 68812, o ligante é um ligante L0 e o domínio de clivagem FokI inclui as mutações de dimerização KKR onde o domínio FokI compreende ainda uma mutação de substituição R416E.

[00241] Outras variantes do domínio FokI que podem ser utilizadas com as ZFPs aqui descritas (incluindo ZFPs derivadas das ZFNs aqui descritas) incluem a adição de uma mutação de Sharkey (S418P+K441E, ver Guo, et al. (2010) *J. Mol Biol.*, doi:10.1016/j.jmb.2010.04.060) e as mutações DAD e RVR em FokI (ver a Patente U.S. N° 8 962 281). Exemplos não limitantes de variantes FokI projetadas que podem ser utilizadas incluem:

[00242] Domínio de clivagem de FokI do tipo selvagem (SEQ ID Nº:139):

[00243] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVEENQ TRNKHINPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00244] FokI-Sharkey (S418P+K441E, SEQ ID Nº:140):

[00245] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVEENQ TRNKHINPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00246] FokI ELD (Q->E @ 486, I->L @499, N->D @496, SEQ ID Nº:141)
 [00247] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00248] FokI ELD, Sharkey (Q->E @ 486, I->L @499, N->D @496,
 S418P+K441E SEQ ID Nº:142)
 [00249] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00250] FokI ELD, R416E (Q->E @ 486, I->L @499, N->D @496,
 R416E, SEQ ID Nº:143)
 [00251] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAENSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00252] FokI ELD, Sharkey, R416E (Q->E @ 486, I->L @499, N->D
 @496, S418P+K441E, R416E, SEQ ID Nº:144)
 [00253] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAENPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483

EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00254] FokI ELD, R416Y (Q->E @ 486, I->L @499, N->D @496,
 R416Y, SEQ ID Nº:145)
 [00255] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAYNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00256] FokI ELD, Sharkey, R416E (Q->E @ 486, I->L @499, N->D
 @496, S418P+K441E, R416E, SEQ ID Nº:146)
 [00257] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAYNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00258] FokI ELD, S418E (Q->E @ 486, I->L @499, N->D @496,
 S418E, SEQ ID Nº:147)
 [00259] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNETQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00260] FokI ELD, Sharkey parcial, S418E (Q->E @ 486, I->L
 @499, N->D @496, K441E, S418E, SEQ ID Nº:148)
 [00261] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNETQDRI

LEMKVMEEFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00262] FokI ELD, K525S (Q->E @ 486, I->L @499, N->D @496,
 K525S, SEQ ID Nº:149)
 [00263] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FSGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00264] FokI ELD, Sharkey K525S (Q->E @ 486, I->L @499, N->D
 @496, S418P+K441E, K525S, SEQ ID Nº:150)
 [00265] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FSGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00266] FokI ELD, I479T (Q->E @ 486, I->L @499, N->D @496,
 I479T, SEQ ID Nº:151)
 [00267] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPTGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00268] FokI ELD, Sharkey, I479T (Q->E @ 486, I->L @499, N->D

@496, S418P+K441E, I479T, SEQ ID N°:152)

[00269] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPTGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00270] FokI ELD, P478D (Q->E @ 486, I->L @499, N->D @496,
 P478D, SEQ ID N°:153)

[00271] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLDIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00272] FokI ELD, Sharkey, P478D (Q->E @ 486, I->L @499, N->D
 @496, S418P+K441E, P478D, SEQ ID N°:154)

[00273] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLDIGQAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00274] FokI ELD, Q481D (Q->E @ 486, I->L @499, N->D @496,
 Q481D, SEQ ID N°:155)

[00275] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGDAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00276] FokI ELD, Sharkey, Q481D (Q->E @ 486, I->L @499, N->D
 @496, S418P+K441E, Q481D, SEQ ID Nº:156)
 [00277] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGDAD 434- 483
 EMERYVEENQ TRDKHLPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00278] FokI KKR (E->K @490, I->K@538, H->R@537, SEQ ID
 Nº:157)
 [00279] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVKENQ TRNKHINPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00280] FokI KKR Sharkey, (E->K @490, I->K@538, H->R@537,
 S418P+K441E, SEQ ID Nº:158)
 [00281] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVKENQ TRNKHINPNE WWKVPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00282] FokI KKR, Q481E (E->K @490, I->K@538, H->R@537,
 Q481E, SEQ ID Nº:159)
 [00283] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGEAD 434- 483

EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00284] FokI KKR, Sharkey Q481E (E->K @490, I->K@538, H->R@537, S418P+K441E, Q481E, SEQ ID N°:160)
 [00285] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGEAD 434- 483
 EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00286] FokI KKR, R416E (E->K @490, I->K@538, H->R@537, R416E, SEQ ID N°:161)
 [00287] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAENSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00288] FokI KKR, Sharkey, R416E (E->K @490, I->K@538, H->R@537, S418P+K441E, R416E, SEQ ID N°:162)
 [00289] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAENPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00290] FokI KKR, K525S (E->K @490, I->K@538, H->R@537, K525S, SEQ ID N°:163)
 [00291] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI

LEMKVMEEFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRVYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FSGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00292] FokI KKR, Sharkey, K525S (E->K @490, I->K@538, H->R@537, S418P+K441E, K525S, SEQ ID Nº:164)
 [00293] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRVYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FSGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00294] FokI KKR, R416Y (E->K @490, I->K@538, H->R@537, R416Y, SEQ ID Nº:165)
 [00295] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAYNSTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRVYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00296] FokI KKR, Sharkey, R416Y (E->K @490, I->K@538, H->R@537, S418P+K441E, R416Y, SEQ ID Nº:166)
 [00297] QLVKSELEEK KSELRHKLKY VPHEYIELIE IAYNPTQDRI
 LEMKVMEEFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQRVYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00298] FokI, KKR I479T (E->K @490, I->K@538, H->R@537,

I479T, SEQ ID N°:167)

[00299] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
LEMKVMEFFM 384- 433 KVGYRGKHL GGSRKPDGAI

YTVGSPIDYG VIVDTKAYSG GYNLPTGQAD 434- 483

EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH

FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00300] FokI, KKR Sharkey I479T (E->K @490, I->K@538, H->R@537, S418P+K441E, I479T, SEQ ID N°:168)

[00301] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
LEMKVMEFFM 384- 433 KVGYRGEHL GGSRKPDGAI

YTVGSPIDYG VIVDTKAYSG GYNLPTGQAD 434- 483

EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH

FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00302] FokI, KKR P478D(E->K @490, I->K@538, H->R@537, P478D, SEQ ID N°:169)

[00303] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
LEMKVMEFFM 384- 433 KVGYRGKHL GGSRKPDGAI

YTVGSPIDYG VIVDTKAYSG GYNLDIGQAD 434- 483

EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH

FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRKFN NGEINF 534- 579

[00304] FokI, KKR Sharkey P478D(E->K @490, I->K@538, H->R@537, P478D, SEQ ID N°:170)

[00305] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
LEMKVMEFFM 384- 433 KVGYRGEHL GGSRKPDGAI

YTVGSPIDYG VIVDTKAYSG GYNLDIGQAD 434- 483

EMQRYVKENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH

FKGNYKAQLT 484- 533 RLNRKTNCNG AVLSVEELLI

GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00306] FokI DAD (R->D@487, N->D@496, I->A@499, SEQ ID Nº:171)
 [00307] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQDYVEENQ TRDKHANPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00308] FokI DAD Sharkey (R->D@487, N->D@496, I->A@499,
 S418P+K441E, SEQ ID Nº:172)
 [00309] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAD 434- 483
 EMQDYVEENQ TRDKHANPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNHITNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00310] FokI RVR (D->R@483, H->R@537, I->V@538, SEQ ID
 Nº:173)
 [00311] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNSTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGKHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAR 434- 483
 EMQRYVEENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
 FKGNYKAQLT 484- 533 RLNRVTNCNG AVLSVEELLI
 GGEMIKAGTL TLEEVRRKFN NGEINF 534- 579
 [00312] FokI RVR Sharkey (D->R@483, H->R@537, I->V@538,
 S418P+K441E, SEQ ID Nº:174)
 [00313] QLVKSELEEK KSELRHKLKY VPHEYIELIE IARNPTQDRI
 LEMKVMEFFM 384- 433 KVYGYRGEHL GGSRKPDGAI
 YTVGSPIDYG VIVDTKAYSG GYNLPIGQAR 434- 483

EMQRYVEENQ TRNKHINPNE WWKVYPSSVT EFKFLFVSGH
FKGNYKAQLT 484- 533 RLNRVTNCNG AVLSVEELLI
GGEMIKAGTL TLEEVRRKFN NGEINE 534- 579

[00314] Todas as combinações em pares de ZFNs foram testadas quanto à funcionalidade e todas demonstraram ser ativas.

Tabela 6: Pares de ZFN específicos para TCRA

55248 5'agGATTGGAA ACCAATCACt gacaggt (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLSR (SEQ ID Nº:53)	L0
	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	KKR Fok C-term
68797 agGATTGGAA CCCAATCACtga caggt (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLSR (SEQ ID Nº:53)	L0
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR Fok C-term
68813 agGATTGGAA CCCAATCACtga caggt (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLSR (SEQ ID Nº:53)	L0
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR R416E Fok C-term
68861 agGATTGGAA CCCAATCACtga caggt (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLSR (SEQ ID Nº:53)	L0
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR Q481E Fok C-term
68877 agGATTGGAA CCCAATCACtga caggt (SEQ ID Nº:14)	DQSNLRA (SEQ ID Nº:42)	TSSNRKT (SEQ ID Nº:43)	LQQTLAD (SEQ ID Nº:51)	QSGNLAR (SEQ ID Nº:34)	RREDLIT (SEQ ID Nº:52)	TSSNLSR (SEQ ID Nº:53)	L0
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR K525S Fok C-term
Sítio E							
Parceira esquerda							
55266 tcAAGCTGGTC GAGaAAAGCTt gaaac (SEQ ID Nº:15)	QSSDLSR (SEQ ID Nº:57)	QSGNRTT (SEQ ID Nº:58)	RSANLAR (SEQ ID Nº:59)	DRSALAR (SEQ ID Nº:33)	RSDVLSE (SEQ ID Nº:60)	KHSTRRV (SEQ ID Nº:61)	N7c
	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	ELD Fok N-term
68798 tcAAGCTGGTC GAGaAAAGCTt gaaac (SEQ ID Nº:15)	QSSDLSR (SEQ ID Nº:57)	QSGNRTT (SEQ ID Nº:58)	RSANLAR (SEQ ID Nº:59)	DRSALAR (SEQ ID Nº:33)	RSDVLSE (SEQ ID Nº:60)	KHSTRRV (SEQ ID Nº:61)	N7c
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	ELD Fok N-term
68846 tcAAGCTGGTC GAGaAAAGCTt gaaac (SEQ ID Nº:15)	QSSDLSR (SEQ ID Nº:57)	QSGNRTT (SEQ ID Nº:58)	RSANLAR (SEQ ID Nº:59)	DRSALAR (SEQ ID Nº:33)	RSDVLSE (SEQ ID Nº:60)	KHSTRRV (SEQ ID Nº:61)	N7c
	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	ELD I479T Fok N-term

Parceira direita							
	TMHQRV E (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHLTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNLDA (SEQ ID Nº:67)	L0
53853 aaCAGGTAAaGA CAGGGGTCTA gcctggg (SEQ ID Nº:16)	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	nenhuma	KKR Fok C-term
	TMHQRV E (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHLTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNLDA (SEQ ID Nº:67)	L0
68879 aaCAGGTAAaGA CAGGGGTCTA gcctggg (SEQ ID Nº:16)	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR K525S Fok C-term
	TMHQRV E (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHLTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNLDA (SEQ ID Nº:67)	L0
68815 aaCAGGTAAaGA CAGGGGTCTA gcctggg (SEQ ID Nº:16)	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR R416E Fok C-term
	TMHQRV E (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHLTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNLDA (SEQ ID Nº:67)	L0
68799 aaCAGGTAAaGA CAGGGGTCTA gcctggg (SEQ ID Nº:16)	nQm5	nenhuma	nQm5	nenhuma	nQm5	nenhuma	KKR Fok C-term
	TMHQRV E (SEQ ID Nº:62)	TSGHLSR (SEQ ID Nº:63)	RSDHLTQ (SEQ ID Nº:64)	DSANLSR (SEQ ID Nº:65)	QSGSLTR (SEQ ID Nº:66)	AKWNLDA (SEQ ID Nº:67)	L0

[00315] Os genes que codificam as ZFNs para cada sítio foram clonados em um plasmídeo de expressão como parceira direta e esquerda, separadas por um peptídeo de autoclivagem 2A em combinações para cada sítio alvo. Os mRNA que codificam as ZFNs foram derivados por meio de métodos padrão de transcrição *in vitro*. Células T ativadas (3 dias após a ativação) foram então tratadas com os vários mRNAs em 3 doses diferentes (12, 6 ou 3 µg in 100 µL, 3E6 células T) por eletroporação. 4 dias após a eletroporação, as células foram analisadas quanto à clivagem nos sítios alvo e no sítio alvo. Os dados são apresentados abaixo nas duas tabelas (uma para cada sítio alvo).

