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

(21) BR 112020006256-1 A2

(22) Data do Depósito: 28/09/2018

(43) Data da Publicação Nacional: 20/10/2020

(54) Título: MÉTODO IN VITRO DE ADMINISTRAÇÃO DE MRNA USANDO NANOPARTÍCULAS DE LIPÍDIO

(51) Int. Cl.: C12N 15/88; A61K 38/46; C12N 9/22; C12N 15/10; C12N 15/113.

(30) Prioridade Unionista: 29/09/2017 US 62/566,232.

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

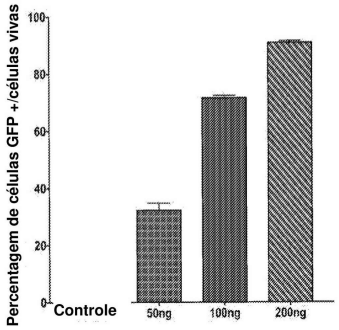
(72) Inventor(es): MORAG HELEN STEWART; MARKUS PARZIVAL VALLASTER; ANTHONY MONTI; POOJA KYATSANDRA NARENDRA; QUAN HO.

(86) Pedido PCT: PCT US2018053569 de 28/09/2018

(87) Publicação PCT: WO 2019/067999 de 04/04/2019

(85) Data da Fase Nacional: 27/03/2020

(57) Resumo: A presente invenção refere-se a composições e métodos para a introdução de um mRNA em células-tronco, como HSPCs, e para a admi-nistração de componentes de edição de genes a essas células in vitro. Por exemplo, a divulgação se refere à modificação de uma sequência gênica usando um complexo de CRISPR-Cas9 em HSPCs, e métodos e sistemas de administração para atingir tal modificação gênica em HSPCs.



Relatório Descritivo da Patente de Invenção para **"MÉTODO IN VITRO DE ADMINISTRAÇÃO DE MRNA USANDO NANOPARTÍCULAS DE LIPÍDIO".**

[0001] O presente pedido reivindica o benefício de prioridade ao Pedido de Patente Provisório US N° 62/566.232, depositado em 29 de setembro de 2017, cujo conteúdo é incorporado neste documento por referência na sua totalidade.

[0002] A introdução de alterações genéticas nas células-tronco, incluindo células-tronco hematopoiéticas (HSCs), e sua progênie é de interesse para os métodos de edição e terapia gênica. Células-tronco como HSCs têm capacidades proliferativas perdidas em células maduras e progenitores comprometidos, tornando-as particularmente úteis para tecnologias de edição de genes. A capacidade de modificar HSCs e células-tronco *in vitro* é importante, por exemplo, e são necessários métodos para fornecer agentes biológicos às HSCs e outras células-tronco em cultura. Existe uma necessidade particular de tecnologias de entrega para HSCs humanas em cultura.

[0003] As HSCs são indispensáveis para a produção de sangue ao longo da vida. As HSCs podem sustentar hematopoiese a longo prazo e funcional devido à sua capacidade de se diferenciar para produzir progênie madura de todas as linhagens sanguíneas mieloides e linfoides ou de se auto-renovar para substituir as células que se comprometem progressivamente com a diferenciação. As HSCs podem ser usadas para restaurar células sanguíneas e imunológicas em receptores de transplante, em pacientes imunocomprometidos ou em outros pacientes. Especificamente, o transplante autólogo ou alogênico de HSCs pode ser usado para o tratamento de pacientes com doenças imunodeficientes e autoimunes herdadas e diversos distúrbios hematopoiéticos para reconstituir as linhagens celulares hematopoiéticas e a defesa do sistema imunológico.

[0004] Métodos para entregar componentes de sistemas de edição de genes CRISPR/Cas às HSCs em cultura são de particular interesse. Métodos de entrega de RNAs, incluindo componentes do sistema CRISPR/Cas para culturas de células hematopoiéticas que incluem HSCs são fornecidos neste documento. Os métodos entregam proteína ativa às células-tronco, incluindo HSCs, cultivadas *in vitro* e incluem o contato das células com uma composição de nanopartículas lipídicas (LNP) que fornece um mRNA que codifica a proteína. Além disso, são fornecidos métodos de edição de genes em células-tronco, como HSCs in vitro, e métodos de produção de uma célula manipulada.

[0005] Em algumas modalidades, são fornecidos métodos de edição de genes nas HSCs *in vitro*e métodos de produção de uma célula HSC manipulada. Em outras modalidades, é fornecido neste documento um método de entrega de um mRNA a uma célula-tronco hematopoiética e/ou célula progenitora (HSPC) ou a uma população de HSPC. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato da população de HSPC ou HSPC com a composição de LNP pré-incubada in vitro. Em algumas modalidades, o método compreende ainda a cultura da população HSPC ou HSPC in vitro. Em algumas modalidades, o método resulta na entrega do mRNA à população de HSPC ou HSPC.

[0006] Em algumas modalidades, é proporcionado neste documento um método de introdução de um mRNA da nuclease Cas e um gRNA a uma célula-tronco, por exemplo, um HSPC. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA de Nuclease Cas, um gRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neu-

tro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato de HSPC com a composição de LNP pré-incubada in vitro. Em algumas modalidades, o método compreende ainda a cultura do HSPC. Em algumas modalidades, o método resulta na introdução do mRNA e do gRNA da nuclease Cas no HSPC.

[0007] Em algumas modalidades, é proporcionado neste documento um método de produção de uma célula-tronco geneticamente modificada, por exemplo, HSPC, in vitro. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA de Nuclease Cas, um gRNA, um mRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato de HSPC com a composição de LNP pré-incubada in vitro. Em algumas modalidades, o método compreende ainda a cultura do HSPC in vitro. Em algumas modalidades, o método resulta na produção de um HSPC geneticamente manipulado.

[0008] Em algumas modalidades, é fornecido um método de entrega de um mRNA a uma população de HSPC ou HSPC, compreendendo a pré-incubação de uma composição de LNP com um fator sérico, entrando em contato com a célula ou população com a composição de LNP pré-incubada *in vitro*; e cultivar a célula ou população *in vitro*; entregando assim o mRNA ao HSPC. Em algumas modalidades, o HSPC é um HSC. Em algumas modalidades, os métodos entregam um mRNA, como um mRNA de nuclease Cas, a uma população de HSPC (por exemplo, uma população de células CD34 +). Em certas modalidades, um RNA guia (gRNA), opcionalmente em combinação com um mRNA da nuclease Cas, é entregue às células.

BREVE DESCRIÇÃO DAS FIGURAS

[0009] A Fig. 1 mostra a entrega de mRNA da proteína fluorescente verde (GFP) nas células da medula óssea CD34+ usando LNPs.

[0010] A Fig. 2 mostra que a entrega de mRNA nas células da medula óssea CD34+ depende da pré-incubação com soro.

[0011] As Figs. 3A e 3B mostram edição de *B2M* nas células da medula óssea CD34+ com pré-incubação de soro, com a Fig. 3A representando a percentagem de células B2M- (nocaute de expressão de proteínas) e a Fig. 3B representando a percentagem de edição alcançada no experimento.

[0012] As Figs. 4A e 4B mostram entrega eficiente com pré-incubação sérica e pré-incubação com ApoE3. A Fig. 4A representa a percentagem de células B2M- e a Fig. 4B fornece a percentagem de edição alcançada na experiência.

[0013] A Fig. 5 mostra o efeito da pré-incubação de LNP com preparações de vários fatores séricos na entrega de LNP às células CD34+.

[0014] As Figs. 6A e 6B mostram dados de viabilidade e edição para células CD34+ que foram expostas ao tratamento com LNP em intervalos variados. A Fig. 6A mostra a viabilidade das células CD34+ após a exposição ao LNP às 2, 6 e 24 horas. A Fig. 6B fornece a percentagem de dados de edição para os grupos de tratamento de 2, 6 e 24 horas.

DESCRIÇÃO DETALHADA

[0015] A presente invenção fornece métodos de uso de composições de LNP de RNAs, incluindo RNAs de componentes CRISPR/Cas (a "carga"), para entrega *in vitro* a células CD34+, por exemplo, populações de células contendo HSC. Os métodos podem exibir propriedades aprimoradas em comparação com as tecnologias de entrega anteriores, por exemplo, os métodos fornecem entrega eficiente dos RNAs, enquanto reduzem a morte celular causada pela transfecção.

[0016] Em algumas modalidades, é fornecido neste documento um método de entrega de um mRNA a uma célula-tronco, por exemplo,

uma população de HSPC ou HSPC. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato da população de HSPC ou HSPC com a composição de LNP pré-incubada *in vitro*. Em algumas modalidades, o método compreende ainda a cultura da população HSPC ou HSPC *in vitro*. Em algumas modalidades, o método resulta na entrega do mRNA à população de HSPC ou HSPC. Em algumas modalidades, o mRNA codifica uma nuclease Cas.

[0017] Em algumas modalidades, é proporcionado neste documento um método de introdução de um mRNA da nuclease Cas e um gRNA a uma célula-tronco, por exemplo, um HSPC ou uma população HSPC. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA de Nuclease Cas, um gRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato de HSPC com a composição de LNP pré-incubada *in vitro*. Em algumas modalidades, o método compreende ainda a cultura do HSPC. Em algumas modalidades, o método resulta na introdução do mRNA e do gRNA da nuclease Cas no HSPC.

[0018] Em algumas modalidades, é proporcionado neste documento um método de produção de uma célula-tronco geneticamente manipulada, por exemplo, HSPC, *in vitro*. Em algumas modalidades, o método compreende a pré-incubação de um fator sérico com uma composição de LNP compreendendo o mRNA de Nuclease Cas, um gRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, o método compreende ainda o contato de HSPC com a composição de LNP pré-incubada *in*

vitro. Em algumas modalidades, o método compreende ainda a cultura do HSPC *in vitro*. Em algumas modalidades, o método resulta na produção de um HSPC geneticamente manipulado.

[0019] Em algumas modalidades, a composição compreende ainda um gRNA. Em algumas modalidades, o mRNA codifica uma nuclease Cas de Classe 2. Em certas modalidades, o componente ou carga de RNA inclui uma mRNA da nuclease Cas, como uma mRNA da nuclease Cas de Classe 2. Em certas modalidades, o componente ou carga de RNA inclui um gRNA do sistema CRISPR/Cas ou ácidos nucleicos que codificam um gRNA. Métodos de edição de genes e métodos de produção de células manipuladas também são fornecidos.

Métodos *in vitro*

[0020] Os presentes métodos entregam RNAs para células CD34+ *in vitro*. "Células CD34+" referem-se às células que expressam em seu marcador CD34 de superfície. As células CD34+ podem ser detectadas e contadas usando, por exemplo, citometria de fluxo e anticorpos CD34 anti-humanos marcados com fluorescência.