Tabela 7a: Clivagem no alvo e fora do alvo no Sítio D

	Sítio D	55254-2A-55248	6879-6-2A-68813	6881-3-2A-68796	68796-2A-68861	68861-2A-68796	68812-2A-68813	68813-2A-68812	68876-2A-68877	68877-2A-68876	Controle
No alvo	12 ug	96,7	99,3	98,8	99,4	99,3	99,9	99,9	99,2	99,1	0,12
	6 ug	98,5	99,2	99,1	99,4	99,4	99	98,9	99,3	99,2	0,14
	3 ug	96	99,1	98,8	99,3	98,9	98,3	97,8	98,7	99,3	0,15
Fora D1	12 ug	39,6	0,29	0,35	0,21	0,18	0,25	0,25	0,2	0,2	0,28
	6 ug	18	0,25	0,3	0,25	0,2	0,28	0,22	0,29	0,23	0,34
	3 ug	7,3	0,28	0,24	0,46	0,26	0,24	0,27	0,22	0,25	0,26
12 µg	Fora soma	42,22	1,67	1,53	3,17	1,19	1,46	1,74	1,33	2,05	1,28
	Dentro/ fora	2,3	59	65	31	84	68	57	75	48	0,09
6 µg	Fora soma	19,14	5,94	1,53	1,53	1,06	1,28	1,36	1,3	1,32	1,22
	Dentro/ fora	5,1	17	65	65	94	77	73	76	75	0,12
3 µg	Fora soma	8,28	4,3	1,18	1,56	1,29	1,22	1,54	1,21	8,13	1,2
	Dentro/ fora	12	23	83	63	77	81	63	82	12	0,13
Soma fora		69,63	11,91	4,23	6,26	3,54	3,96	4,65	3,84	11,5	3,7
Dentro/fora média		6,3	33	71	53	85	75	65	77	45	0,11

Tabela 7b: Clivagem no alvo e fora do alvo no Sítio E

	Sítio E	55266-2A-53853	55266-2A-68815	68815-2A-55266	55266-2A-68879	68879-2A-55266	68798-2A-68815	68815-2A-68798	68846-2A-53853	53853-2A-68846	Controle do Sítio E
No alvo	12 ug	96,7	97,6	86,5	96,3	95,5	97,5	96,4	96,5	97	0,19
	6 ug	95,3	94,6	81,3	95,2	91,5	96,5	97,2	94,4	NA	0,34
	3 ug	95,3	NA	0,17							
Fora E1	12 ug	1,24	0,32	0,23	0,24	0,3	0,23	0,27	0,19	0,21	0,29
	6 ug	0,79	0,23	0,24	0,27	0,25	0,22	0,22	0,18	0,23	0,25
	3 ug	0,5	0,26	0,18	0,2	0,23	0,2	0,23	0,23	0,23	0,26
Fora E2	12 ug	19,69	1,05	0,51	0,95	1,04	0,37	0,36	0,18	0,23	0,24
	6 ug	11,09	NA	0,34	0,67	0,69	0,31	0,26	0,17	0,22	0,17
	3 ug	4,05	0,36	0,28	0,34	0,33	0,24	0,26	0,23	0,22	0,13
Fora E3	12 ug	4,32	0,14	0,19	0,4	0,19	0,17	0,19	0,18	0,16	0,19
	6 ug	1,33	0,13	0,13	0,21	0,17	0,19	0,14	0,11	0,19	0,21
	3 ug	0,47	0,13	0,15	0,2	0,18	0,14	0,15	0,12	0,14	0,14
12 µg	Fora soma	25,24	1,51	0,93	1,59	1,53	0,77	0,82	0,54	0,6	0,71
	Dentro/ fora	3,8	65	93	61	65	127	117	177	161	0,27
6 µg	Fora soma	13,21	0,36	0,72	1,15	1,11	0,72	0,61	0,46	0,64	0,63
	Dentro/ fora	7,2	261	113	83	82	135	160	204	NA	0,54
3 µg	Fora soma	5,02	0,74	0,61	0,74	0,74	0,57	0,64	0,58	0,55	0,52
	Dentro/ fora	18,98	NA	0,32							
Soma fora		43,47	2,62	2,26	3,48	3,38	2,06	2,07	1,59	1,79	1,86
Dentro/fora média		10	163	103	72	72	131	139	191	161	0,38

[00316] Assim, após as modificações, os reagentes ZFN mantiveram a excelente atividade de corte no alvo, frequentemente enquanto diminuindo a atividade fora do alvo ao valor de fundo (compare, por exemplo, a atividade de clivagem no alvo do par parental 55254/55248 com o par modificado 68861/68796, mostrando 96,7 e 99,3 por cento de clivagem no alvo nas doses saturantes de 12 µg, respectivamente, enquanto também mantendo uma atividade total fora do alvo nessa dose de 42,22 por cento, par original, e 1,19% no par modificado – semelhante ao nível controle de 1,28.

[00317] Quanto às ZFNs para TRAC ZFNs: os aminoácidos em contato potencial com fosfato foram modificados no domínio FokI das proteínas B2M. Modificações exemplares dos componentes ZFP ("desenhos") são mostradas abaixo na Tabela 8.

Tabela 8: Otimização de ZFNs específicas para B2M

SBS no (sítio alvo, 5'-3')	Desenho						Ligante	
	[Seqüência de hélice, SEQ ID]							
	[Mutações ao esqueleto do dedo]							
	F1	F2	F3	F4	F5	F6		
SBS57531 5' gaGTAGCGcGAGCACAGC taaggccacg (SEQ ID Nº:126)	AQCCLFH (SEQ ID Nº:128)	DQSNLRA (SEQ ID Nº:42)	RSANLTR (SEQ ID Nº:129)	RSDDLTR (SEQ ID Nº:130)	QSGSLTR (SEQ ID Nº:66)	N/A	N6a KKR Fok N-term	
SBS68957 5' gaGTAGCGcGAGCACAGC taaggccacg (SEQ ID Nº:126)	AQCCLFH (SEQ ID Nº:128)	DQSNLRA (SEQ ID Nº:42)	RSANLTR (SEQ ID Nº:129)	RSDDLTR (SEQ ID Nº:130)	QSGSLTR (SEQ ID Nº:66)	N/A	N6a	
	nenhuma	nenhuma	Nenhuma	nenhuma	nenhuma	N/A	KKR K525S Fok N-term	
SBS72678 5' gaGTAGCGcGAGCACAGC taaggccacg (SEQ ID Nº:126)	AQCCLFH (SEQ ID Nº:128)	DQSNLRA (SEQ ID Nº:42)	RSANLTR (SEQ ID Nº:129)	RSDDLTR (SEQ ID Nº:130)	QSGSLTR (SEQ ID Nº:66)	N/A	N6a	
	nenhuma	nenhuma	Nenhuma	nenhuma	nenhuma	N/A	KKR R416Y Fok N-term	
SBS57071 gcCACGGAgCGAGACATC TCGgcoccaa (SEQ ID Nº:127)	RSDDLSK (SEQ ID Nº:131)	DSSARKK (SEQ ID Nº:132)	DRSNLSR (SEQ ID Nº:22)	QRTHLRD (SEQ ID Nº:133)	QSGHLAR (SEQ ID Nº:29)	DSSNREA (SEQ ID Nº:134)	L0 ELD Fok C-term	
SBS72732 gcCACGGAgCGAGACATC TCGgcoccaa (SEQ ID Nº:127)	RSDDLSK (SEQ ID Nº:131)	DSSARKK (SEQ ID Nº:132)	DRSNLSR (SEQ ID Nº:22)	QRTHLRD (SEQ ID Nº:133)	QSGHLAR (SEQ ID Nº:29)	DSSNREA (SEQ ID Nº:134)	L0	
	nenhuma	nenhuma	Nenhuma	nenhuma	nenhuma	nenhuma	ELD P478D Fok C-term	
SBS72748 gcCACGGAgCGAGACATC TCGgcoccaa (SEQ ID Nº:127)	RSDDLSK (SEQ ID Nº:131)	DSSARKK (SEQ ID Nº:132)	DRSNLSR (SEQ ID Nº:22)	QRTHLRD (SEQ ID Nº:133)	QSGHLAR (SEQ ID Nº:29)	DSSNREA (SEQ ID Nº:134)	L0	
	nenhuma	nenhuma	Nenhuma	nenhuma	nenhuma	N/A	ELD Q481D Fok C-term	

[00318] Os reagentes B2M modificados foram testados quanto à atividade como acima e foram analisados para o *knockout* fenotípico por análise FACs com um anticorpo específico para HLA. Todas as combinações em pares (57531/57071; 57531/72732; 57531/72748; 68957/57071; 68957/72732; 68957/72748; 72678/57071; 72678/72732; 72678/72748) mostraram ser ativas com resultados exemplares para os pares indicados mostrados abaixo na Tabela 9 e demonstram que as variantes modificadas são ativas.

Tabela 9: Análise fenotípica de ZFN específica para B2M

Par de ZFNs (2A mRNA)	Concentração de ZFN (µg/mL)			
	30	60	90	120
% Indels				
57071/68957	74	79	83	81
72732/57531	83	86	87	85
72732/72678	86	nt	nt	87
72748/68957	37	nt	nt	80

nt: não testado.

[00319] Análises dentro e fora do alvo foram também realizadas utilizando MiSeq para cada um dos pares listados acima na Tabela 9. Os resultados são mostrados abaixo para cada par nas Tabelas 10A-10D, e demonstram que esses reagentes são altamente específicos.

Tabela 10A: Análise de fora do alvo para o par 57071/68957

	ZFP		GFP			
	corrigido	bruto	corrigido	bruto	valor p	curadoria
57071/68957						
Alvo	91,64	91,94	0,19	0,25	0,00	positiva
OT1	0,08	0,39	0,04	0,35	0,12	negativa
OT2	0,03	0,33	0,01	0,24	0,06	negativa
OT3	0,08	1,22	0,03	1,00	0,05	negativa
OT4	0,02	0,16	0,03	0,14	1,00	negativa
OT5	0,04	0,48	0,02	0,41	1,00	negativa
OT6	0,04	0,27	0,03	0,22	1,00	talvez
OT7	nt	nt	nt	nt	nt	nt
OT8	0,02	0,18	0,02	0,13	1,00	negativa
OT9	0,04	0,72	0,06	0,58	1,00	negativa
OT10	0,03	0,15	0,03	0,12	1,00	negativa

Tabela 10B: Análise de fora do alvo para o par 72732/57531

	ZFP		GFP			
72732/57531	corrigido	bruto	corrigido	bruto	valor p	curadoria
Alvo	95,75	96,88	0,25	0,31	0,00	positiva
OT1	0,03	0,26	0,02	0,26	1,00	negativa
OT2	0,08	0,52	0,06	0,41	1,00	negativa
OT3	0,06	0,19	0,05	0,21	1,00	negativa
OT4	0,06	0,47	0,04	0,40	1,00	negativa
OT5	0,03	0,19	0,02	0,19	1,00	negativa
OT6	0,02	0,77	0,02	0,84	1,00	negativa
OT7	0,04	0,98	0,06	0,79	1,00	negativa
OT8	0,07	7,42	0,07	7,45	1,00	negativa
OT9	0,02	0,14	0,02	0,16	1,00	negativa
OT10	0,03	0,27	0,03	0,28	1,00	negativa

Tabela 10C: Análise de fora do alvo para o par 72732/72678

	ZFP		GFP			
72732/72678	corrigido	bruto	corrigido	bruto	valor p	curadoria
Alvo	94,76	95,23	0,17	0,21	0,00	positiva
OT1	0,09	0,48	0,02	0,36	0,00	negativa
OT2	0,05	0,37	0,02	0,39	0,43	talvez
OT3	0,03	0,28	0,03	0,19	1,00	negativa
OT4	0,02	0,18	0,01	0,15	1,00	negativa
OT5	0,01	0,09	0,03	0,11	1,00	negativa
OT6	0,09	0,42	0,03	0,41	0,00	negativa
OT7	1,02	17,40	2,35	19,23	1,00	negativa
OT8	0,07	0,71	0,04	0,58	1,00	negativa
OT9	0,02	0,21	0,05	0,20	1,00	negativa
OT10	0,03	0,25	0,02	0,18	1,00	negativa

Tabela 10D: Análise de fora do alvo para o par 72748/68957

	ZFP		GFP			
72748/68957	corrigido	bruto	corrigido	bruto	valor p	curadoria
Alvo	93,39	93,50	0,16	0,20	0,00	positiva
OT1	0,05	0,30	0,02	0,24	0,69	negativa
OT2	0,02	0,14	0,02	0,14	1,00	negativa
OT3	0,05	2,24	0,04	2,29	1,00	negativa
OT4	0,02	0,33	0,03	0,31	1,00	negativa
OT5	0,05	7,57	0,07	7,21	1,00	negativa
OT6	0,03	1,03	0,03	1,03	1,00	negativa
OT7	0,76	1,86	0,59	1,79	1,00	negativa
OT8	0,02	0,14	0,02	0,13	1,00	negativa
OT9	0,03	0,23	0,03	0,29	1,00	negativa
OT10	0,33	94,52	0,29	94,49	1,00	negativa

[00320] As ZFNs modificadas específicas para TRAC e B2M foram testadas em combinação e avaliadas quanto à eficiência do *knockout*, tanto por análise Miseq como por análise fenotípica examinando a quantidade de células CD3+ ou HLA+ por análise FACs. A análise foi realizada em células T, utilizando duas concentrações diferentes do mRNA codificador de ZFN adicionado (90 µg/mL ou 120 µg/mL). Os resultados são mostrados abaixo na Tabela 11 e demonstram que esses reagentes são altamente eficientes.

Tabela 11: Clivagem de TRAC/B2M

Reagentes ZFN (2A-mRNAs)		Triagem fenotípica		Análise Miseq	
68846-2A-53853 (TRAC) µg/mL	72732-2A-72678 (B2M) µg/mL	% de CD3-neg	% de HLA-I-neg	% de TRAC indels	% de B2M indels
0	30	-	86	-	95
60	0	98	-	92	-
90	90	95	86	90	95
120	90	94	86	90	94
90	120	94	86	90	95
120	120	95	87	91	95

[00321] Os reagentes foram também testados em combinação na presença ou ausência de uma construção doadora de GFP direcionada por um promotor PGK. Os resultados são mostrados na Tabela 12, em que a inserção foi efetuada no lócus clivado de B2M ou TRAC. Em cada caso, o doador PGK-GFP foi entregue por AAV6 e compreendia braços de homologia com homologia flanqueando os sítios de corte de TRAC ou B2M. O par de ZFNs específico usado para construção de TRAC foi 68846-2A-53853 enquanto a construção para o par específico para B2M foi 72732-2A-72678.

Tabela 12: Atividade de *knockout* duplo em dois doadores de células T

Doador de células T nº 1			Doador de células T nº 2		
Amostra	Lócus visado	% indel	Amostra	Lócus visado	% indel
Simulação	B2M	0,3	Simulação	B2M	0,04
TRAC + B2M	B2M	84,14	TRAC + B2M	B2M	75,33
TRAC + B2M PGK-GFP	B2M	83,55	TRAC + B2M PGK-GFP	B2M	80,96
Simulação	TRAC	0,08	Simulação	TRAC	0,38
TRAC + B2M	TRAC	88,05	TRAC + B2M	TRAC	85,09
TRAC + B2M PGK-GFP	TRAC	78,94	TRAC + B2M PGK-GFP	TRAC	74,54

[00322] Assim, os pares otimizados de ZFNs específicas para B2M foram construídos, escolhendo-se uma variante de FokI (ver acima) em combinação com um domínio de ligação ao DNA de ZFP.

[00323] As sequências de aminoácidos otimizadas para o domínio de ligação ao DNA das ZFNs para B2M, 72732 e 72678, são mostradas abaixo:

[00324] 72732 N term:

[00325] RPFQCRICMRNFSRSDDL SKHIRTHTGEKPFACDICGRKF
ADSSARKKKHTKIHTGEKPFQCRICMRNFSDRSNLSRHIRTHTGEKPF
ACDICGRKFAQRTHLRDHTKIHTHPRAPIPKPFQCRICMRNFSQLGHL
ARHIRTHTGEKPFACDICGRKFADSSNREAHTKIH (SEQ ID Nº:175)

[00326] 72678 C-term:

[00327] RPFQCRICMRKFAAQCCLFHHTKIHTGEKPFQCRICMRNF
SDQSNLRAHIRHTGEKPFACDICGRKFARSANLTRHTKIHTHPRAPI
PKPFQCRICMRNFSSRSDDLTRHIRTHTGEKPFACDICGRKFAQSGSL
TRHTKIH (SEQ ID Nº:176)

[00328] ZFNs adicionais compreendendo as ZFPs modificadas das ZFNs aqui descritas (por exemplo, SEQ ID Nº:175 e SEQ ID Nº:176)

são também geradas utilizando domínios FokI e/ou ligante diferentes.

[00329] Do mesmo modo, os pares otimizados de ZFNs específicas para TRAC foram construídos, escolhendo-se uma variante de FokI (ver, por exemplo, acima) em combinação com um domínio de ligação ao DNA de ZFP. As sequências de aminoácidos otimizadas para o domínio de ligação ao DNA das ZFNs para B2M, 68846 e 53853, são mostradas abaixo:

[00330] 68846 C-term:

[00331] RPFQCRICMQNFSQSSDLSRHIRTHTGEKPFACDICGRKF
AQSGNRTTHTKIHTHPRAPIPKPFQCRICMQNFSRSANLARHIRTHTG
EKPFACDICGRKFADRSALARHTKIHTGSQKPFQCRICMQNFSRSDV
LSEHIRTHTGEKPFACDICGRKFAKHSTRRVHTKIH (SEQ ID Nº:177)

[00332] 53853 N-term:

[00333] RPFQCRICMRNFSTMHQRVEHIRTHTGEKPFACDICGRK
FATSGHLSRHTKIHTGSQKPFQCRICMRNFSRSDHLTQHIRTHTGEK
PFACDICGRKFADSANLSRHTKIHTHPRAPIPKPFQCRICMRNFSQSG
SLTRHIRTHTGEKPFACDICGRKFAAKWNLDAHTKIH SEQ ID
Nº:178).

[00334] As ZFNs podem ser montadas com o domínio de ligação ao DNA no N-terminal em relação ao domínio FokI, em que a sequência ligadora entre o domínio de ligação ao DNA e o domínio FokI foi o ligante L0: LRGS. Alternativamente, se a ZFN é montada de tal modo que o domínio FokI seja N-terminal em relação ao domínio de ligação ao DNA, o ligante usado foi o ligante N7c: SGAIRCHDEFWF (SEQ ID Nº:179).

[00335] Características adicionais foram adicionadas às construções incluindo 3x TAG FLAG na região do N-terminal (DYKDHDGDYKDHDIDYKDDDDK, SEQ ID Nº:180), e uma sequência de localização nuclear (PKKKRKV, SEQ ID Nº:181).

[00336] Além disso, em algumas construções, sequências que

codificam o par de ZFNs de interesse são unidas em uma sequência de DNA, onde as fases de leitura aberta para cada ZFN parceira são separadas por uma sequência 2A. Tal sequência de DNA, de 68846-2A-53853, é mostrada abaixo:

[00337] 5'ATGGACTACAAAGACCATGACGGTGATTATAAAGATC
ATGACATCGATTACAAGGATGACGATGACAAGATGGCCCCAAGA
AGAAGAGGAAGGTGGCATCCACGGGTACCCGCCGCTATGGGA
CAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCG
GCACAAGCTGAAGTACGTCCCCACGAGTACATCGAGCTGATCG
AGATCGCCAGGAACAGCACCCAGGACCGCATCCTGGAGATGAAG
GTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGAAAGCAC
CTGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGG
CAGCCCCATCGATTACGGCGTGATCGTGGACACAAAGGCCTACA
GCGGCGGCTACAATCTGCCTACCGGCCAGGCCGACGAGATGGAG
AGATACGTGGAGGAGAACCAAGACCCGGATAAGCACCTAACCC
CAACGAGTGGTGGAAAGGTGTACCCTAGCAGCGTGACCGAGTTCA
AGTTCCCTGTTCGTGAGCGGCCACTTCAAGGGCAACTACAAGGCC
AGCTGACCAGGCTGAACCACATCACCAACTGCAATGGCGCCGTG
CTGAGCGTGGAGGAGCTGCTGATGGCGGCGAGATGATCAAAGC
CGGCACCCCTGACACTGGAGGAGGTGGCGCAAGTTCAACAAACG
GCGAGATCAACTTCAGCGGCCATCAGATGCCACGACGAGTTCA
TGGTTCAGGCCCTCCAGTGTGAATCTGCATGCAGAACTTCAGT
CAGTCCTCCGACCTGTCCGCCACATCCGCACCCACACCGCGA
GAAGCCTTTGCCTGTGACATTGTGGAGGAAATTGCCAGTC
CGGCAACCGCACCAACCATACCAAGATAACACACGCATCCCAGGG
CACCTATTCCAAGCCCTCCAGTGTGAATCTGCATGCAGAACT
TCAGTCGCTCCGCCAACCTGGCCGCCACATCCGCACCCACACC
GGCGAGAAGCCTTTGCCTGTGACATTGTGGAGGAAATTGCC
GACCGCTCCGCCCTGGCCGCCATACCAAGATAACACACGGGATC
TCAGAAGCCCTCCAGTGTGAATCTGCATGCAGAACTTCAGTCG