[0021] Em algumas modalidades, é fornecido um método de entrega de um mRNA a uma célula-tronco, por exemplo, uma população de HSPC ou HSPC, o método compreendendo (a) pré-incubando um fator sérico com uma composição de LNP compreendendo o mRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG; (b) entrando em contato com a população de HSPC ou HSPC com a composição pré-incubada de LNP *in vitro*; e (c) cultivando a população de HSPC ou de HSPC *in vitro*; entregando dessa forma o mRNA ao HSPC. Em algumas modalidades, o mRNA codifica uma nuclease Cas, como uma nuclease Cas de Classe 2. Em alguns aspectos, o mRNA da nuclease Cas da Classe 2 é um mRNA de Cas9 ou um mRNA de Cpf1. Em certas modalidades, a nuclease Cas da Classe 2 é uma *S.pyogenes* Cas9. Em algumas modalidades, a com-

posição de LNP compreende ainda um gRNA. Em modalidades adicionais, os métodos introduzem um mRNA da nuclease Cas e um gRNA em um HSPC, o método compreendendo (a) pré-incubando um fator sérico com uma composição de LNP compreendendo o mRNA da nuclease Cas, um gRNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG; (b) entrando em contato com o HSPC com a composição pré-incubada de LNP *in vitro*; e (c) cultivando o HSPC; introduzindo dessa forma a nuclease Cas e o gRNA no HSPC.

[0022] Em várias modalidades, os gRNAs dos métodos descritos nestes documentos podem ser um RNA de guia duplo (dgRNA) ou um RNA de guia único (sgRNA).

[0023] Em algumas modalidades dos métodos *in vitro*, a transfecção de LNP pode reduzir a morte celular de HSPC ou CD34+ em comparação com tecnologias conhecidas como eletroporação. Em algumas modalidades, a transfecção de LNP pode causar menos de 5%, menos de 10%, menos de 20%, menos de 30% ou menos de 40% de morte celular. Em certas modalidades, a sobrevida de células pós-transfecção é de pelo menos 60%, 70%, 80%, 90% ou 95%.

[0024] As células-tronco são caracterizadas pela capacidade de se auto-renovar e se diferenciar em uma gama diversificada de tipos de células. Os dois tipos amplos de células-tronco de mamíferos são células-tronco embrionárias (ES) e células-tronco adultas. As células-tronco adultas ou células progenitoras podem repor células especializadas. A maioria das células-tronco adultas tem restrição de linhagem e pode ser referida por sua origem tecidual. As linhas de células ES são derivadas do tecido epiblasto da massa celular interna de um blastocisto ou embriões em estágio inicial de mórula. As células ES são pluripotentes e dão origem a derivados das três camadas germinais, ou seja, o ectoderma, endoderma e mesoderma. As células-tronco plu-

ripotentes induzidas (iPSCs) são células adultas que foram reprogramadas geneticamente para um estado semelhante a uma célula-tronco embrionária, sendo forçadas a expressar genes e fatores importantes para manter as propriedades definidoras das células-tronco embrionárias. Uma "célula-tronco" pode ser uma ESC, uma iPSC, uma célula progenitora ou um HSPC, por exemplo.

[0025] Os termos "célula-tronco hematopoiética e/ou célula progenitora" e "HSPC" são usados de forma intercambiável e se referem a uma população de células compreendendo HSCs e células progenitoras hematopoiéticas ("HPCs""). Tais células são caracterizadas, por exemplo, como CD34+. Em modalidades exemplares, os HSPCs são isolados da medula óssea. Em outras modalidades exemplares, os HSPCs são isolados do sangue periférico. Em outras modalidades exemplares, os HSPCs são isolados do sangue do cordão umbilical.

[0026] Os HSPCs podem ser derivados da medula óssea, sangue periférico ou sangue do cordão umbilical e podem ser autólogos (as células-tronco do próprio paciente) ou alogênicos (as células-tronco são de um doador).

[0027] O termo "células progenitoras hematopoiéticas" ou "HPCs", conforme usado neste documento, refere-se às células hematopoiéticas primitivas que têm uma capacidade limitada de auto-renovação e o potencial de diferenciação de várias linhagens (por exemplo, mieloide, linfoide), diferenciação de linhagem (por exemplo, mieloide ou linfoide) ou diferenciação restrita do tipo celular (por exemplo, progenitor eritroide), dependendo do posicionamento dentro da hierarquia hematopoiética (Doulatov et al., Cell Stem Cell 2012).

[0028] O termo "células-tronco hematopoiéticas" ou "HSCs", conforme usado neste documento, também se refere às células sanguíneas imaturas com capacidade de se auto-renovar e se diferenciar em células sanguíneas mais maduras que compreendem granulócitos (por

exemplo, promielócitos, neutrófilos, eosinófilos, basófilos), eritrócitos (por exemplo, reticulócitos, eritrócitos), trombócitos (por exemplo, megacarioblastos, megacariócitos produtores de plaquetas, plaquetas) e monócitos (por exemplo, monócitos, macrófagos). É conhecido na técnica que essas células podem ou não incluir células CD34+. As células CD34+ são células imaturas que expressam o marcador de superfície celular CD34. Acredita-se que as células CD34+ incluam uma subpopulação de células com as propriedades de células-tronco definidas acima. O transplante de populações de células, como HSPCs que contêm HSCs multipotentes, pode ser usado para tratar leucemia, linfoma e outros distúrbios.

[0029] As HSCs são células multipotentes que podem dar origem às células progenitoras primitivas (por exemplo, células progenitoras multipotentes) e/ou células progenitoras comprometidas com linhagens hematopoiéticas específicas (por exemplo, células progenitoras linfoides). As células-tronco comprometidas com linhagens hematopoiéticas específicas podem ser de linhagem de células T, linhagem de células B, linhagem de células dendríticas, linhagem de células de Langerhans e/ou linhagem de células de macrófagos específicas de tecido linfoide. Além disso, os HSCs também se referem a HSC de longo prazo (LT-HSC) e HSC de curto prazo (ST-HSC). ST-HSCs são mais ativos e mais proliferativos que LT-HSCs. No entanto, o LT-HSC possui autorenovação ilimitada (isto é, eles sobrevivem ao longo da vida adulta), enquanto o ST-HSC possui autorrenovação limitada (ou seja, eles sobrevivem apenas por um período limitado de tempo). Qualquer um desses HSCs pode ser usado em qualquer um dos métodos descritos neste documento. Opcionalmente, os ST-HSCs são úteis porque são altamente proliferativos e, portanto, aumentam rapidamente o número de HSCs e sua progênie.

[0030] HSCs, HPCs e HSPC s são opcionalmente obtidos de pro-

dutos derivados de sangue. Um produto sanguíneo inclui um produto obtido do corpo ou de um órgão do corpo contendo células de origem hematopoiética. Tais fontes incluem medula óssea, cordão umbilical, sangue periférico (por exemplo, sangue periférico mobilizado, por exemplo, mobilizado com um agente de mobilização como G-CSF ou Plerixafor® (AMD3100)), fígado, timo, linfa e baço. Todos os produtos sanguíneos mencionados acima (por exemplo, nas formas bruta, não fracionada ou fracionada) podem ser enriquecidos para células com características HSC de maneiras conhecidas pelos versados na técnica. Da mesma forma, eles podem ser enriquecidos pelas características da população de HPC e/ou HSPC. Numa modalidade, as HSCs são caracterizadas como CD34+/CD38-/CD90+/CD45RA-. Nas modalidades, as HSCs são caracterizadas como células CD34+/CD90+/CD49f+. Em modalidades adicionais, as HSCs são caracterizadas como Lineage-CD34+/CD38-/CD90+/CD45RA-. Nas modalidades, as HSCs são caracterizadas como células Lineage-CD34+/CD90+/CD49f+, em que "linhagem" significa omitir marcadores para células diferenciadas terminalmente, por exemplo, células T, células B etc. Estes podem ser excluídos pela coloração das células com anticorpos contra marcadores de superfície expressos por células que se comprometeram com uma linhagem hematopoiética. Estes podem incluir, mas não estão limitados a: CD3 (célula T), CD19 (célula B), CD33 (mieloide), CD56 (célula NK), CD235a (células eritroides), CD71 (células eritroides).

[0031] "Enriquecido" quando usado no contexto da população celular refere-se a uma população celular selecionada com base na presença de um ou mais marcadores, por exemplo, CD34+. Uma população de células, como uma população de células-tronco ou uma população de HSPC, refere-se a células de mamíferos eucarióticos, preferencialmente humanos, isoladas de fontes biológicas, por exemplo, produtos ou tecidos sanguíneos e derivados de mais de uma célula.

[0032] Durante a pré-incubação, um fator sérico pode entrar em contato com uma composição de LNP, antes da entrega para a célula HSPC *in vitro*.

[0033] Algumas modalidades dos métodos *in vitro* compreendem a pré-incubação de um fator sérico e a composição de LNP por cerca de 30 segundos durante a noite. Em algumas modalidades, a etapa de pré-incubação compreende pré-incubar um fator sérico e a composição de LNP por cerca de 1 minuto a 1 hora. Em algumas modalidades, compreende a pré-incubação por cerca de 1-30 minutos. Em outras modalidades, compreende a pré-incubação por cerca de 1-10 minutos. Ainda outras modalidades compreendem a pré-incubação por cerca de 5 minutos. Em certas modalidades, os pontos finais dos intervalos e os valores fornecidos acima podem ser ± 0,5, 1, 2, 3 ou 4 minutos.

[0034] Em certas modalidades, a etapa de pré-incubação ocorre a cerca de 4 °C. Em certas modalidades, a etapa de pré-incubação ocorre a cerca de 25 °C. Em certas modalidades, a etapa de pré-incubação ocorre a cerca de 37 °C. A etapa de pré-incubação pode compreender um tampão, como bicarbonato de sódio ou HEPES. Em certas modalidades, o tampão pode compreender um meio de cultura HSPC. Em modalidades adicionais, o tampão pode consistir em meio de cultura HSPC.

[0035] A pré-incubação de uma composição de LNP com um fator sérico pode compreender pré-incubação com soro, com uma fração sérica ou com um fator sérico isolado. Em algumas modalidades, a composição de LNP é pré-incubada com soro. O soro pode ser de mamífero, camundongo, primata ou soro humano. Em algumas modalidades, a composição de LNP é pré-incubada com um fator sérico isolado. Em certas modalidades, o fator sérico é uma ApoE. Em certas modalidades, o fator sérico é escolhido entre ApoE2, ApoE3 e ApoE4. Em modalidades adicionais, a ApoE é uma proteína recombinante,

como uma proteína humana recombinante. A ApoE pode ser ApoE3 humana recombinante. Pode ser ApoE4 humano recombinante.

[0036] Em algumas modalidades, os métodos compreendem o contato com a célula-tronco, por exemplo, HSPC ou população de células-tronco, por exemplo, população de HSPC, após a etapa de pré-incubação, por exemplo, o contato das células com uma composição de LNP pré-incubada. Em algumas modalidades, os métodos compreendem o contato com a população de células-tronco, como a população de ES ou iPSC após a etapa de pré-incubação, por exemplo, o contato das células com uma composição de LNP pré-incubada. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 1 minuto a cerca de 72 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 1 hora a cerca de 24 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 4 horas a cerca de 24 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 4 horas a cerca de 12 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 2 horas a cerca de 12 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 6 horas a cerca de 8 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 6 horas a cerca de 24 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por cerca de 6 horas a cerca de 24 horas. Em algumas modalidades, os métodos compreendem o contato das células

com uma composição de LNP pré-incubada por cerca de 4 horas a cerca de 12 horas. Em algumas modalidades, os métodos compreendem o contato das células com uma composição de LNP pré-incubada por pelo menos cerca de 0,5, 1, 2, 4, 6, 8, 10 ou 12 horas. Em algumas modalidades, os métodos compreendem uma etapa de lavagem após a etapa de contato. A etapa de lavagem pode compreender meios.