CTCCGACGTGCTGTCCGAGCACATCCGCACCCACACC GGCGAGA
AGCCTTTGCCTGTGACATTGTGGGAGGAATTGCCAAGCACT
CCACCCGCCGCGTGCATAACCAAGATAACACCTGCGGCAGAAGGAC
AGATCTGGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACCTGCGG
TGACGTGGAGGAGAATCCC GCCCTAGGACCATGGACTACAAAG
ACCATGACGGTGATTATAAGATCATGACATCGATTACAAGGATGA
CGATGACAAGATGGCCCCAAGAAGAAGAGGAAGGT CGGCATTC
ATGGGGTACCCGCCGCTATGGCTGAGAGGCCCTCCAGT GTCGA
ATCTGCATGCGTAAC TTCA GTACCATGCACCAGCGCGTGGAGCAC
ATCCGCACCCACACCGGCGAGAAGCCTTGCCTGTGACATTGT
GGGAGGAAATTGCCACCTCCGGCCACCTGTCCC GCCATACCAA
GATACACACGGGCAGCCAAAAGCCCTTCCAGT GTCGAATCTGCAT
GCGTAAC TTCA GTCGCTCCGACCACCTGACCCAGCACATCCGCAC
CCACACCGGCGAGAAGCCTTGCCTGTGACATTGTGGAGGAA
ATTGCCGACTCCGCCAACCTGTCCC CCATACCAAGATAACACAC
GCACCCCGCGCGCCCCGATCCCGAAGCCCTTCCAGT GTCGAATCT
GCATGCGTAAC TTCA GTCA GTCCGGCTCCCTGACCCGCCACATCC
GCACCCACACCGGCGAGAAGCCTTGCCTGTGACATTGTGGGA
GGAAATTGCCCCAAGTGGAACCTGGACGCCATACCAAGATAAC
ACCTGCGGGGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAA
GAAGTCCGAGCTGCGGCACAAGCTGAAGTACGTGCCCCACGAGT
ACATCGAGCTGATCGAGATGCCAGGAACAGCACCCAGGACCGC
ATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGC
TACAGGGAAAGCACCTGGCGGAAGCAGAAAGCCTGACGGCG
CCATCTATACAGTGGCAGCCCCATCGATTACGGCGTACGTGG
ACACAAAGGCCTACAGCGCGGCTACAATCTGCCTATGCCAG
GCCGACGAGATGCAGAGATA CGTGAAGGAGAAC CAGACCCGGAA
TAAGCACATCAACCCCAACGAGTGGTGGAGGTGTACCC TAGCAG
CGTGACCGAGTTCAAGTTCTGTTGTGAGCGGCCACTTCAAGGG
CAACTACAAGGCCAGCTGACCAGGCTGAACCGAAAACCAACT

GCAATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGC
 CGAGATGATCAAAGCCGGCACCTGACACTGGAGGAGGTGCG
 GCAAGTTCAACAAACGGCGAGATCAACTTCTGATAA (SEQ ID
 N°:182).

[00338] A sequência de aminoácidos do quadro de leitura aberta de 68846-2A-53853 é:

[00339] MDYKDHDGDY KDHIDYKDD DDKMAPKKR
KVGIHGVPA MGQLVKSELE EKKSELRHKL 1-60 KYVPHEYIEL
IEIARNSTQD RILEMKVMEF FMKVYGYRGK HLGGSRKPDG
AIYTVGSPID 61-120 YGVIVDTKAY SGGYNLPTGQ ADEMERYVEE
NQTRDKHLNP NEWWKVYPSS VTEFKFLFVS 121-180
GHFKGNYKAQ LTRLNHITNC NGAVLSVEEL LIGGEMIKAG
TLTLEEVRRK FNNGEINFSG 181-240 AIRCHDEFWF RPFQCRICMQ
NFSQSSDLSR HIRHTGEKP FACDICGRKF AQSGNRTTHT 241-300
KIHTHPRAPI PKPFQCRICM QNFSRSANLA RHIRHTGEK
PFACDICGRK FADRSALARH 301-360 TKIHTGSQKP FQCRICMQNF
SRSDVLSEHI RTHTGEKPFA CDICGRKFAK HSTRRVHTKI 361-420
HLRQKDRSGG GEGRGSLLTC GDVEENPGPR TMDYKDHDGD
YKDHDIDYKD DDDKMAPKK 421-480 RKVGIHGVPA
AMAERPFQCR ICMRNFSTMH QRVEHIRHT GEKPFACDIC
GRKFATSGHL 481-540 SRHTKIHTGS QKPFQCRICM RNFSRSDHLT
QHIRHTGEK PFACDICGRK FADSANLSRH 541-600 TKIHTHPRAP
IPKPFQCRIC MRNFSQSGSL TRHIRHTGE KPFACDICGR
KFAAKWNLDA 601-660 HTKIHLRGSQ LVKSELEEK
SELRHKLKYV PHEYIELIEI ARNSTQDRIL EMKVMEFFMK 661-720
VGYRGKHLG GSRKPDGAIY TVGSPIDYGV IVDTKAYS
YNLPIGQADE MQRYVKENQT 721-780 RNKHINPNEW
WKVYPSSVTE FKFLFVSGHF KGNYKAQLTR LNRKTNCNGA
VLSVEELLIG 781-840 GEMIKAGTLT LEEVRRKFNN GEINF (SEQ ID
 N°:183) 841-865

[00340] As características desse polipeptídeo são detalhadas abaixo na Tabela 13.

Tabela 13: Características da sequência peptídica de 68846-2A-53853

Característica	Designação	Localização (dentro da SEQ ID Nº:183)
Sequência 3x FLAG	<u>xx</u>	2-23
Sequência de localização nuclear	<u>xx</u>	26-32
Domínio ELD I479T Fok1	xx	43- 238
Ligante N7c	<u>xx</u>	239-250
Domínio de ligação ao DNA 68846	<u>xx</u>	251-421
Ligante 2A	<u>xx</u>	432- 449
Sequência 3x FLAG	<u>xx</u>	452- 474
Sequência de localização nuclear	<u>xx</u>	477- 483
Domínio de ligação ao DNA 53853	<u>xx</u>	495- 665
Ligante L0	<u>xx</u>	666-669
Domínio KKR FokI	xx	670-865

[00341] A sequência do quadro de leitura aberta de 72732-2A-72678 é mostrada abaixo:

[00342]

ATGGACTACA	AAGACCATGA	CGGTGATTAT
AAAGATCATG	ACATCGATTA	CAAGGATGAC
TGGCCCCCAA	GAAGAACAGAGG	AAGGTCGGCA
ACCCGCCGCT	ATGGCTGAGA	GGCCCTTCCA
TGCATGCGTA	ACTTCAGTCG	TAGTGACGAC
ACATCCGCAC	CCACACAGGC	GAGAACCTT
CATTGTGGG	AGGAAATTG	CCGACAGCAG
AAGCATACCA	AGATACACAC	GGCGAGAACAG
GTCGAATCTG	CATGCGTAAC	TTCAGTGACC
GTCCCGCCAC	ATCCGCACCC	ACACCGGCGA
GCCTGTGACA	TTTGTGGGAG	GAAGCCTTT
		CAGCGCACCC

ACCTGCGCGA	CCATACCAAG	ATACACACGC	ACCCGCGCGC
CCCGATCCCC	AAGCCCTTCC	AGTGTGGAAT	CTGCATGCGT
AACTTCAGTC	AGTCCGGCCA	CCTGGCCCGC	CACATCCGCA
CCCACACCGG	CGAGAAGCCT	TTTGCCTGTG	ACATTGTGG
GAGGAAATT	GCCGACTCCT	CCAACCAGCGA	GGCCCATAACC
AAGATACACC	TGCGGGGATC	CCAGCTGGTG	AAGAGCGAGC
TGGAGGAGAA	GAAGTCCGAG	CTGCGGCACA	AGCTGAAGTA
CGTGCCCCAC	GAGTACATCG	AGCTGATCGA	GATGCCAGG
AACAGCACCC	AGGACCGCAT	CCTGGAGATG	AAGGTGATGG
AGTTCTTCAT	GAAGGTGTAC	GGCTACAGGG	GAAAGCACCT
GGGCGGAAGC	AGAAAGCCTG	ACGGCGCCAT	CTATACAGTG
GGCAGCCCCA	TCGATTACGG	CGTGATCGT	GACACAAAGG
CCTACAGCGG	CGGCTACAAT	CTGGACATCG	GCCAGGCCGA
CGAGATGGAG	AGATACGTGG	AGGAGAACCA	GACCCGGGAT
AAGCACCTCA	ACCCCCAACGA	GTGGTGGAAAG	GTGTACCCCTA
GCAGCGTGAC	CGAGTTCAAG	TTCCTGTTCG	TGAGCGGCCA
CTTCAAGGGC	AACTACAAGG	CCCAGCTGAC	CAGGCTGAAC
CACATCACCA	ACTGCAATGG	CGCCGTGCTG	AGCGTGGAGG
AGCTGCTGAT	CGGCGGCGAG	ATGATCAAAG	CCGGCACCCCT
GACACTGGAG	GAGGTGCGGC	GCAAGTTCAA	CAACGGCGAG
ATCAAACCTCA	GATCTGGCGG	CGGAGAGGGC	AGAGGAAGTC
TTCTAACCTG	CGGTGACGTG	GAGGAGAAC	CCGGCCCTAG
GACCATGGAC	TACAAAGACC	ATGACGGTGA	TTATAAAGAT
CATGACATCG	ATTACAAGGA	TGACGATGAC	AAGATGGCCC
CCAAGAAGAA	GAGGAAGGTC	GGCATTATG	GGGTACCCGC
CGCTATGGGA	CAGCTGGTGA	AGAGCGAGCT	GGAGGAGAAG
AAGTCCGAGC	TGCAGGCACAA	GCTGAAGTAC	GTGCCACCG
AGTACATCGA	GCTGATCGAG	ATCGCCTACA	ACAGCACCCA
GGACCGCATC	CTGGAGATGA	AGGTGATGGA	GTTCTTCATG
AAGGTGTACG	GCTACAGGGG	AAAGCACCTG	GGCGGAAGCA

GAAAGCCTGA CGGCGCCATC TATACAGTGG GCAGCCCCAT
 CGATTACGGC GTGATCGTGG ACACAAAGGC CTACAGCGGC
 GGCTACAATC TGCCTATCGG CCAGGCCGAC GAGATGCAGA
 GATACGTGAA GGAGAACCAAG ACCCGGAATA AGCACATCAA
 CCCCAACGAG TGGTGGAAGG TGTACCCTAG CAGCGTGACC
 GAGTTCAAGT TCCTGTTCGT GAGCGGCCAC TTCAAGGGCA
 ACTACAAGGC CCAGCTGACC AGGCTGAACC GCAAAACCAA
 CTGCAATGGC GCCGTGCTGA GCGTGGAGGA GCTGCTGATC
 GGC GGCGAGA TGATCAAAGC CGGCACCCCTG ACACTGGAGG
 AGGTGCGGCG CAAGTTCAAC AACGGCGAGA TCAACTTCAG
 CGGCGCTCAG GGATCTACCC TGGACTTTAG GCCCTTCCAG
 TGTCGAATCT GCATGCGTAA GTT GCGGCC CAGTGTGTC
 TGTTCCACCA TACCAAGATA CACACGGGCG AGAAGCCCTT
 CCAGTGTGCA ATCTGCATGC GTA ACTTCAG TGACCAGTCC
 AACCTGCGCG CCCACATCCG CACCCACACC GGCGAGAAGC
 CTTTGCGCTG TGACATTGT GGGAGGAAAT TTGCCC GCTC
 CGCCAACCTG ACCCGCCATA CCAAGATA CACGCACCCG
 CGCGCCCCGA TCCCGAAGCC CTTCCAGTGT CGAATCTGCA
 TGCGTAACTT CAGTCGCTCC GACGACCTGA CCCGCCACAT
 CCGCACCCAC ACCGGCGAGA AGCCTTTGC CTGTGACATT
 TGTGGGAGGA AATTGCCC A GTCCGGCTCC CTGACCCGCC
 ATACCAAGAT ACACCTGCGG CAGAAGGACT GATAA (SEQ ID
 Nº:184)

[00343] A sequência de aminoácidos da leitura aberta de 72732-2A-72678 é mostrada abaixo.

[00344] MDYKDHDG DY KDH DIDYKDD DDKMAPKKKR
KVGIHGVPA A MAERPFQCRI CMRNFSRSDD 1-60 LSKHIRTHTG
EKPFACDICG RKFADSSARK KHTKIHTGEK PFQCRICMRN
FSDRSNLSRH 61-120 IRTHTGEKPF ACDICGRKFA QRTHLRDHTK
IHTHPRAPIP KPFQCRICMR NFSQSGHLAR 121-180 HIRTHTGEKP

<u>FACDICGRKF</u>	<u>ADSSNREAHT</u>	<u>KIHLRGSQLV</u>	KSELEEKKSE
LRHKLKYVPH	181-240 EYIELIEIAR	NSTQDRILEM	KVMEFFMKVY
GYRGKHLGGS	RKPDGAIYTV	GSPIDYGVIV	241-300
DTKAYSGGYN	LDIGQADEME	RYVEENQTRD	KHLNPNEWWK
VYPSSVTEFK	FLFVSGHFKG	301-360	NYKAQLTRLN
HITNCNGAVL	SVEELLIGGE	MIKAGTLTLE	EVRRKFNNGE
INFRSGGGEG	361-420	<u>RGSLLTCGDV</u>	<u>EENPGPRTMD</u>
<u>YKDHDG DYKD HDIDYKDDDD KMAPKKKRKV</u>	<u>GIHGVPAA MG</u>	421-	
480 QLVKSELEEK	KSELRHKLKY	VPHEYIELIE	IAYNSTQDRI
LEMKVMEFFM	KVYGYRGKHL	481-540	GGSRKPDGAI
YTVGSPIDYG	VIVDTKAYSG	GYNLPIGQAD	EMQRYVKENQ
TRNKHINPNE	541-600	WWKVYPSSVT	EFKFLFVSGH
FKGNYKAQLT	RLNRKTNCNG	AVLSVEELLI	GGEMIKAGTL
601- 660 TLEEVRKFN	NGEINFSGAQ	<u>GSTLDFRPFQ</u>	<u>CRICMRKFAA</u>
<u>QCCLFHHTKI</u>	<u>HTGEKPFQCR</u>	661-720	<u>ICMRNFSDQS</u>
<u>NLRAHIRHT</u>	<u>GEKPFACDIC</u>	<u>GRKFARSANL</u>	<u>TRHTKIHTHP</u>
<u>RAPIPKPFQC</u>	721-780 <u>RICMRNFSRS</u>	<u>DDLTRHIRTH</u>	<u>TGEKPFACDI</u>
<u>CGRKFAQSGS</u>	<u>LTRHTKIHLR</u>	781-833	

[00345] (SEQ ID N°:185)

[00346] As características da sequência de aminoácidos de 72732-2A-72678 são mostradas abaixo in Tabela 14.

Tabela 14: Características da sequência de aminoácidos de 72732-2A-72678

Característica	Designação	Localização (na SEQ ID Nº:185)
Sequência 3x FLAG	<u>xx</u>	2-23
Sequência de localização nuclear	<u>xx</u>	26-32
Domínio de ligação ao DNA 72732	<u>xx</u>	44- 213
Ligante L0	xx	214-217
Domínio ELD P478D FokI	<u>xx</u>	218-413
Ligante 2A	<u>xx</u>	419- 436
Sequência 3x FLAG	<u>xx</u>	440-461
Sequência de localização nuclear	<u>xx</u>	464-470
Domínio KKR R416Y FokI	<u>xx</u>	481-676
Ligante N6a	xx	677-686
Domínio de ligação ao DNA 72678	<u>xx</u>	687-828

Exemplo 6: Testes *in vivo* de reagentes ZFN

[00347] Células T como aqui descritas são administradas a modelos animais da doença do enxerto contra hospedeiro e/ou câncer (por exemplo, camundongos *nude* injetados com linhagens celulares de câncer, como mieloma múltiplo, para estabelecer modelos de tumor). Por exemplo, células T humanas ativadas são eletroporadas com mRNAs que codificam as ZFNs específicas para B2M e TRAC, em que cada par é codificado por um único mRNA separado por uma sequência codificadora de um peptídeo de autoclivagem 2A (MacLeod, *et al.* (2017) *Mol Ther.* 25(4):949-961). As células são também transduzidas com partículas de AAV compreendendo um doador de CAR (por exemplo, CAR CD19). As células são então cultivadas e coradas para expressão do CAR e ausência de células CD3+. Quaisquer células CD3+ residuais são depletadas por separação

magnética. Camundongos NSG são injetados por via intravenosa com células Raji expressando luciferase de vagalume (Raji-ffLuc) e, depois de quatro dias, são injetados com as células T CAR CD3-/anti-CD19. O enxerto e crescimento das células Raji-ffLuc é evidente em quatro dias após a injeção e aumenta significativamente em camundongos não tratados. Os picos de frequência de células T CAR no sangue de camundongos tratados são observados no dia 8, atingindo ~10% de células no sangue periférico no grupo de dose alta. Nos dias 17–19, todos os camundongos nos grupos controle mostram evidência de carga tumoral significativa, especialmente na medula espinhal e medula óssea, resultando em paralisia completa das patas traseiras, e são sacrificados. Em contraste, todos os grupos de camundongos tratados com as células T CAR anti-CD19 não mostram evidência de crescimento tumoral no dia 11 e permaneceram livres de tumor até o dia 32 do estudo.

[00348] Nenhuma doença residual ou mínima é detectada no tecido dos animais (por exemplo, medula óssea, baço, pulmões, fígado, coração, etc.) que receberam as células T como aqui descritas. Em contraste, os indivíduos controles apresentam células tumorais detectáveis na maioria dos tecidos.

[00349] Todas as patentes, os pedidos de patente e publicações mencionados acima são aqui incorporados, em sua totalidade, por referência neste pedido de patente.

[00350] Embora a descrição tenha sido fornecida em alguns detalhes a título de ilustração e exemplo por questões de clareza e entendimento, será evidente para os técnicos no assunto que várias alterações e modificações podem ser praticadas sem que, contudo, se desviem do espírito e âmbito da invenção. Assim, a descrição anterior e os exemplos a seguir não devem ser interpretados como limitações.

REIVINDICAÇÕES

1. Nuclease de dedo de zinco, caracterizada pelo fato de que compreende:

uma ZFP de uma ZFN designada 68957, 72678, 72732 ou 72748;

um domínio de clivagem FokI projetado; e

um ligante entre o domínio de clivagem FokI e a ZFP.

2. Nuclease de dedo de zinco, de acordo com a reivindicação 1, caracterizada pelo fato de que compreende uma primeira e uma segunda ZFNs como segue: uma ZFN designada 57531 e uma ZFN designada 72732; uma ZFN designada 57531 e uma ZFN designada 72748; uma ZFN designada 68957 e uma ZFN designada 57071; uma ZFN designada 68957 e uma ZFN designada 72732; uma ZFN designada 68957 e uma ZFN designada 72748; uma ZFN designada 72678 e uma ZFN designada 57071; uma ZFN designada 72678 e uma ZFN designada 72732; e uma compreendendo uma ZFP ZFN designada 72678 e uma ZFN designada 727482.

3. Nuclease de dedo de zinco compreendendo uma primeira e uma segunda ZFNs, de acordo com a reivindicação 1 ou 2, caracterizada pelo fato de ser como a seguir: uma ZFN compreendendo uma ZFP da ZFN designada 72678 e uma ZFN compreendendo uma ZFP da ZFN designada 72732.

4. Polinucleotídeo, caracterizado pelo fato de que codifica a nuclease de dedo de zinco como definida em qualquer uma das reivindicações 1 a 3.

5. Polinucleotídeo, de acordo com a reivindicação 4, caracterizado pelo fato de que compreende uma sequência 2A entre as sequências que codificam a primeira e a segunda ZFNs.

6. Célula compreendendo a nuclease de dedo de zinco,

como definida em qualquer uma das reivindicações 1 a 3, ou um polinucleotídeo como definido nas reivindicações 4 e 5, caracterizada pelo fato de que o genoma da célula é modificado pela nuclease de dedo de zinco.

7. Célula, de acordo com a reivindicação 6, caracterizada pelo fato de que a célula é uma célula-tronco ou célula precursora.

8. Célula, de acordo com a reivindicação 7, caracterizada pelo fato de que a célula é uma célula humana.

9. Célula, de acordo com qualquer uma das reivindicações 6 a 8, caracterizada pelo fato de que a modificação genômica é selecionada a partir do grupo que consiste em inserções, deleções e combinações das mesmas.

10. Célula, de acordo com a reivindicação 6, caracterizada pelo fato de que compreende ainda uma ou mais modificações genômicas adicionais.

11. Célula, de acordo com a reivindicação 10, caracterizada pelo fato de que as modificações genômicas adicionais compreendem modificação de um gene do receptor de célula T (TCR), modificação de um gene HLA-A, modificação de um gene HLA-B, modificação de um gene HLA-C, modificação de um gene TAP, modificação de um gene CTLA-4, modificação de um gene PD1, modificação de um gene CISH, modificação de um gene tet-2 e/ou inserção de um transgene no genoma.

12. Célula, de acordo com a reivindicação 11, caracterizada pelo fato de que o transgene codifica pelo menos um receptor quimérico de antígeno (CAR).

13. Célula, de acordo com a reivindicação 12, caracterizada pelo fato de que a célula é uma célula T efetora ou uma célula T reguladora.

14. Célula ou linhagem celular, caracterizada pelo fato de

que descende da célula, como definida em qualquer uma das reivindicações 6 a 13.

15. Composição farmacêutica, caracterizada pelo fato de que compreende a nuclease de dedo de zinco de acordo com o polinucleotídeo, como definido na reivindicação 4 ou 5, ou a célula, como definida na reivindicação 12 ou 13.

16. Método para modificação de um gene endógeno de beta-2-microglobulina (B2M) em uma célula, caracterizado pelo fato de que compreende administrar o polinucleotídeo, como definido na reivindicação 4 ou 5, à célula de modo que o gene B2M endógeno seja modificado.

17. Método, de acordo com a reivindicação 16, caracterizado pelo fato de que compreende ainda introduzir uma sequência exógena na célula de modo que a sequência exógena seja inserida no gene B2M endógeno.

18. Método, de acordo com a reivindicação 16 ou 17, caracterizado pelo fato de que a modificação compreende uma deleção.

19. Método para produção de uma célula geneticamente modificada compreendendo uma modificação genômica dentro de um gene B2M endógeno, caracterizado pelo fato de que compreende as etapas de:

a) colocar uma célula em contato com o polinucleotídeo como definido na reivindicação 4 ou 5;

b) submeter a célula a condições que conduzam à expressão da proteína de fusão a partir do polinucleotídeo; e

c) modificar o gene B2M endógeno com a proteína de fusão expressa suficiente para produzir a célula geneticamente modificada.