[0037] Em algumas modalidades, os métodos compreendem um mRNA da nuclease Cas. Em algumas modalidades, os métodos compreendem um mRNA da nuclease Cas de Classe 2. Em algumas modalidades, os métodos compreendem um ácido nucleico de gRNA, como um gRNA. Em certas modalidades, os métodos compreendem pelo menos dois ácidos nucleicos de gRNA. Em modalidades adicionais, os métodos compreendem 3 ou mais ácidos nucleicos de gRNA. Em algumas modalidades, um mRNA, como um mRNA da nuclease Cas e um gRNA, é formulado em uma única composição de LNP. Em algumas modalidades, os métodos compreendem um mRNA, como um mRNA de nuclease Cas e um ácido nucleico de gRNA que é co-encapsulado na composição de LNP. Em modalidades adicionais, os métodos compreendem um mRNA e um ácido nucleico de gRNA que são encapsulados separadamente em LNPs. Em certas modalidades, um mRNA é formulado em uma primeira composição de LNP e um ácido nucleico de gRNA é formulado em uma segunda composição de LNP. Em algumas modalidades, a primeira e a segunda composições de LNP são administradas simultaneamente. Em outras modalidades, a primeira e a segunda composições de LNP são administradas sequencialmente. Em algumas modalidades dos métodos *in vitro*, a primeira e a segunda composições de LNP são combinadas antes da etapa de pré-incubação. Em algumas modalidades, a primeira e a segunda composições de LNP são pré-incubadas separadamente.

[0038] Em uma modalidade, uma composição de LNP compreen-

dendo um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2, pode ser administrada a uma célula ou população de células, como, por exemplo, uma população de HSPC ou HSPC, separadamente da administração de uma composição compreendendo um gRNA. Em uma modalidade, uma composição de LNP compreendendo um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2 e um gRNA, pode ser administrada, como a uma população de HSPC ou HSPC, separadamente da administração de um ácido nucleico modelo na célula. Em uma modalidade, uma composição de LNP compreendendo um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2, pode ser administrada, como a uma população de HSPC ou HSPC, seguida pela administração sequencial de uma composição de LNP compreendendo um gRNA e, em seguida, um modelo para a célula ou população. Nas modalidades em que uma composição de LNP compreendendo um mRNA que codifica uma nuclease Cas é administrada antes de uma composição de LNP compreendendo um gRNA, as administrações podem ser separadas por cerca de 4, 6, 8, 12, 24, 36, 48, ou 72 horas; ou cerca de 1, 2, ou 3 dias.

[0039] Em algumas modalidades dos métodos *in vitro* descritos neste documento, a população de células-tronco, HSPC ou HSPC pode ser cultivada *in vitro* após a transfecção através de LNPs.

[0040] Em algumas modalidades, a população de células-tronco transfectadas, HSPC ou HSPC é expandida em um meio de cultura de células-tronco, como um meio de cultura de HSPC. "Expansão" ou "expandido" no contexto de células refere-se a um aumento no número de um tipo de célula característico, ou tipos de células, de uma população inicial de células, que pode ou não ser idêntica. As células iniciais usadas para expansão podem não ser as mesmas que as células geradas a partir da expansão. Algumas modalidades dos métodos *in vitro*

compreendem a cultura da população HSPC ou HSPC em um meio de cultura HSPC. Algumas modalidades compreendem ainda a expansão dos HSPCs em um meio de cultura HSPC que compreende um expansor de células-tronco. Ver, por exemplo, WO2010/059401 (por exemplo, composto do Exemplo 1), WO2013/110198 e WO2017115268, que são incorporados neste documento por referência em relação a compostos adequados para expansão de células-tronco. "Expansor de células-tronco" refere-se a um composto que faz com que células, por exemplo, HSPCs, HSCs e/ou HPCs proliferem, por exemplo, aumentem em número, a uma taxa mais rápida em relação aos mesmos tipos de células ausentes do referido agente. Em um aspecto exemplar, o expansor de células-tronco é um inibidor da através do receptor de aril-hidrocarboneto.

[0041] Em modalidades adicionais, os métodos *in vitro* compreendem ainda alterar o meio de cultura entre as etapas de contato e cultura. Em ainda outras modalidades, a etapa de cultura compreende o meio de cultura celular que inclui trombopoietina (Tpo), ligante Flt3 (Flt-3L) e fator de células-tronco humanas (SCF). Em modalidades, o meio de cultura de células inclui ainda interleucina-6 humana (IL-6). Nas modalidades, o meio de cultura de células inclui trombopoietina (Tpo), ligante Flt3 (Flt-3L) e fator de células-tronco humanas (SCF).

Carga CRISPR/Cas

[0042] A carga CRISPR/Cas entregue através da formulação de LNP inclui uma molécula de mRNA que codifica uma proteína de interesse. Por exemplo, um mRNA para expressar uma proteína como a proteína fluorescente verde (GFP) e o agente de ligação ao DNA guiado por RNA, ou uma nuclease Cas, está incluída. São fornecidas composições de LNP que incluem um mRNA de nuclease Cas, por exemplo, um mRNA da nuclease de Classe 2 que permite a expressão em uma célula de uma proteína Cas9. Além disso, a carga pode conter um

ou mais RNAs guia ou ácidos nucleicos que codificam RNAs guia. Um ácido nucleico modelo, por exemplo, para reparo ou recombinação, também pode ser incluído na composição ou um ácido nucleico modelo pode ser utilizado nos métodos descritos neste documento.

[0043] "mRNA" refere-se a um polinucleotídeo que não é DNA e que compreende um quadro de leitura aberto que pode ser traduzido em um polipeptídeo (isto é, pode servir como um substrato para tradução por um ribossomo e tRNAs amino-acilados). O mRNA pode compreender uma estrutura de fosfato-açúcar incluindo resíduos de ribose ou seus análogos, por exemplo, resíduos de 2'-metóxi-ribose. Em algumas modalidades, os açúcares de uma estrutura de mRNA fosfato-açúcar consistem essencialmente em resíduos de ribose, resíduos de 2'-metóxi-ribose ou uma combinação destes. Em geral, os mRNAs não contêm uma quantidade substancial de resíduos de timidina (por exemplo, 0 resíduos ou menos que 30, 20, 10, 5, 4, 3 ou 2 resíduos de timidina; ou menos que 10%, 9%, 8%, 7%, 6%, 5%, 4%, 4%, 3%, 2%, 1%, 0,5%, 0,2% ou 0,1% de timidina). Um mRNA pode conter uridinas modificadas em algumas ou todas as suas posições de uridina.

Sistemas Nuclease CRISPR/Cas

[0044] Um componente das formulações divulgadas é um mRNA que codifica o agente de ligação ao DNA guiado por RNA, como uma nuclease Cas.

[0045] Como usado neste documento, um "agente de ligação a DNA guiado por RNA" significa um polipeptídeo ou complexo de polipeptídeos com atividade de ligação a RNA e DNA, ou uma subunidade de ligação a DNA desse complexo, em que a atividade de ligação a DNA é específica da sequência e depende da sequência do RNA. Agentes de ligação a DNA guiados por RNA exemplificativos incluem cleavases/nickases Cas e suas formas inativadas ("agentes de ligação ao DNA dCas"). "Nuclease Cas", como utilizado neste documento,

abrange agentes de ligação ao DNA cleavases, nickases Cas e dCas. Os agentes de ligação ao DNA cleavases/nickases e dCas incluem um complexo Csm ou Cmr de um sistema CRISPR tipo III, a subunidade Cas10, Csm1 ou Cmr2, um complexo em Cascata de um sistema CRISPR tipo I, a subunidade Cas3 e a nucleases Cas de Classe 2. Como utilizado neste documento, uma "nuclease Cas de Classe 2" é um polipeptídeo de cadeia única com atividade de ligação ao DNA guiada por RNA. As nucleases Cas de Classe 2 incluem cleavases/nickases Cas da Classe 2 (por exemplo, variantes H840A, D10A ou N863A), que possuem ainda atividade de cleavases ou nickases de DNA guiada por RNA e agentes de ligação de DNA dCas de Classe 2, nas quais a atividade da cleavases/nickases é inativado. As nucleases Cas de Classe 2 incluem, por exemplo, Cas9, Cpf1, C2c1, C2c2, C2c3, HF Cas9 (por exemplo, variantes N497A, R661A, Q695A, Q926A), HypaCas9 (por exemplo, variantes N692A, M694A, Q695A, H698A), eSPCas9 (1,0) (por exemplo, variantes K810A, K1003A, R1060A) e eSPCas9 (1,1) (por exemplo, variantes K848A, K1003A, R1060A) e suas modificações. A proteína Cpf1, Zetsche et al., *Cell*, 163: 1-13 (2015), é homóloga a Cas9 e contém um domínio de nuclease semelhante a RuvC. As sequências Cpf1 de Zetsche são incorporadas por referência na sua totalidade. *Veja*, Por exemplo, Zetsche, Tabelas S1 e S3. Ver, , Makarova et al., *Nat Rev Microbiol*, 13(11): 722-36 (2015); Shmakov et al., *Molecular Cell,* 60:385-397 (2015).

[0046] Em algumas modalidades, o agente de ligação ao DNA guiado por RNA é uma nuclease Cas de Classe 2. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA possui atividade de cleavase, que também pode ser referida como atividade de endonuclease de fita dupla. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA compreende uma nuclease Cas, como uma nuclease Cas de Classe 2 (que pode ser, por exemplo, uma nuclease

Cas do Tipo II, V ou VI). As nucleases Cas de Classe 2 incluem, por exemplo, proteínas Cas9, Cpf1, C2c1, C2c2 e C2c3 e suas modificações. Exemplos de nucleases Cas9 incluem aquelas dos sistemas CRISPR do tipo II de *S. pyogenes, S. aureus*e outros procariontes (ver, por exemplo, a lista no próximo parágrafo) e versões modificadas (por exemplo, manipulado ou mutantes). Ver, por exemplo, US2016/0312198 A1; US 2016/0312199 A1. Outros exemplos de nucleases Cas incluem um complexo Csm ou Cmr de um sistema CRISPR tipo III ou a sua subunidade Cas10, Csm1 ou Cmr2; e um complexo Cascade de um sistema CRISPR tipo I ou a sua subunidade Cas3. Em algumas modalidades, a nuclease Cas pode ser de um sistema Tipo IIA, Tipo IIB ou Tipo IIC. Para discussão de vários sistemas CRISPR e nucleases Cas, ver, por exemplo, Makarova et al., Nat. Rev. Microbiol. 9:467-477 (2011); Makarova et al., Nat. Rev. Microbiol, 13: 722-36 (2015); Shmakov et al., Molecular Cell, 60:385-397 (2015).