20. Kit, caracterizado pelo fato de que compreende o polinucleotídeo, como definido na reivindicação 4 ou 5.

21. Método de tratamento e prevenção de um câncer em um indivíduo, caracterizado pelo fato de que compreende administrar, ao indivíduo, a célula como definida em qualquer uma das reivindicações 6 a 14, ou a composição farmacêutica como definida na reivindicação 15.

22. Método de tratamento ou prevenção de uma doença autoimune em um indivíduo, caracterizado pelo fato de que compreende administrar, ao indivíduo, a célula como definida em qualquer uma das reivindicações 6 a 14, ou a composição farmacêutica como definida na reivindicação 15.

23. Nuclease de dedo de zinco (ZFN), caracterizada pelo fato de que compreende ZFNs esquerda e direita como segue: uma ZFN designada 68796 e uma ZFN designada 68813; uma ZFN designada 68796 e uma ZFN designada 68861; uma ZFN designada 68812 e uma ZFN designada 68813; uma ZFN designada 68876 e uma ZFN designada 68877; uma ZFN designada 68815 e uma ZFN designada 55266; uma ZFN designada 68879 e uma ZFN designada 55266; uma ZFN designada 68798 e uma ZFN designada 68815; ou uma ZFN designada 68846 e uma ZFN designada 53853.

24. Polinucleotídeo codificador de uma ou mais nucleases de dedo de zinco como definidas na reivindicação 23, caracterizado pelo fato de que o polinucleotídeo é opcionalmente mRNA.

25. Polinucleotídeo, de acordo com a reivindicação 24, caracterizado pelo fato de que compreende uma sequência 2A entre as sequências que codificam as ZFNs esquerda e direita.

26. Célula, caracterizada pelo fato de que compreende a nuclease de dedo de zinco, como definida em qualquer uma das reivindicações 23 a 25, em que o genoma da célula é modificado pela nuclease de dedo de zinco.

27. Célula, de acordo com a reivindicação 26, caracterizada

pelo fato de que a célula é uma célula-tronco ou célula precursora.

28. Célula, de acordo com a reivindicação 26 ou 27, caracterizada pelo fato de que a célula é uma célula humana.

29. Célula, de acordo com qualquer uma das reivindicações 26 a 28, caracterizada pelo fato de que a modificação genômica é selecionada a partir do grupo que consiste em inserções, deleções e combinações da mesma.

30. Célula, de acordo com a reivindicação 29, caracterizada pelo fato de que compreende ainda uma ou mais modificações genômicas adicionais.

31. Célula, de acordo com a reivindicação 30, caracterizada pelo fato de que as modificações genômicas adicionais compreendem modificação de um gene B2M, modificação de um gene HLA-A, modificação de um gene HLA-B, modificação de um gene HLA-C, modificação de um gene TAP, modificação de um gene CTLA-4, modificação de um gene PD1 e/ou inserção de um transgene no genoma.

32. Célula, de acordo com a reivindicação 31, caracterizada pelo fato de que o transgene codifica pelo menos um receptor quimérico de antígeno (CAR).

33. Célula, de acordo com a reivindicação 31 ou 32, caracterizada pelo fato de que a célula é uma célula-tronco.

34. Célula ou linhagem celular, caracterizada pelo fato de que descende da célula como definida em qualquer uma das reivindicações 26 a 33.

35. Composição farmacêutica, caracterizada pelo fato de que compreende a nuclease de dedo de zinco de acordo com o polinucleotídeo, como definido na reivindicação 24 ou 25, ou uma célula, como definida em qualquer uma das reivindicações 26 a 34.

36. Método para modificação de um gene endógeno do

receptor de células T (TCR) em uma célula, caracterizado pelo fato de que o método compreende administrar o polinucleotídeo como definido na reivindicação 24 ou 25, à célula de modo que o gene TCR endógeno seja modificado.

37. Método, de acordo com a reivindicação 36, caracterizado pelo fato de que compreende ainda introduzir uma sequência exógena na célula de modo que a sequência exógena seja inserida no gene TCR endógeno.

38. Método, de acordo com a reivindicação 36, caracterizado pelo fato de que a modificação compreende uma deleção.

39. Método de produção de uma célula geneticamente modificada compreendendo uma modificação genômica dentro de um gene TCR endógeno, caracterizado pelo fato de que compreende as etapas de:

a) colocar uma célula em contato com o polinucleotídeo como definido na reivindicação 24 ou 25;

b) submeter a célula a condições que conduzam à expressão da proteína de fusão a partir do polinucleotídeo; e

c) modificar o gene TCR endógeno com a proteína de fusão expressa suficiente para produzir a célula geneticamente modificada.

40. Kit, caracterizado pelo fato de que compreende o polinucleotídeo como definido na reivindicação 24 ou 25.

41. Método de tratamento e prevenção de um câncer ou doença do enxerto contra hospedeiro e um indivíduo, caracterizado pelo fato de que compreende administrar, ao indivíduo, a célula como definida em qualquer uma das reivindicações 26 a 34.

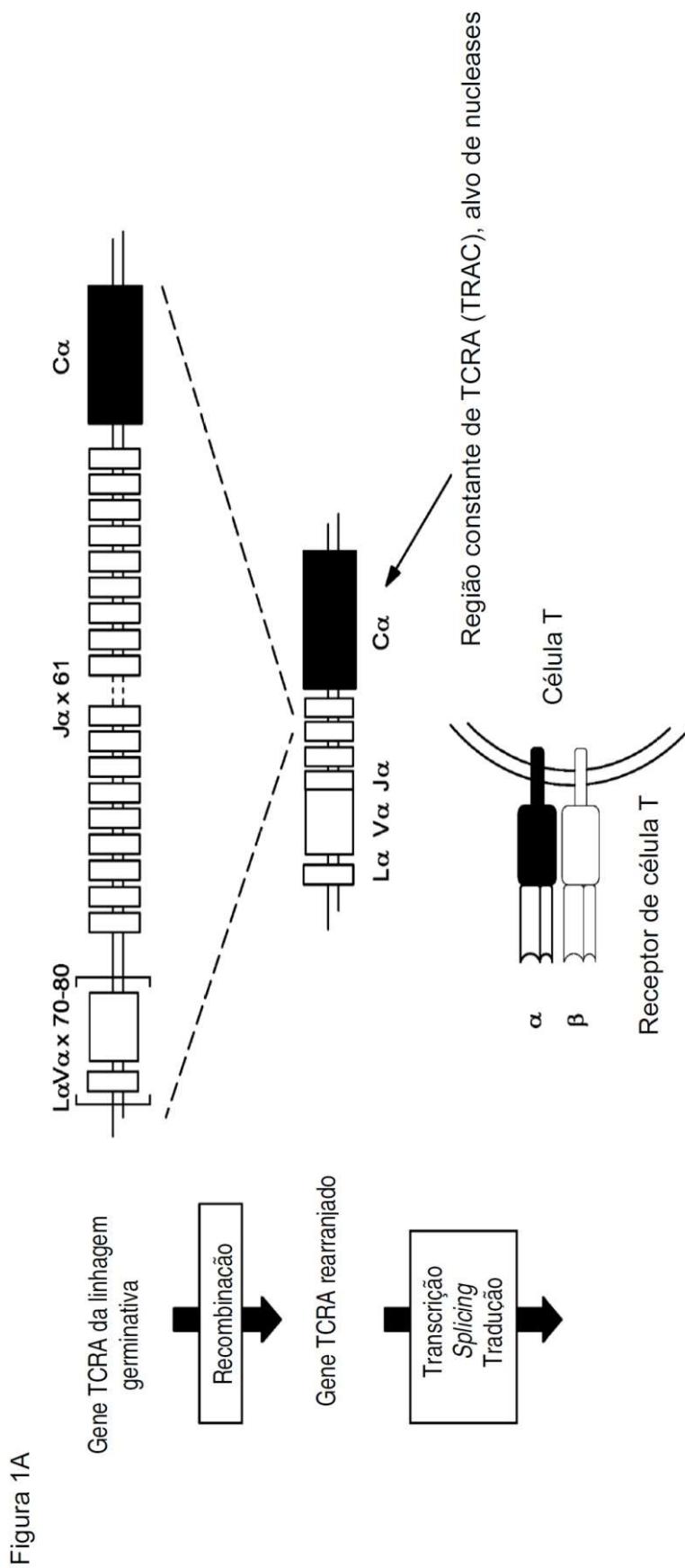


figura 1B

Exon c1

1 aATCAGAACCTGACCTGGCGTGTACCGCTAATCCAGTCAAGCTGAGACTTAATCCAGTTGATTCTAACAAATGT

101 GTCACAAAGTAAGGATTCTGTATATCACAGACAAAACTGTGCTAGACATGAGGTCTATGGACTCAAGGCAACAGTCTGGAGCAAC

201 AAATCTGACTTGGCATGGCAAAAGCCTTCACAAACGATTATTCCAGAAGACCTTCTCCCCAGCCCAAGgttaaggcagcttttgtgccttcgac

Exon c2

2101 tctggatgtaaaatgtctgtttccctttaAAAGCTCCTGTGATGTCAAGCTTGAACAGgtaaagacagggtcttagcc
E

Exon c3

3021 aagccataaccgctgtggccttgcgtttacagatacGAACCTAAACTTCAAAACCTGTTCCGAAATCCTCTCCTGAAAGTGGCC
Arg D Lys

Figura 2A

Sítio A, B e D de TRAC modificado

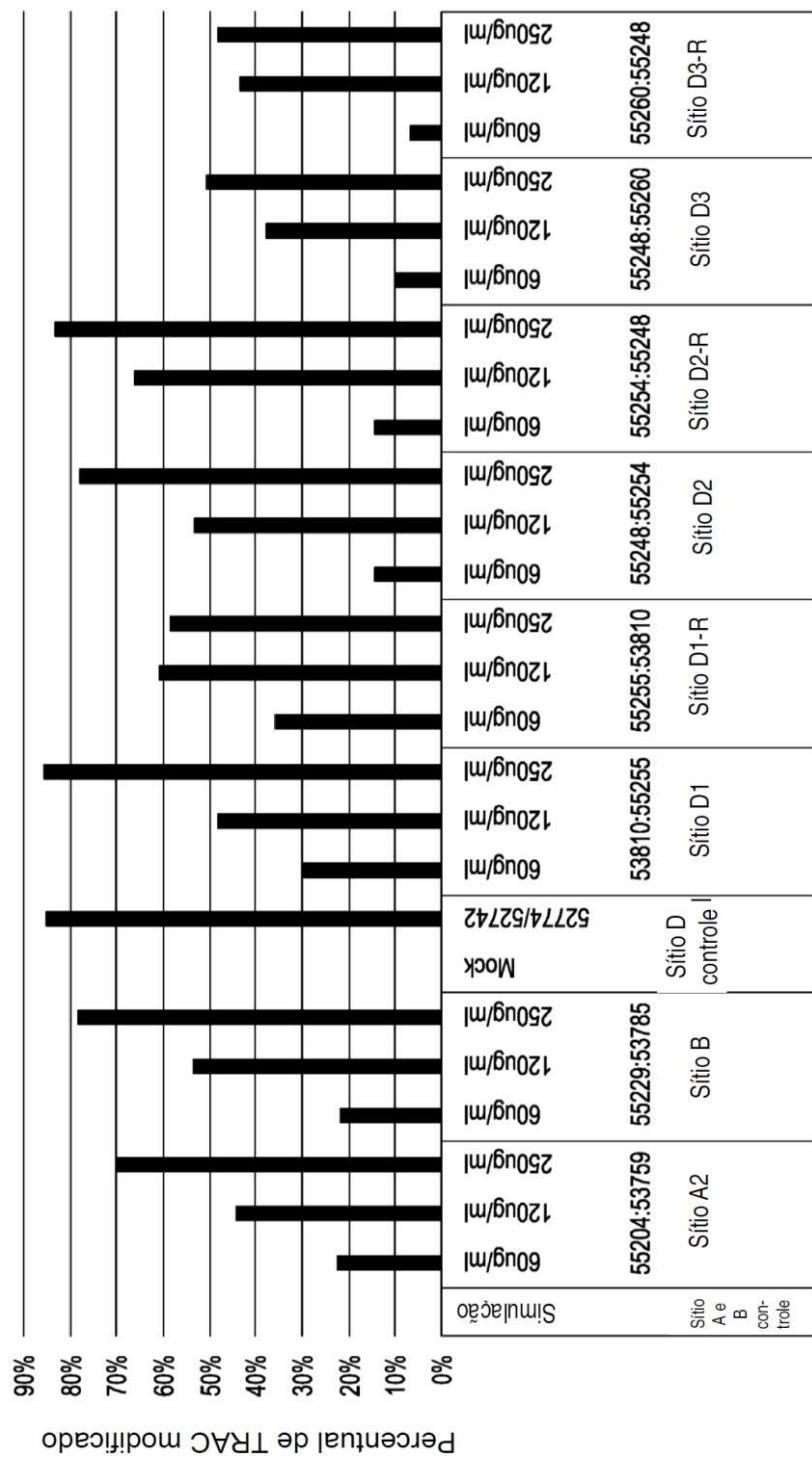
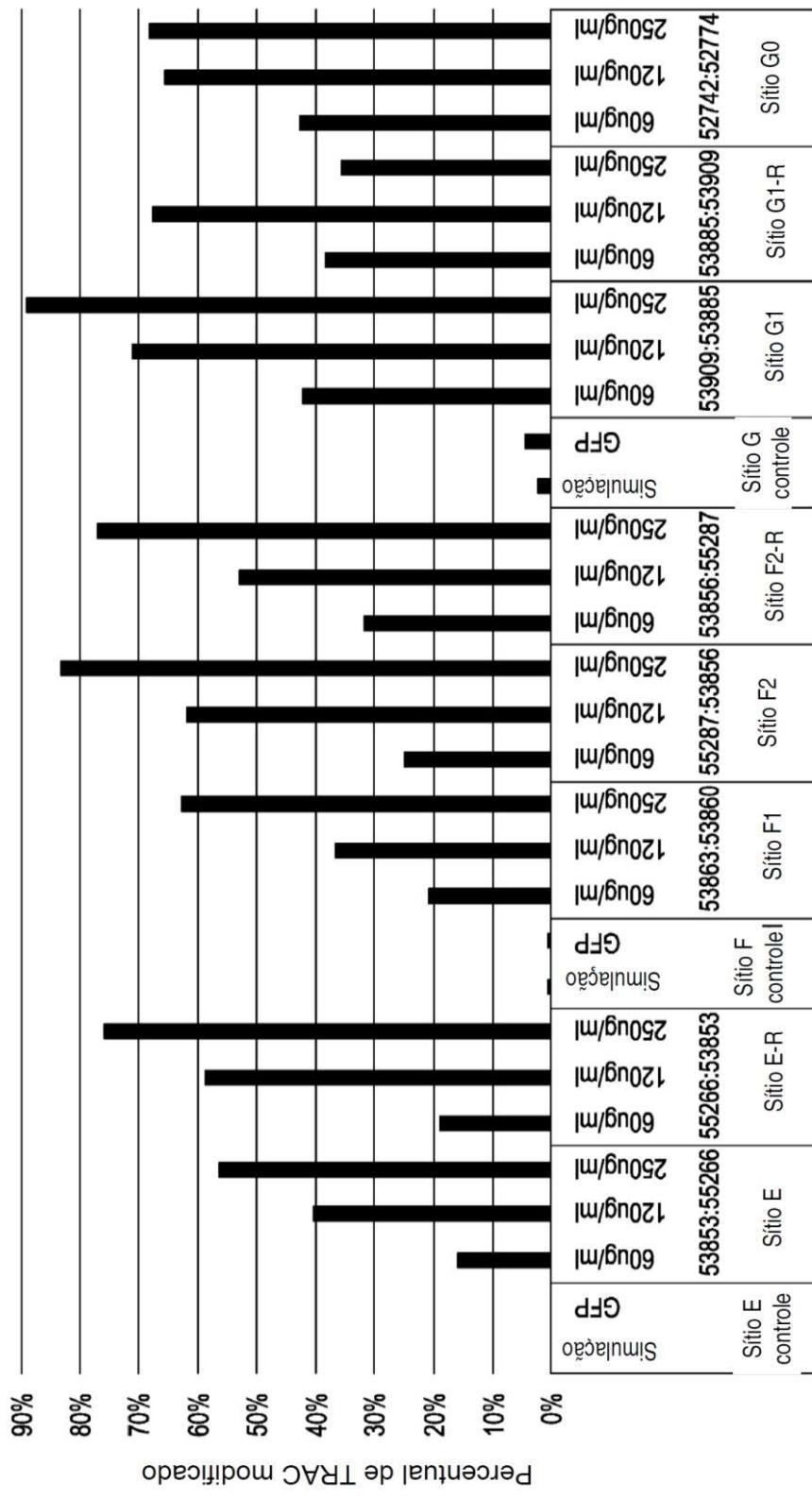
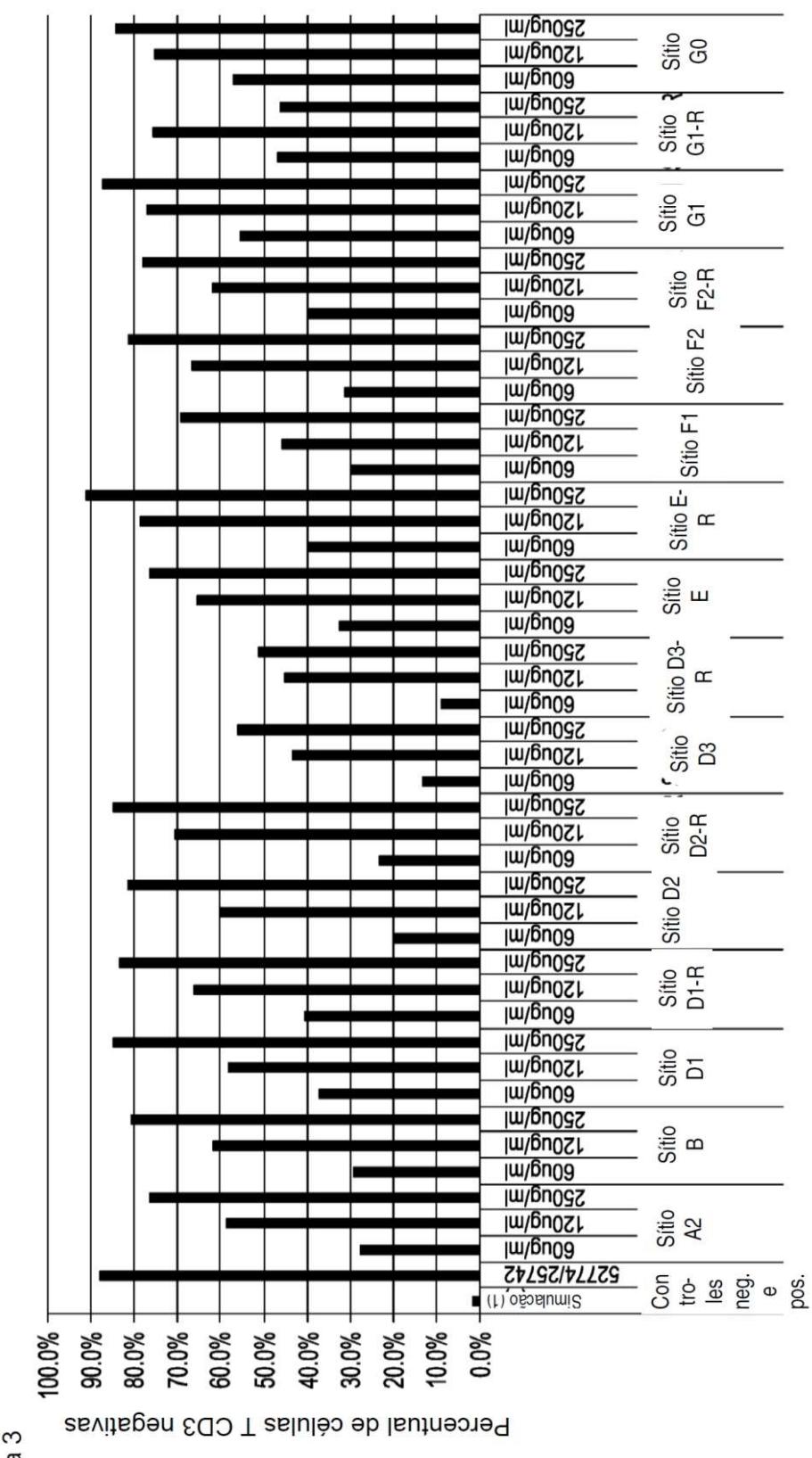