[0047] As espécies exemplares não limitantes das quais a nuclease de Cas pode ser derivada incluem *Streptococcus pyogenes, Streptococcus thermophilus, Streptococcus sp., Staphylococcus aureus, Listeria innocua, Lactobacillus gasseri, Francisella novicida, Wolinella succinogenes, Sutterella wadsworthensis, Gammaproteobacterium, Neisseria meningitidis, Campylobacter jejuni, Pasteurella multocida, Fibrobacter succinogene, Rhodospirillum rubrum, Nocardiopsis dassonvillei, Streptomyces pristinaespiralis, Streptomyces viridochromogenes, Streptomyces viridochromogenes, Streptosporangium roseum, Streptosporangium roseum, Alicyclobacillus acidocaldarius, Bacillus pseudomycoides, Bacillus selenitireducens, Exiguobacterium sibiricum, Lactobacillus delbrueckii, Lactobacillus salivarius, Lactobacillus buchneri, Treponema denticola, Microscilla marina, Burkholderiales bacterium, Polaromonas naphthalenivorans, Polaromonas sp., Crocosphaera watsonii, Cyanothece sp., Microcystis aeruginosa, Synechococcus*

sp., Acetohalobium arabaticum, Ammonifex degensii, Caldicelulosiruptor becscii, Candidatus Desulforudis, Clostridium botulinum, Clostridium difficile, Finegoldia magna, Natranaerobius thermophilus, Pelotomaculum thermopropionicum, Acidithiobacillus caldus, Acidithiobacillus ferrooxidans, Allochromatium vinosum, Marinobacter sp., Nitrosococcus halophilus, Nitrosococcus watsoni, Pseudoalteromonas haloplanktis, Ktedonobacter racemifer, Methanohalobium evestigatum, Anabaena variabilis, Nodularia spumigena, Nostoc sp., Arthrospira maxima, Arthrospira platensis, Arthrospira sp., Lyngbya sp., Microcoleus chthonoplastes, Oscillatoria sp., Petrotoga mobilis, Thermosipho africanus, Streptococcus pasteurianus, Neisseria cinerea, Campylobacter lari, Parvibaculum lavamentivorans, Corynebacterium diphtheria, Acidaminococcus sp., Lachnospiraceae bacterium ND2006, e *Acaryochloris marina*.

[0048] Em algumas modalidades, a nuclease Cas é a nuclease Cas9 de *Streptococcus pyogenes*. Em algumas modalidades, a nuclease Cas é a nuclease Cas9 de *Streptococcus thermophilus*. Em algumas modalidades, a nuclease Cas é a nuclease Cas9 de *Neisseria meningitidis*. Em algumas modalidades, a nuclease Cas é a nuclease Cas9 é de *Staphylococcus aureus*. Em algumas modalidades, a nuclease Cas é a nuclease Cpf1 de *Francisella novicida*. Em algumas modalidades, a nuclease Cas é a nuclease Cpf1 de *Acidaminococcus sp.* Em algumas modalidades, a nuclease Cas é a nuclease Cpf1 da *bactéria* Lachnospiraceae ND2006. Em outras modalidades, a nuclease Cas é a nuclease Cpf1 de *Francisella tularensis, bactéria Lachnospiraceae, Butyrivibrio proteoclasticus, bactéria Peregrinibacteria, bactéria Parcubacteria, Smithella, Acidaminococcus, Candidatus Methanoplasma termitum, Eubacterium previraens, Moraxella bovoculi, Moraxella bovoculi* ou *Porphyromonas macacae*. Em certas modalidades, a nuclease Cas é uma nuclease Cpf1 de um *Acidaminococcus* ou

Lachnospiraceae.

[0049] O Cas9 do tipo selvagem possui dois domínios de nuclease: RuvC e HNH. O domínio RuvC quebra a fita de DNA não alvo e o domínio HNH quebra a fita alvo de DNA. Em algumas modalidades, a nuclease Cas9 compreende mais de um domínio RuvC e/ou mais de um domínio HNH. Em algumas modalidades, a nuclease Cas9 é um Cas9 do tipo selvagem. Em algumas modalidades, o Cas9 é capaz de induzir uma quebra de fita dupla no DNA alvo. Em certas modalidades, a nuclease Cas pode clivar dsDNA, pode clivar uma fita de dsDNA ou pode não ter atividade de cleavase ou nickase de DNA. Uma sequência de aminoácidos Cas9 exemplar é fornecida como SEQ ID NO: 3. Uma sequência exemplar de ORF mRNA de Cas9, que inclui códons de iniciação e parada, é fornecida como SEQ ID NO: 4. Uma sequência de codificação de mRNA Cas9 exemplar, adequada para inclusão em uma proteína de fusão, é fornecida como SEQ ID NO: 10.

[0050] Em algumas modalidades, são utilizadas nucleases quiméricas de Cas, em que um domínio ou região da proteína é substituído por uma porção de uma proteína diferente. Em algumas modalidades, um domínio de nuclease Cas pode ser substituído por um domínio de uma nuclease diferente, como Fok1. Em algumas modalidades, uma nuclease Cas pode ser uma nuclease modificada.

[0051] Em outras modalidades, a nuclease Cas pode ser de um sistema CRISPR/Cas Tipo I. Em algumas modalidades, a nuclease Cas pode ser um componente do complexo Cascade de um sistema CRISPR/Cas Tipo I. Em algumas modalidades, a nuclease Cas pode ser uma proteína Cas3. Em algumas modalidades, a nuclease Cas pode ser de um sistema CRISPR/Cas Tipo III. Em algumas modalidades, a nuclease Cas pode ter uma atividade de clivagem de RNA.

[0052] Em algumas modalidades, o agente de ligação ao DNA guiado por RNA possui atividade de nickase de fita simples, ou seja,

pode cortar uma fita de DNA para produzir uma quebra de fita única, também conhecida como "nick." Em algumas modalidades, o agente de ligação ao DNA guiado por RNA compreende uma nickase Cas. Uma nickase é uma enzima que cria um nick no dsDNA, ou seja, corta uma fita, mas não a outra da dupla hélice do DNA. Em algumas modalidades, uma nickase Cas é uma versão de uma nuclease Cas (por exemplo, uma nuclease Cas discutida acima) na qual um sítio ativo endonucleolítico é inativado, por exemplo, por uma ou mais alterações (por exemplo, mutações pontuais) em um domínio catalítico. Ver, por exemplo, Pat. US N°. 8.889.356 para discussão de nickases Cas e alterações exemplares no domínio catalítico. Em algumas modalidades, uma nickase Cas como uma nickase Cas9 tem um domínio RuvC ou HNH inativado.

[0053] Em algumas modalidades, o agente de ligação ao DNA guiado por RNA é modificado para conter apenas um domínio de nuclease funcional. Por exemplo, a proteína do agente pode ser modificada de modo que um dos domínios da nuclease seja mutado ou completamente ou parcialmente deletado para reduzir sua atividade de clivagem de ácido nucleico. Em algumas modalidades, uma nickase é usada tendo um domínio RuvC com atividade reduzida. Em algumas modalidades, uma nickase é usada tendo um domínio RuvC inativo. Em algumas modalidades, uma nickase é usada tendo um domínio HNH com atividade reduzida. Em algumas modalidades, uma nickase é usada tendo um domínio HNH inativo.

[0054] Em algumas modalidades, um aminoácido conservado dentro de um domínio de nuclease da proteína Cas é substituído para reduzir ou alterar a atividade da nuclease. Em algumas modalidades, uma nuclease Cas pode compreender uma substituição de aminoácidos no domínio da nuclease RuvC ou do tipo RuvC. Substituições de aminoácidos exemplares no domínio nuclease RuvC ou tipo RuvC in-

cluem D10A (com base na proteína Cas9 de S. *pyogenes*). Ver, Por exemplo, Zetsche et al. (2015) *Cell* 22 de outubro: 163 (3): 759-771. Em algumas modalidades, a nuclease Cas pode compreender uma substituição de aminoácidos no domínio da nuclease HNH ou do tipo HNH. Substituições de aminoácidos exemplares no domínio nuclease HNH ou do tipo HNH incluem E762A, H840A, N863A, H983A e D986A (com base na proteína Cas9 de S. *pyogenes*). *Ver*, Por exemplo, Zetsche et al. (2015). Outras substituições de aminoácidos exemplares incluem D917A, E1006A e D1255A (com base na sequência de *Francisella novicida* U112 Cpf1 (FnCpf1) (UniProtKB - A0Q7Q2 (CPF1_FRATN)).

[0055] Em algumas modalidades, um mRNA que codifica uma nickase é fornecido em combinação com um par de RNAs guia que são complementares às fitas antisense e sense da sequência alvo, respectivamente. Nesta modalidade, os RNAs de guia direcionam a nickase para uma sequência alvo e introduzem um DSB gerando um nick em fitas opostas da sequência alvo (isto é, duplo corte). Em algumas modalidades, o uso de duplo corte pode melhorar a especificidade e reduzir os efeitos fora do alvo. Em algumas modalidades, uma nickase é usada junto com dois RNAs guia separados, direcionando fitas opostas de DNA, para produzir uma nick dupla no DNA alvo. Em algumas modalidades, uma nickase é usada junto com dois RNAs guia separados que são selecionados para estarem próximos para produzir uma nick dupla no DNA alvo.

[0056] Em algumas modalidades, o agente de ligação ao DNA guiado por RNA carece de atividade de cleavase e nickase. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA compreende um polipeptídeo de ligação ao DNA dCas. Um polipeptídeo dCas possui atividade de ligação ao DNA, enquanto essencialmente carece de atividade catalítica (cleavase/nickase). Em algumas modalidades, o polipeptídeo dCas é um polipeptídeo dCas9. Em algumas

modalidades, o agente de ligação ao DNA guiado por RNA sem atividade de cleavase e nickase ou o polipeptídeo de ligação a DNA é uma versão de uma nuclease Cas (por exemplo, uma nuclease Cas discutida acima) na qual seus sítios ativos endonucleolíticos são inativados, por exemplo, por uma ou mais alterações (por exemplo, mutações pontuais) em seus domínios catalíticos. Ver, por exemplo, US 2014/0186958 A1; US 2015/0166980 A1.

[0057] Em algumas modalidades, o agente de ligação ao DNA guiado por RNA compreende um ou mais domínios funcionais heterólogos (por exemplo, é ou compreende um polipeptídeo de fusão).

[0058] Em algumas modalidades, o domínio funcional heterólogo pode facilitar o transporte do agente de ligação ao DNA guiado por RNA para o núcleo de uma célula. Por exemplo, o domínio funcional heterólogo pode ser um sinal de localização nuclear (NLS). Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com 1-10 NLS(s). Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com 1-5 NLS(s). Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com um NLS. Quando um NLS é usado, o NLS pode ser ligado no terminal N ou no terminal C da sequência do agente de ligação ao DNA guiada por RNA. Também pode ser inserido na sequência do agente de ligação ao DNA guiado por RNA. Em outras modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com mais de um NLS. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com 2, 3, 4 ou 5 NLSs. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com dois NLSs. Em certas circunstâncias, os dois NLSs podem ser os mesmos (Por exemplo, dois SV40 NLSs) ou diferentes. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA é fundido a duas sequências SV40 NLS ligadas no

terminal carbóxi. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com dois NLSs, um ligado no terminal N e um no terminal C. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido com 3 NLSs. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser fundido sem NLS. Em algumas modalidades, o NLS pode ser uma sequência monopartida, como, Por exemplo, o SV40 NLS, PKK-KRKV ou PKKKRRV. Em algumas modalidades, o NLS pode ser uma sequência bipartida, como o NLS da nucleoplasmina, KRPAATK-KAGQAKKKK. Em uma modalidade específica, um único PKKKRKV NLS pode ser ligado no terminal C do agente de ligação ao DNA guiado por RNA. Um ou mais ligantes são opcionalmente incluídos no sítio de fusão.

[0059] Em algumas modalidades, o domínio funcional heterólogo pode ser capaz de modificar a meia-vida intracelular do agente de ligação ao DNA guiado por RNA. Em algumas modalidades, a meia-vida do agente de ligação ao DNA guiado por RNA pode ser aumentada. Em algumas modalidades, a meia-vida do agente de ligação ao DNA guiado por RNA pode ser reduzida. Em algumas modalidades, o domínio funcional heterólogo pode ser capaz de aumentar a estabilidade do agente de ligação ao DNA guiado por RNA. Em algumas modalidades, o domínio funcional heterólogo pode ser capaz de reduzir a estabilidade do agente de ligação ao DNA guiado por RNA. Em algumas modalidades, o domínio funcional heterólogo pode atuar como um peptídeo sinal para degradação de proteínas. Em algumas modalidades, a degradação da proteína pode ser mediada por enzimas proteolíticas, como, por exemplo, proteassomas, proteases lisossômicas ou proteases de calpaína. Em algumas modalidades, o domínio funcional heterólogo pode compreender uma sequência PEST. Em algumas modalidades, o agente de ligação ao DNA guiado por RNA pode ser

modificado pela adição de ubiquitina ou uma cadeia de polubiquitina. Em algumas modalidades, a ubiquitina pode ser uma proteína do tipo ubiquitina (UBL). Exemplos não limitativos de proteínas do tipo ubiquitina incluem modificador pequeno do tipo ubiquitina (SUMO), proteína reativa cruzada da ubiquitina (UCRP, também conhecida como gene 15 estimulado por interferon (ISG15)), modificador 1 relacionado à ubiquitina (URM1), proteína 8 desregulada no desenvolvimento expressa em células precursoras dos neurônios (NEDD8, também chamada Rub1 no *S. cerevisiae*), associada ao antígeno leucocitário humano F (FAT10), autofagia-8 (ATG8) e -12 (ATG12), Fau proteína do tipo ubiquitina (FUB1), UBL ancorado na membrana (MUB), modificador 1 de dobra de ubiquitina (UFM1) e proteína 5 do tipo de ubiquitina (UBL5).

[0060] Em algumas modalidades, o domínio funcional heterólogo pode ser um domínio marcador. Exemplos não limitativos de domínios marcadores incluem proteínas fluorescentes, marcadores de purificação, marcadores epítopos e sequências de genes repórteres. Em algumas modalidades, o domínio marcador pode ser uma proteína fluorescente. Exemplos não limitativos de proteínas fluorescentes adequadas incluem proteínas fluorescentes verdes (Por exemplo, GFP, GFP-2, tagGFP, turboGFP, sfGFP, EGFP, Esmeralda, Verde Azami, Verde Azami Monomérico, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarelas (Por exemplo, YFP, EYFP, Citrino, Vênus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azuis (Por exemplo, EBFP, EBFP2, Azurita, mKalamal, GFPuv, safira, safira-T), proteínas fluorescentes cianas (Por exemplo, ECFP, cerúleo, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes vermelhas (Por exemplo, mKate, mKate2, mPlum, monômero DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-monômero, HcRed-Tandem, HcRed1, AsRed2, eqFP611, eRasberry, mStrawberry, Jred) e proteínas fluorescentes laranja (mOrange, mKO, Kusabira-Orange, Komabira-Mono-

mérica-Orange, mTangerina, tdTomato) ou qualquer outra proteína fluorescente adequada. Em outras modalidades, o domínio do marcador pode ser uma tag de purificação e/ou uma tag de epítopo. Tags exemplares não limitantes incluem glutationa-S-transferase (GST), proteína de ligação à quitina (CBP), proteína de ligação à maltose (MBP), tioredoxina (TRX), poli(NANP), tag de purificação por afinidade em tandem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, 8xHis, proteína carreadora de biotina carboxil (BCCP), poli-His e calmodulina. Genes repórter exemplares não limitativos incluem glutationa-S-transferase (GST), peroxidase de rábano silvestre (HRP), cloranfenicol acetiltransferase (CAT), beta-galactosidase, beta-glucuronidase, luciferase ou proteínas fluorescentes.

[0061] Em modalidades adicionais, o domínio funcional heterólogo pode direcionar o agente de ligação ao DNA guiado por RNA para uma organela, tipo de célula, tecido ou órgão específico. Em algumas modalidades, o domínio funcional heterólogo pode direcionar o agente de ligação ao DNA guiado por RNA para mitocôndrias.

[0062] Em outras modalidades, o domínio funcional heterólogo pode ser um domínio efetor. Quando o agente de ligação ao DNA guiado por RNA é direcionado para sua sequência alvo, Por exemplo, quando uma nuclease Cas é direcionada para uma sequência alvo por um gRNA, o domínio efetor pode modificar ou afetar a sequência alvo. Em algumas modalidades, o domínio efetor pode ser escolhido entre um domínio de ligação de ácido nucleico, um domínio de nuclease (por exemplo, um domínio de nuclease não-Cas), um domínio de modificação epigenética, um domínio de ativação transcricional ou um domínio repressor da transcrição. Em algumas modalidades, o domínio funcional heterólogo é uma nuclease, como uma nuclease FokI. Ver, por exemplo, Pat. US N°. 9.023.649. Em algumas modalidades, o domínio

funcional heterólogo é um ativador ou repressor da transcrição. Veja, por exemplo, Qi et al., "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression," *Cell* 152:1173-83 (2013); Perez-Pinera et al., "RNA-guided gene activation by CRISPR-Cas9-based transcription factors", *Nat. Methods* 10:973-6 (2013); Mali et al., "CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering, "*Nat. Biotechnol.* 31:833-8 (2013); Gilbert et al., "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes", *Cell* 154:442-51 (2013). Como tal, o agente de ligação ao DNA guiado por RNA torna-se essencialmente um fator de transcrição que pode ser direcionado para ligar uma sequência alvo desejada usando um RNA guia. Em certas modalidades, o domínio de modificação do DNA é um domínio de metilação, como um domínio de desmetilação ou metiltransferase. Em certas modalidades, o domínio efetor é um domínio de modificação de DNA, como um domínio de edição de base. Em modalidades particulares, o domínio de modificação do DNA é um domínio de edição de ácido nucleico que introduz uma modificação específica no DNA, como um domínio desaminase. Ver, Por exemplo, WO 2015/089406; US 2016/0304846. Os domínios de edição de ácido nucleico, domínios desaminase e variantes de Cas9 descritos em WO 2015/089406 e US 2016/0304846 são incorporados por referência neste documento.

[0063] A nuclease pode compreender pelo menos um domínio que interage com um RNA guia ("gRNA"). Além disso, a nuclease pode ser direcionada para uma sequência alvo por um gRNA. Nos sistemas de nuclease Cas de Classe 2, o gRNA interage com a nuclease e com a sequência alvo, de modo que direciona a ligação à sequência alvo. Em algumas modalidades, o gRNA fornece a especificidade para a clivagem direcionada, e a nuclease pode ser universal e emparelhada com

diferentes gRNAs para clivar diferentes sequências alvo. A nuclease Cas de classe 2 pode emparelhar com uma estrutura de andaime de gRNA dos tipos, ortólogos e espécies exemplares listadas acima.

Guia RNA (gRNA)

[0064] Em algumas modalidades da presente invenção, a carga para a formulação de LNP inclui pelo menos um gRNA. O gRNA pode guiar a nuclease Cas ou nuclease Cas de Classe 2 para uma sequência alvo em uma molécula de ácido nucleico alvo. Em algumas modalidades, um gRNA se liga e fornece especificidade de clivagem por uma nuclease Cas de Classe 2. Em algumas modalidades, o gRNA e a nuclease Cas podem formar uma ribonucleoproteína (RNP), Por exemplo, um complexo CRISPR/Cas, como um complexo CRISPR/Cas9. Em algumas modalidades, o complexo CRISPR/Cas pode ser um complexo CRISPR/Cas9 Tipo II. Em algumas modalidades, o complexo CRISPR/Cas pode ser um complexo CRISPR/Cas Tipo V, como um complexo Cpf1/RNA guia. As nucleases Cas e os gRNAs cognatos podem ser emparelhados. As estruturas em andaime de gRNA que se emparelham com cada nuclease Cas de Classe 2 variam com o sistema CRISPR/Cas específico.

[0065] "RNA guia", "gRNA" e simplesmente "guia" são usados neste documento de forma intercambiável para se referir a um crRNA (também conhecido como CRISPR RNA) ou à combinação de um crRNA e um trRNA (também conhecido como tracrRNA). O crRNA e o trRNA podem ser associados como uma única molécula de RNA (RNA guia único, sgRNA) ou em duas moléculas de RNA separadas (RNA guia duplo, dgRNA). "RNA guia" ou "gRNA" refere-se a cada tipo. O trRNA pode ser uma sequência que ocorre naturalmente ou uma sequência de trRNA com modificações ou variações em comparação com as sequências que ocorrem naturalmente.

[0066] Como utilizado neste documento, uma "sequência guia" re-

fere-se a uma sequência dentro de um RNA guia que é complementar a uma sequência alvo e funciona para direcionar um RNA guia para uma sequência alvo para ligação ou modificação (por exemplo, clivagem) por um agente de ligação DNA guiado por RNA. Uma "sequência guia" também pode ser referida como "sequência alvo" ou "sequência espaçadora." Uma sequência guia pode ter 20 pares de bases de comprimento, por exemplo, no caso de *Streptococcus pyogenes* (por exemplo, Spy Cas9) e homólogos/ortólogos Cas9 relacionados. Sequências mais curtas ou mais longas também podem ser usadas como guias, por exemplo, 15-, 16-, 17-, 18-, 19-, 21-, 22-, 23-, 24- ou 25-nucleotídeos de comprimento. Em algumas modalidades, a sequência alvo está em um gene ou em um cromossomo, por exemplo, e é complementar à sequência guia. Em algumas modalidades, o grau de complementaridade ou identidade entre uma sequência guia e sua sequência alvo correspondente pode ser de cerca de 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a sequência guia e a região alvo podem ser 100% complementares ou idênticas. Em outras modalidades, a sequência guia e a região alvo podem conter pelo menos uma incompatibilidade. Por exemplo, a sequência guia e a sequência alvo podem conter 1, 2, 3 ou 4 desencontros, em que o comprimento total da sequência alvo é pelo menos 17, 18, 19, 20 ou mais pares de bases. Em algumas modalidades, a sequência guia e a região alvo podem conter 1-4 incompatibilidades, onde a sequência guia compreende pelo menos 17, 18, 19, 20 ou mais nucleotídeos. Em algumas modalidades, a sequência guia e a região alvo podem conter 1, 2, 3 ou 4 desencontros, em que a sequência guia compreende 20 nucleotídeos.

[0067] As sequências alvo das proteínas Cas incluem as fitas positivas e negativas do DNA genômico (ou seja, a sequência fornecida e o complemento inverso da sequência), pois um substrato de ácido nu-

cleico para uma proteína Cas é um ácido nucleico de fita dupla. Por conseguinte, quando se diz que uma sequência guia é "complementar a uma sequência alvo", deve ser entendido que a sequência guia pode direcionar um RNA guia para se ligar ao complemento reverso de uma sequência alvo. Assim, em algumas modalidades, onde a sequência guia liga o complemento reverso de uma sequência alvo, a sequência guia é idêntica a certos nucleotídeos da sequência alvo (por exemplo, a sequência alvo que não inclui o PAM), exceto a substituição de U por T na sequência guia.