Figura 2B
Sítio E, F e G de TRAC modificado





6/9

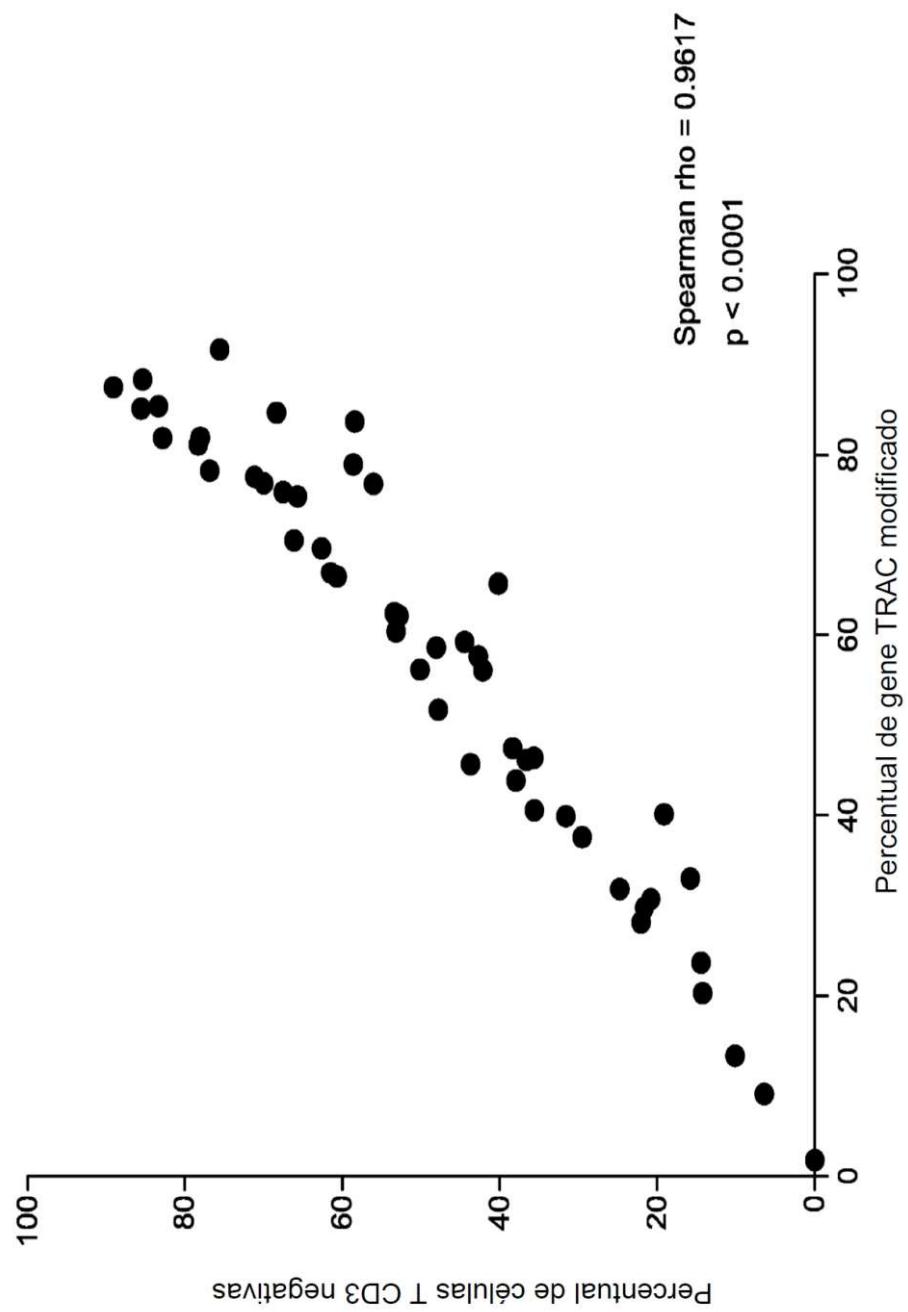
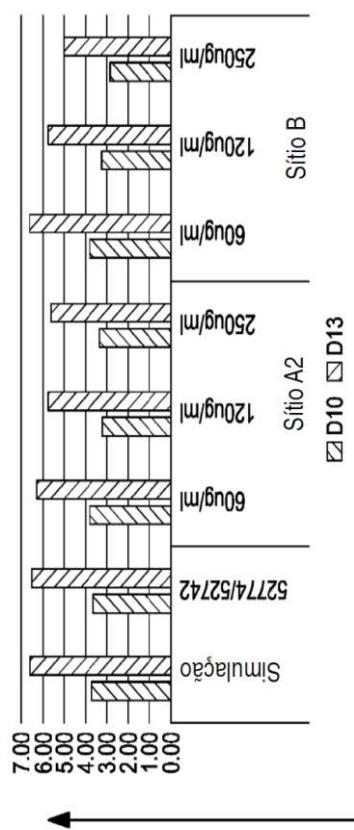
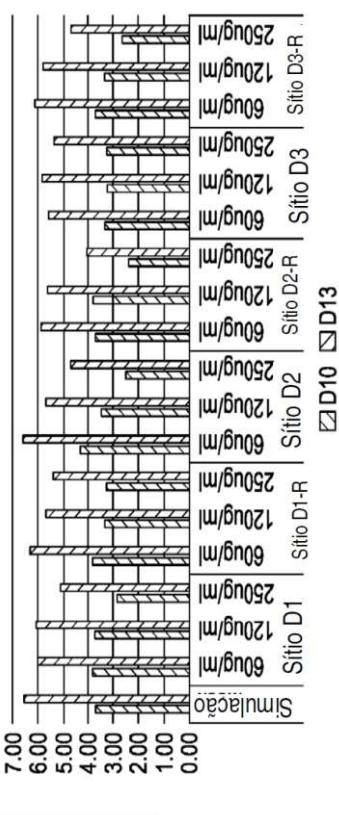


Figura 4

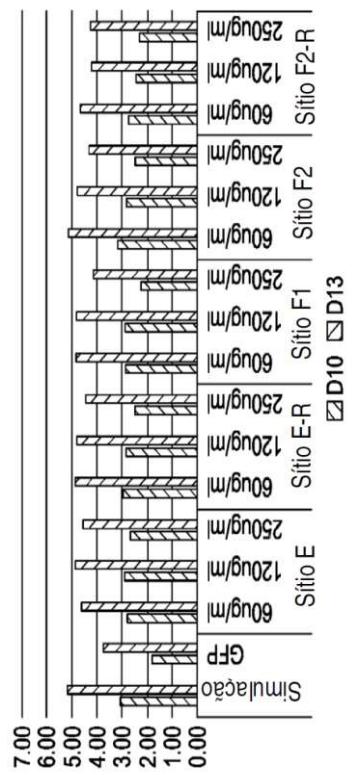
Crescimento celular para os Sítios A e B

**FIG. 5A**

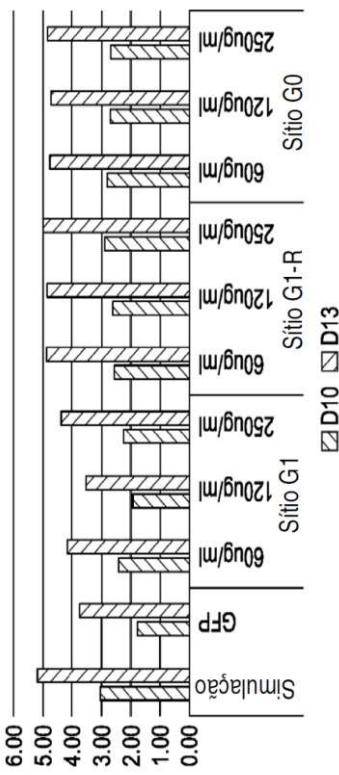
Crescimento celular para os Sítios D

**FIG. 5C**

Crescimento celular para os Sítios E e F

**FIG. 5B**

Crescimento celular para os Sítios G

**FIG. 5D**

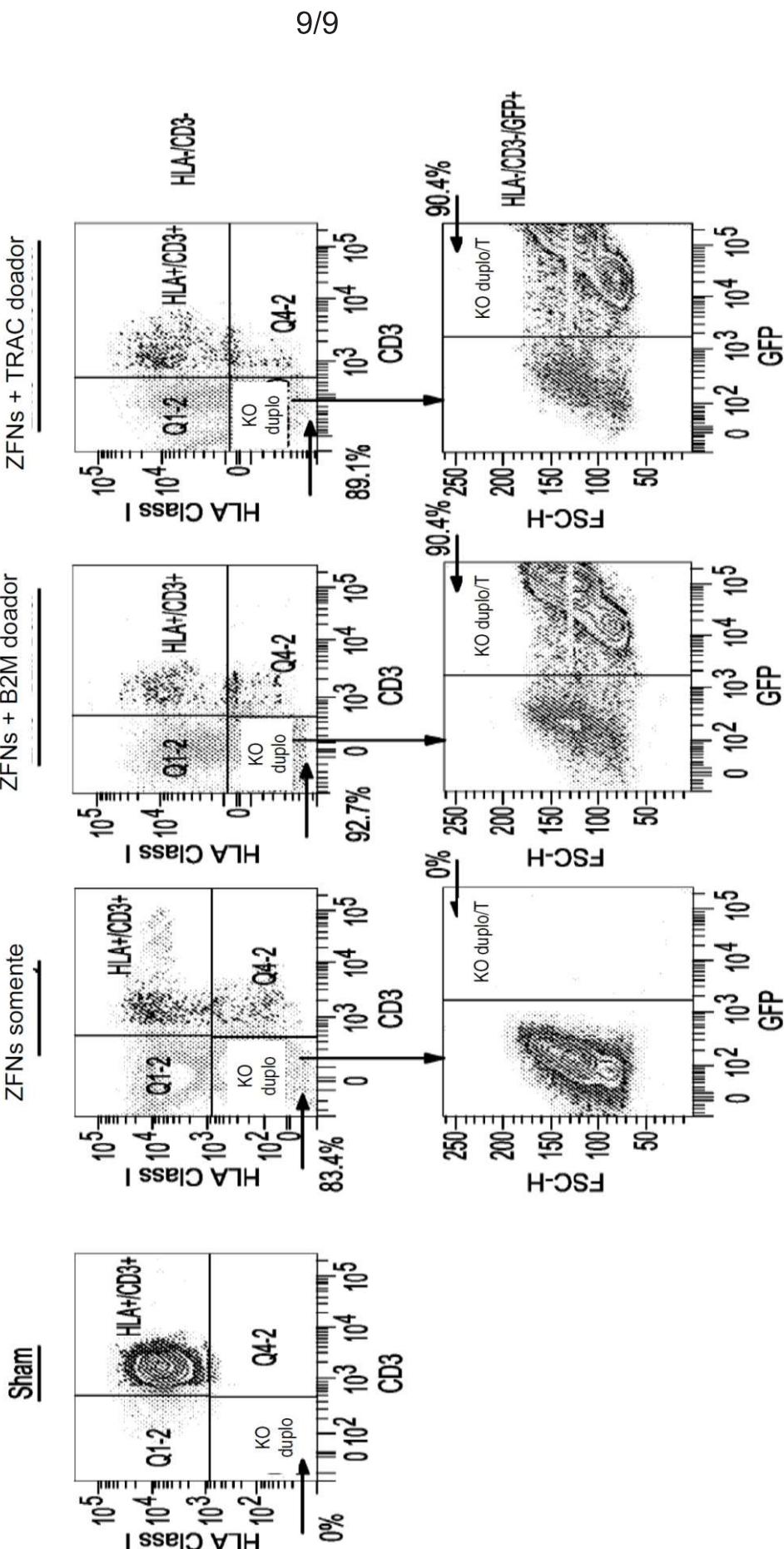
Duplicações de células Totais

figura 6

8/9

		FACS			
		D10		D10	
μg/ml	TRAC-	B2M-	DUPLO-	GFP+ Total	GFP+ DUPLO-
Sham	0,5	0,1	0,0	0,0	0,0
TRAC / B2M KO somente	85,0	83,6	80,0	0,0	0,0
TRAC / B2M KO + 1E5vg / célula TRAC lócus AAV doador	91,9	92,7	89,3	80,8	83,0
TRAC / B2M KO + 3E4vg / célula TRAC lócus AAV doador	91,2	93,4	89,1	71,9	74,3
TRAC / B2M KO + 1E5vg / célula B2M lócus AAV doador	88,2	90,5	86,4	54,9	59,6
TRAC / B2M KO + 3E4vg / célula B2M lócus AAV doador	89,8	92,2	87,9	43,2	46,7

Figura 7



RESUMO

Patente de Invenção: **"DISRUPÇÃO DIRECIONADA DE RECEPTORES DE CÉLULA T E/OU HLA".**

A presente invenção refere-se a métodos e composições para inativar os genes de TCR e/ou HLA, utilizando nucleases projetadas que compreendem pelo menos um domínio de ligação ao DNA e um domínio de clivagem ou meio-domínio de clivagem em condições capazes de preservar a viabilidade celular. Além disso, são providos também polinucleotídeos codificadores de nucleases, vetores compreendendo polinucleotídeos codificadores de nucleases e células compreendendo polinucleotídeos codificadores de nucleases e/ou células compreendendo nucleases.

Este anexo apresenta o código de controle da listagem de sequências biológicas.

Código de Controle

Campo 1



Campo 2



Outras Informações:

- Nome do Arquivo: LISTAGEM DE SEQUÊNCIA 13.12.19 P244266.TXT
- Data de Geração do Código: 13/12/2019
- Hora de Geração do Código: 15:58:33
- Código de Controle:
 - Campo 1: F36963B7E565108D
 - Campo 2: 15FC5D0B5AC4D08D