[0068] O comprimento da sequência de direcionamento pode depender do sistema CRISPR/Cas e dos componentes utilizados. Por exemplo, diferentes nucleases Cas de Classe 2 de diferentes espécies bacterianas têm diferentes comprimentos ótimos de sequência de direcionamento. Por conseguinte, a sequência de direcionamento pode compreender 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 ou mais de 50 nucleotídeos de comprimento. Em algumas modalidades, o comprimento da sequência de direcionamento é 0, 1, 2, 3, 4 ou 5 nucleotídeos maior ou menor que a sequência guia de um sistema CRISPR/Cas de ocorrência natural. Em certas modalidades, a nuclease Cas e a estrutura em andaime gRNA serão derivados do mesmo sistema CRISPR/Cas. Em algumas modalidades, a sequência de direcionamento pode compreender ou consistir em 18-24 nucleotídeos. Em algumas modalidades, a sequência de direcionamento pode compreender ou consistir em 19-21 nucleotídeos. Em algumas modalidades, a sequência de direcionamento pode compreender ou consistir em 20 nucleotídeos.

[0069] Em algumas modalidades, o sgRNA é um "sgRNA Cas9" capaz de mediar a clivagem de DNA guiada por RNA por uma proteína Cas9. Em algumas modalidades, o sgRNA é um "sgRNA Cpf1" capaz de mediar a clivagem de DNA guiada por RNA por uma proteína Cpf1.

Em certas modalidades, o gRNA compreende um RNA crRNA e tracr suficiente para formar um complexo ativo com uma proteína Cas9 e mediar a clivagem de DNA guiada por RNA. Em certas modalidades, o gRNA compreende um crRNA suficiente para formar um complexo ativo com uma proteína Cpf1 e mediar a clivagem de DNA guiada por RNA. *Veja* Zetsche 2015.

[0070] Certas modalidades da invenção também fornecem ácidos nucleicos, Por exemplo, cassetes de expressão, que codificam o gRNA descrito neste documento. Um "ácido nucleico de RNA guia" é usado neste documento para se referir a um RNA guia (Por exemplo, um sgRNA ou um dgRNA) e um cassete de expressão de RNA guia, que é um ácido nucleico que codifica um ou mais RNAs guia.

[0071] Em algumas modalidades, o ácido nucleico pode ser uma molécula de DNA. Em algumas modalidades, o ácido nucleico pode compreender uma sequência de nucleotídeos que codifica um crRNA. Em algumas modalidades, a sequência de nucleotídeos que codifica o crRNA compreende uma sequência de direcionamento flanqueada por toda ou uma porção de uma sequência de repetição de um sistema CRISPR/Cas de ocorrência natural. Em algumas modalidades, o ácido nucleico pode compreender uma sequência de nucleotídeos que codifica um RNA tracr. Em algumas modalidades, o crRNA e o RNA tracr podem ser codificados por dois ácidos nucleicos separados. Em outras modalidades, o crRNA e o RNA tracr podem ser codificados por um único ácido nucleico. Em algumas modalidades, o crRNA e o RNA tracr podem ser codificados por cadeias opostas de um único ácido nucleico. Em outras modalidades, o crRNA e o RNA tracr podem ser codificados pela mesma fita de um único ácido nucleico. Em algumas modalidades, o ácido nucleico do gRNA codifica um sgRNA. Em algumas modalidades, o ácido nucleico do gRNA codifica um sgRNA da nuclease Cas9. Nas modalidades que se seguem, o ácido nucleico do

gRNA codifica um sgRNA da nuclease de Cpf1.

[0072] A sequência nucleotídica que codifica o RNA guia pode ser operacionalmente ligada a pelo menos uma sequência de controle transcricional ou regulatória, como um promotor, um 3'-UTR ou um 5'-UTR. Em um exemplo, o promotor pode ser um promotor de tRNA, por exemplo, $tRNA^{Lys3}$ ou uma quimera de tRNA. *Veja* Mefferd et al., *RNA*. 2015 21:1683-9; Scherer et al., *Nucleic Acids Res*. 2007 35: 2620–2628. Em certas modalidades, o promotor pode ser reconhecido pela RNA polimerase III (Pol III). Exemplos não limitativos de promotores de Pol III também incluem promotores U6 e H1. Em algumas modalidades, a sequência nucleotídica que codifica o RNA guia pode ser operacionalmente ligada a um camundongo ou promotor U6 humano. Em algumas modalidades, o ácido nucleico do gRNA é um ácido nucleico modificado. Em certas modalidades, o ácido nucleico de gRNA inclui um nucleosídeo ou nucleotídeo modificado. Em algumas modalidades, o ácido nucleico do gRNA inclui uma modificação na extremidade 5', por exemplo, um nucleosídeo ou nucleotídeo modificado para estabilizar e impedir a integração do ácido nucleico. Em algumas modalidades, o ácido nucleico de gRNA compreende um DNA de fita dupla tendo uma modificação na extremidade 5' em cada fita. Em certas modalidades, o ácido nucleico do gRNA inclui um didesóxi-T invertido ou um nucleosídeo ou nucleotídeo abásico invertido como a modificação da extremidade 5'. Em algumas modalidades, o ácido nucleico do gRNA inclui um marcador como biotina, destiobioteno-TEG, digoxigenina e marcadores fluorescentes, incluindo, por exemplo, FAM, ROX, TAMRA e AlexaFluor.

[0073] Em certas modalidades, mais de um ácido nucleico de gRNA, como um gRNA, pode ser usado com um sistema de nuclease CRISPR/Cas. Cada ácido nucleico de gRNA pode conter uma sequência de direcionamento diferente, de modo que o sistema CRISPR/Cas cliva mais de uma sequência alvo. Em algumas modalidades, um ou

mais gRNAs podem ter propriedades iguais ou diferentes, como atividade ou estabilidade dentro de um complexo CRISPR/Cas. Onde mais de um gRNA é usado, cada gRNA pode ser codificado no mesmo ou em diferentes ácidos nucleicos de gRNA. Os promotores utilizados para conduzir a expressão de mais de um gRNA podem ser iguais ou diferentes.

RNAs modificados

[0074] Em certas modalidades, as composições de LNP compreendem RNAs modificados.

[0075] Os nucleosídeos ou nucleotídeos modificados podem estar presentes em um RNA, por exemplo, um gRNA ou mRNA. Um gRNA ou mRNA compreendendo um ou mais nucleosídeos ou nucleotídeos modificados, por exemplo, é chamado de RNA "modificado" para descrever a presença de um ou mais componentes ou configurações de ocorrência não naturais e/ou naturais que são usados em vez de ou além disso aos resíduos canônicos A, G, C e U. Em algumas modalidades, um RNA modificado é sintetizado com um nucleosídeo ou nucleotídeo não canônico, aqui chamado "modificado."

[0076] Os nucleosídeos e nucleotídeos modificados podem incluir um ou mais dos seguintes: (i) alteração, Por exemplo, substituição, de um ou de ambos os oxigênios de fosfato não-ligados e/ou de um ou mais dos oxigênios de fosfato de ligação na ligação da estrutura principal fosfodiéster (uma modificação da estrutura principal exemplar); (ii) alteração, Por exemplo, substituição, de um constituinte do açúcar ribose, Por exemplo, do hidroxil 2' no açúcar ribose (uma modificação exemplar no açúcar); (iii) substituição por atacado da fração fosfato por ligantes "defosfo" (uma modificação exemplar da estrutura principal); (iv) modificação ou substituição de uma nucleobase de ocorrência natural, incluindo uma nucleobase não canônica (uma modificação de base exemplar); (v) substituição ou modificação da estrutura principal ribose-fosfato (uma modificação exemplar da estrutura principal); (vi)

modificação da extremidade 3' ou da extremidade 5' do oligonucleotídeo, Por exemplo, remoção, modificação ou substituição de um grupo fosfato terminal ou conjugação de uma fração, cap ou ligante (tais modificações de cap 3' ou 5' podem compreender uma modificação de açúcar e/ou estrutura principal); e (vii) modificação ou substituição do açúcar (uma modificação exemplar do açúcar). Certas modalidades compreendem uma modificação na extremidade 5' de um mRNA, gRNA ou ácido nucleico. Certas modalidades compreendem uma modificação na extremidade 3' de um mRNA, gRNA ou ácido nucleico. Um RNA modificado pode conter modificações nas extremidades 5' e 3'. Um RNA modificado pode conter um ou mais resíduos modificados em locais não terminais. Em certas modalidades, um gRNA inclui pelo menos um resíduo modificado. Em certas modalidades, um mRNA inclui pelo menos um resíduo modificado.

[0077] Como utilizado neste documento, uma primeira sequência é considerada "compreendendo uma sequência com pelo menos X% de identidade para" uma segunda sequência se um alinhamento da primeira sequência com a segunda sequência mostrar que X% ou mais das posições da segunda sequência em sua totalidade é correspondida pela primeira sequência. Por exemplo, a sequência AAGA compreende uma sequência com 100% de identidade com a sequência AAG porque um alinhamento daria 100% de identidade, pois existem correspondências para todas as três posições da segunda sequência. As diferenças entre RNA e DNA (geralmente a troca de uridina por timidina ou vice-versa) e a presença de análogos de nucleosídeos, como uridinas modificadas, não contribuem para diferenças de identidade ou complementaridade entre polinucleotídeos, desde que os nucleotídeos relevantes (como timidina, uridina ou uridina modificada) tenham o mesmo complemento (por exemplo, adenosina para toda timidina, uridina ou uridina modificada; outro exemplo é citosina e 5-metilcitosina,

ambas com guanosina ou guanosina modificada como complemento). Assim, por exemplo, a sequência 5'-AXG em que X é qualquer uridina modificada, como pseudouridina, N1-metil pseudouridina ou 5-metoxiuridina, é considerada 100% idêntica à AUG, pois ambas são perfeitamente complementares à mesma sequência (5'-CAU). Exemplos de algoritmos de alinhamento são os algoritmos Smith-Waterman e Needleman-Wunsch, que são bem conhecidos na técnica. Um versado na técnica entenderá que escolha de algoritmo e configurações de parâmetro é apropriada para um determinado par de sequências a serem alinhadas; para sequências de comprimento e identidade geralmente semelhantes e esperadas >50% para aminoácidos ou >75% para nucleotídeos, o algoritmo Needleman-Wunsch com configurações padrão da interface do algoritmo Needleman-Wunsch fornecida pelo EBI no www.ebi.ac.uk servidor web é geralmente apropriado.

mRNAs

[0078] Em algumas modalidades, uma composição ou formulação divulgada compreende um mRNA compreendendo um quadro de leitura aberto (ORF), tal como, por exemplo um ORF que codifica um agente de ligação a DNA guiado por RNA, como uma nuclease Cas ou nuclease Cas de Classe 2, conforme descrito neste documento. Em algumas modalidades, um mRNA compreendendo um ORF que codifica um agente de ligação a DNA guiado por RNA, como uma nuclease Cas ou nuclease Cas de Classe 2, é fornecido, usado ou administrado. Em algumas modalidades, o ORF que codifica um agente de ligação ao DNA guiado por RNA é um "agente de ligação ao DNA guiado por RNA modificado ORF" ou simplesmente um "ORF modificado", que é usado como abreviação para indicar que o ORF é modificado em um ou mais das seguintes maneiras: (1) o ORF modificado tem um conteúdo de uridina que varia de seu conteúdo mínimo de uridina a 150% do conteúdo mínimo de uridina; (2) o ORF modificado possui um con-

teúdo de dinucleotídeo de uridina que varia desde seu conteúdo mínimo de dinucleotídeo de uridina até 150% do conteúdo mínimo de dinucleotídeo de uridina; (3) o ORF modificado possui pelo menos 90% de identidade com qualquer uma das SEQ ID NOs: 1, 4, 10, 14, 15, 17, 18, 20, 21, 23, 24, 26, 27, 29, 30, 50, 52, 54, 65 ou 66; (4) o ORF modificado consiste em um conjunto de códons dos quais pelo menos 75% dos códons são códons mínimos de uridina para um dado aminoácido, por exemplo, códons com o menor número de uridinas (geralmente 0 ou 1, exceto para um códon para fenilalanina, onde o códon mínimo de uridina possui 2 uridinas); ou (5) o ORF modificado compreende pelo menos uma uridina modificada. Em algumas modalidades, o ORF modificado é modificado em pelo menos duas, três ou quatro das maneiras anteriores. Em algumas modalidades, o ORF modificado compreende pelo menos uma uridina modificada e é modificado em pelo menos um, dois, três ou todos os itens (1)-(4) acima.

[0079] "Uridina modificada" é usado neste documento para se referir a um nucleosídeo diferente da timidina com os mesmos aceitadores de ligação de hidrogênio que a uridina e uma ou mais diferenças estruturais da uridina. Em algumas modalidades, uma uridina modificada é uma uridina substituída, isto é, uma uridina na qual um ou mais substituintes não prótons (por exemplo, alcóxi, como metóxi) tomam o lugar de um próton. Em algumas modalidades, uma uridina modificada é pseudouridina. Em algumas modalidades, uma uridina modificada é uma pseudouridina substituída, ou seja, uma pseudouridina na qual um ou mais substituintes não prótons (por exemplo, alquil, como metil) tomam o lugar de um próton. Em algumas modalidades, uma uridina modificada é uma uridina substituída, pseudouridina ou pseudouridina substituída.

[0080] "Posição da uridina", conforme utilizado neste documento, refere-se a uma posição em um polinucleotídeo ocupado por uma uridina ou uma uridina modificada. Assim, por exemplo, um polinucleotí-

deo no qual "100% das posições de uridina são uridinas modificadas" contém uma uridina modificada em todas as posições que seriam uridinas em um RNA convencional (onde todas as bases são bases padrão A, U, C ou G) da mesma sequência. A menos que indicado de outra forma, um U em uma sequência polinucleotídica de uma tabela de sequências ou listagem de sequências, ou que acompanha, esta descrição pode ser uma uridina ou uma uridina modificada.

Tabela 1. Códons mínimos de uridina

	Aminoácido	**Códon mínimo de uridina**
A	Alanina	GCA ou GCC ou GCG
G	Glicina	GGA ou GGC ou GGG
V	Valina	GUC ou GUA ou GUG
D	Ácido aspártico	GAC
E	Ácido glutâmico	GAA ou GAG
I	Isoleucina	AUC ou AUA ou AUG
T	Treonina	ACA ou ACC ou ACG
N	Asparagina	AAC
K	Lisina	AAG ou AAA
S	Serina	AGC
R	Arginina	AGA ou AGG
L	Leucina	CUG ou CUA ou CUC
P	Prolina	CCG ou CCA ou CCC
H	Histidina	CAC ou CAA ou CAG
Q	Glutamina	CAG ou CAA
F	Fenilalanina	UUC
Y	Tirosina	UAC
C	Cisteína	UGC
W	Triptofano	UGG
M	Metionina	AUG

[0081] Em qualquer uma das modalidades anteriores, o ORF modificado pode consistir em um conjunto de códons dos quais pelo menos 75%, 80%, 85%, 90%, 95%, 98%, 99% ou 100% dos códons são códons listados na Tabela de códons mínimos de uridina. Em qualquer uma das modalidades anteriores, o ORF modificado pode compreender uma sequência com pelo menos 90%, 95%, 98%, 99% ou 100% de identidade com qualquer uma das SEQ ID NO: 1, 4, 10, 14, 15, 17, 18, 20, 21, 23, 24, 26, 27, 29, 30, 50, 52, 54, 65 ou 66.

[0082] Em qualquer uma das modalidades anteriores, o ORF modificado pode ter um conteúdo de uridina que varia de seu conteúdo mínimo de uridina a 150%, 145%, 140%, 135%, 130%, 125%, 120%, 115%, 110%, 105%, 104%, 103%, 102% ou 101% do conteúdo mínimo de uridina.

[0083] Em qualquer uma das modalidades anteriores, o ORF modificado pode ter um conteúdo de dinucleotídeo de uridina que varia de seu conteúdo mínimo de dinucleotídeo de uridina a 150%, 145%, 140%, 135%, 130%, 125%, 120%, 115%, 110%, 105%, 104%, 103%, 102% ou 101% do teor mínimo de dinucleotídeo de uridina.

[0084] Em qualquer uma das modalidades anteriores, a ORF modificada pode compreender uma uridina modificada pelo menos em uma, uma pluralidade de, ou todas as posições de uridina. Em algumas modalidades, a uridina modificada é uma uridina modificada na posição 5, por exemplo, com um halogênio, metil ou etila. Em algumas modalidades, a uridina modificada é uma pseudouridina modificada na posição 1, por exemplo, com um halogênio, metil ou etila. A uridina modificada pode ser, por exemplo, pseudouridina, N1-metil-pseudouridina, 5-metoxiuridina, 5-iodouridina ou uma combinação destes. Em algumas modalidades, a uridina modificada é 5-metoxiuridina. Em algumas modalidades, a uridina modificada é 5-iodouridina. Em algumas modalidades, a uridina modificada é pseudouridina. Em algumas mo-

dalidades, a uridina modificada é N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de N1-metil pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de 5-iodouridina e N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de 5-iodouridina e 5-metoxiuridina.

[0085] Em algumas modalidades, pelo menos 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% ou 100% das posições de uridina em um mRNA de acordo com a descrição são uridinas modificadas. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são uridinas modificadas, por exemplo, 5-metoxiuridina, 5-iodouridina, N1-metil pseudouridina, pseudouridina ou uma combinação destes. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são 5-metoxiuridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são N1-metil pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%,

65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são 5-iodouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são 5-metoxiuridina, e o restante é N1-metil pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um mRNA de acordo com a descrição são 5-iodouridina e o restante são N1-metil pseudouridina.

[0086] Em qualquer uma das modalidades anteriores, o ORF modificado pode compreender um conteúdo reduzido de dinucleotídeo de uridina, como o menor conteúdo possível de dinucleotídeo de uridina (UU), por exemplo, um ORF que (a) use um códon mínimo de uridina (como discutido acima) em todas as posições e (b) codifica a mesma sequência de aminoácidos que o ORF fornecido. O conteúdo de dinucleotídeo de uridina (UU) pode ser expresso em termos absolutos como a enumeração de dinucleotídeos de UU em um ORF ou em uma base de taxa como a porcentagem de posições ocupadas pelas uridinas de dinucleotídeos de uridina (por exemplo, AUUAU teria um conteúdo de dinucleotídeo de uridina de 40% porque 2 de 5 posições são ocupadas pelas uridinas de um dinucleotídeo de uridina). Os resíduos de uridina modificados são considerados equivalentes às uridinas com o objetivo de avaliar o conteúdo mínimo de dinucleotídeo de uridina.

[0087] Em algumas modalidades, o mRNA compreende pelo menos uma UTR de um mRNA de mamífero expresso, como um mRNA expresso constitutivamente. Um mRNA é considerado constitutivamente expresso em um mamífero se for continuamente transcrito em pelo menos um tecido de um mamífero adulto saudável. Em algumas modalidades, o mRNA compreende um 5' UTR, 3' UTR ou 5' e 3' UTRs de um RNA de mamífero expresso, tal como um mRNA de mamífero ex-

presso constitutivamente. O mRNA da actina é um exemplo de um mRNA expresso constitutivamente.

[0088] Em algumas modalidades, o mRNA compreende pelo menos uma UTR da hidroxisteroide 17-beta desidrogenase 4 (HSD17B4 ou HSD), por exemplo, uma 5' UTR da HSD. Em algumas modalidades, o mRNA compreende pelo menos uma UTR de um mRNA de globina, por exemplo, mRNA de alfa globina humana (HBA), mRNA de beta globina humana (HBB) ou mRNA de Xenopus laevis beta globina (XBG). Em algumas modalidades, o mRNA compreende um 5' UTR, 3' UTR ou 5' e 3' UTRs de um mRNA de globina, como HBA, HBB ou XBG. Em algumas modalidades, o mRNA compreende uma 5' UTR do hormônio de crescimento bovino, citomegalovírus (CMV), Hba-a1 de camundongo, HSD, um gene da albumina, HBA, HBB ou XBG. Em algumas modalidades, o mRNA compreende uma 3' UTR do hormônio de crescimento bovino, citomegalovírus, Hba-a1 de camundongo, HSD, um gene de albumina, HBA, HBB ou XBG. Em algumas modalidades, o mRNA compreende 5' e 3' UTRs do hormônio de crescimento bovino, citomegalovírus, Hba-a1 de camundongo, HSD, um gene da albumina, HBA, HBB, XBG, proteína de choque térmico 90 (Hsp90), gliceraldeído 3-fosfato desidrogenase (GAPDH), beta-actina, alfa-tubulina, proteína tumoral (p53) ou receptor do fator de crescimento epidérmico (EGFR).

[0089] Em algumas modalidades, o mRNA compreende 5' e 3' UTRs que são da mesma fonte, por exemplo, um mRNA expressado constitutivamente como actina, albumina ou globina como HBA, HBB ou XBG.

[0090] Em algumas modalidades, o mRNA não compreende uma 5' UTR, por exemplo, não há nucleotídeos adicionais entre o cap 5' e o códon de iniciação. Em algumas modalidades, o mRNA compreende uma sequência de Kozak (descrita abaixo) entre o cap 5' e o códon de

iniciação, mas não possui nenhuma 5' UTR adicional. Em algumas modalidades, o mRNA não compreende uma 3' UTR, por exemplo, não há nucleotídeos adicionais entre o códon de parada e a tail poli-A.

[0091] Em algumas modalidades, o mRNA compreende uma sequência de Kozak. A sequência de Kozak pode afetar o início da tradução e o rendimento geral de um polipeptídeo traduzido de um mRNA. Uma sequência de Kozak inclui um códon de metionina que pode funcionar como o códon de iniciação. Uma sequência Kozak mínima é NNNRUGN em que pelo menos um dos seguintes é verdadeiro: o primeiro N é A ou G e o segundo N é G. No contexto de uma sequência nucleotídica, R significa uma purina (A ou G). Em algumas modalidades, a sequência de Kozak é RNNRUGN, NNNRUGG, RNNRUGG, RNNAUGN, NNNAUGG ou RNNAUGG. Em algumas modalidades, a sequência de Kozak é rccRUGg com zero incompatibilidades ou com até uma ou duas incompatibilidades para posições minúsculas. Em algumas modalidades, a sequência de Kozak é rccAUGg com zero incompatibilidades ou com até uma ou duas incompatibilidades para posições minúsculas. Em algumas modalidades, a sequência Kozak é gccRccAUGG com zero incompatibilidades ou com até uma, duas ou três incompatibilidades para posições minúsculas. Em algumas modalidades, a sequência de Kozak é gccAccAUG com zero incompatibilidades ou com até uma, duas, três ou quatro incompatibilidades para posições minúsculas. Em algumas modalidades, a sequência de Kozak é GCCACCAUG. Em algumas modalidades, a sequência Kozak é gccgccRccAUGG com zero incompatibilidades ou com até uma, duas, três ou quatro incompatibilidades para posições minúsculas.

[0092] Em algumas modalidades, o mRNA que compreende um ORF que codifica um agente de ligação a DNA guiado por RNA compreende uma sequência com pelo menos 90% de identidade com a SEQ ID NO: 43, opcionalmente em que o ORF da SEQ ID NO: 43 (ou

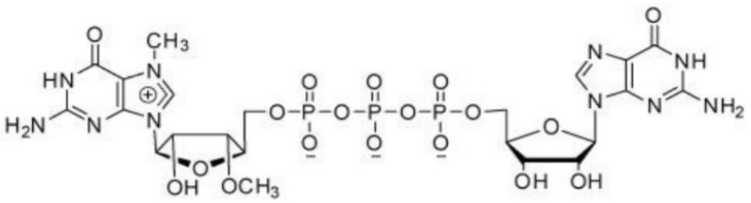
seja, SEQ ID NO: 4) é substituído por um ORF alternativo. Em algumas modalidades, o mRNA compreende qualquer uma das SEQ ID NOs: 10, 14, 15, 17, 18, 20, 21, 23, 24, 26, 27, 29, 30, 50, 52, 54, 65 ou 66.

[0093] Em algumas modalidades, o grau de identidade com a sequência opcionalmente substituída da SEQ ID NO: 43 é de 95%. Em algumas modalidades, o grau de identidade com a sequência opcionalmente substituída da SEQ ID NO: 4 é de 98%. Em algumas modalidades, o grau de identidade com a sequência opcionalmente substituída da SEQ ID NO: 43 é de 99%. Em algumas modalidades, o grau de identidade com a sequência opcionalmente substituída da SEQ ID NO: 43 é de 100%.

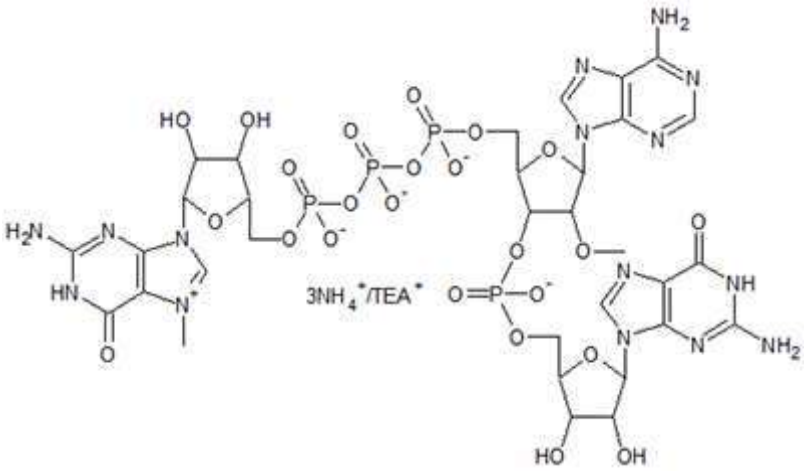
[0094] Em algumas modalidades, um mRNA divulgado neste documento compreende um cap 5', como uma Cap0, Cap1 ou Cap2. Um cap 5' é geralmente um ribonucleotídeo de 7-metilguanina (que pode ser modificado ainda mais, conforme discutido abaixo, por exemplo, com relação ao ARCA) ligado através de um 5'-trifosfato à posição 5' do primeiro nucleotídeo da cadeia 5'-a- 3' do mRNA, isto é, o primeiro nucleotídeo cap-proximal. No Cap0, as riboses do primeiro e do segundo nucleotídeos cap-proximais do mRNA compreendem um 2'-hidroxila. No Cap1, as riboses do primeiro e do segundo nucleotídeos transcritos do mRNA compreendem 2'-metóxi e 2'-hidróxila, respectivamente. No Cap2, as riboses do primeiro e do segundo nucleotídeos cap-proximais do mRNA compreendem um 2'-metóxi. Veja, por exemplo, Katibah et al. (2014) *Proc Natl Acad Sci USA* 111(33):12025-30; Abbas et al. (2017) *Proc Natl Acad Sci USA* 114(11):E2106-E2115. A maioria dos mRNAs eucarióticos endógenos superiores, incluindo os mRNAs de mamíferos, como os mRNAs humanos, compreendem Cap1 ou Cap2. Cap0 e outras estruturas de caps diferentes de Cap1 e Cap2 podem ser imunogênicas em mamíferos, como seres humanos,

devido ao reconhecimento como "não-próprio" por componentes do sistema imunológico inato, como IFIT-1 e IFIT-5, que podem resultar em níveis elevados de citocinas, incluindo interferon tipo I. Os componentes do sistema imunológico inato, como o IFIT-1 e o IFIT-5, também podem competir com o eIF4E pela ligação de um mRNA com um cap diferente de Cap1 ou Cap2, potencialmente inibindo a tradução do mRNA.

[0095] Um cap pode ser incluído co-transcricionalmente. Por exemplo, ARCA (análogo de tampa anti-reversa; Thermo Fisher Scientific Cat. N°. AM8045) é um análogo de cap que compreende um 3'-metóxi-5'-trifosfato de 7-metilguanina ligado à posição 5' de um ribonucleotídeo de guanina que pode ser incorporado *in vitro* em um transcrito no início. O ARCA resulta em um cap Cap0 na qual a posição 2' do primeiro nucleotídeo cap-proximal é hidroxila. Ver, por exemplo, Stepinski et al., (2001) "Synthesis and properties of mRNAs containing the novel 'anti-reverse' cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'deoxy)GpppG," *RNA* 7: 1486–1495. A estrutura ARCA é mostrada abaixo.



[0096] CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies Cat. N°. N-7113) ou CleanCap™ GG (m7G (5') ppp (5') (2'OMeG) pG; TriLink Biotechnologies ref. N°. N-7133) pode ser usado para fornecer uma estrutura Cap1 co-transcricionalmente. As versões 3'-O-metiladas do CleanCap™AG e CleanCap™GG também estão disponíveis na TriLink Biotechnologies como Cat. Nºs N-7413 e N-7433, respectivamente. A estrutura CleanCap™ AG é mostrada abaixo.



[0097] Alternativamente, um cap pode ser adicionado a um RNA pós-transcricionalmente. Por exemplo, a enzima de cobertura Vaccinia está disponível comercialmente (New England Biolabs Cat. N°. M2080S) e possui atividades de RNA trifosfatase e guanililtransferase, fornecidas por sua subunidade D1, e guanina metiltransferase, fornecidas por sua subunidade D12. Como tal, pode adicionar uma 7-metilguanina a um RNA, de modo a originar Cap0, na presença de S-adenosilmetionina e GTP. Ver, por exemplo, Guo, P. e Moss, B. (1990) *Proc. Natl. Acad. Sci. USA* 87, 4023-4027; Mao, X. e Shuman, S. (1994) *J. Biol. Chem*. 269, 24472-24479.

[0098] Em algumas modalidades, o mRNA compreende ainda um tail poli-adenilada (poli-A). Em algumas modalidades, o tail poli-A compreende pelo menos 20, 30, 40, 50, 60, 70, 80, 90 ou 100 adeninas, opcionalmente até 300 adeninas. Em algumas modalidades, o tail poli-A compreende 95, 96, 97, 98, 99 ou 100 nucleotídeos de adenina. Em alguns casos, o tail poli-A é "interrompido" com uma ou mais "âncoras" de nucleotídeos não-adenina em um ou mais locais dentro do tail poli-A. Os tails de poli-A podem compreender pelo menos 8 nucleotídeos de adenina consecutivos, mas também compreender um ou mais nucleotídeos não-adenina. Como usado neste documento, "nucleotídeos não adeninos" referem-se a nucleotídeos naturais ou não naturais que não compreendem adenina. Os nucleotídeos de guanina, timina e cito-

sina são nucleotídeos não-adenina exemplares. Assim, os tails de poli-A no mRNA descritos neste documento podem compreender nucleotídeos de adenina consecutivos localizados 3' a nucleotídeos que codificam um agente de ligação a DNA guiado por RNA ou uma sequência de interesse. Em alguns casos, os tails de poli-A no mRNA compreendem nucleotídeos de adenina não consecutivos localizados 3' em nucleotídeos que codificam um agente de ligação a DNA guiado por RNA ou uma sequência de interesse, em que nucleotídeos não de adenina interrompem os nucleotídeos de adenina regularmente ou irregularmente intervalos espaçados.

[0099] Como usado neste documento, "nucleotídeos não adeninos" referem-se a nucleotídeos naturais ou não naturais que não compreendem adenina. Os nucleotídeos de guanina, timina e citosina são nucleotídeos não-adenina exemplares. Assim, os tails de poli-A no mRNA descritos neste documento podem compreender nucleotídeos de adenina consecutivos localizados 3' a nucleotídeos que codificam um agente de ligação a DNA guiado por RNA ou uma sequência de interesse. Em alguns casos, os tails de poli-A no mRNA compreendem nucleotídeos de adenina não consecutivos localizados 3' em nucleotídeos que codificam um agente de ligação a DNA guiado por RNA ou uma sequência de interesse, em que nucleotídeos não de adenina interrompem os nucleotídeos de adenina regularmente ou irregularmente intervalos espaçados.

[00100] Em algumas modalidades, o mRNA é purificado. Em algumas modalidades, o mRNA é purificado usando um método de precipitação (Por exemplo, precipitação com LiCl, precipitação com álcool ou um método equivalente, Por exemplo, como descrito neste documento). Em algumas modalidades, o mRNA é purificado usando um método baseado em cromatografia, como um método baseado em HPLC ou um método equivalente (Por exemplo, como descrito neste docu-

mento). Em algumas modalidades, o mRNA é purificado usando um método de precipitação (Por exemplo, precipitação com LiCl) e um método baseado em HPLC.

[00101] Em algumas modalidades, pelo menos um gRNA é fornecido em combinação com um mRNA divulgado neste documento. Em algumas modalidades, um gRNA é fornecido como uma molécula separada do mRNA. Em algumas modalidades, um gRNA é fornecido como uma parte, como uma parte de uma UTR, de um mRNA divulgado neste documento.

gRNAs

[00102] Em um aspecto, a presente invenção fornece métodos de entrega de um sistema de edição de genoma (por exemplo, um sistema de nuclease de dedo de zinco, um sistema TALEN, um sistema de meganuclease ou um sistema CRISPR/Cas) a uma célula (ou população de células), para exemplo, um HSPC (ou população de HSPCs), por exemplo, uma célula CD34+ (ou população de células CD34+), em que o resultado é uma célula (ou sua descendência) que aumentou a expressão da hemoglobina fetal (por exemplo, quando a referida célula é diferenciada em um eritrócitos). Divulgadas neste documento são sequências guia úteis para alcançar esse efeito. Nas modalidades, o sistema de edição de genoma compreende um ou mais vetores, por exemplo, mRNA, que codifica os componentes do sistema de edição de genoma. Em outras modalidades, o sistema de edição de genoma compreende um ou mais polipeptídeos. Num aspecto preferido, os métodos compreendem a entrega de um sistema CRISPR/Cas. Em modalidades, o sistema CRISPR/Cas compreende um gRNA e uma nuclease Cas, por exemplo, complexada na forma de um complexo de proteínas ribonucleares (RNP). Em outras modalidades, o sistema CRISPR/Cas compreende um ou mais vetores que codificam um gRNA e/ou uma nuclease Cas. Em outras modalidades, o sistema

CRISPR/Cas compreende um ou mais vetores, por exemplo, mRNA, que codifica uma nuclease de Cas (por exemplo, uma nuclease Cas de Classe 2) e um ou mais gRNAs. Em aspectos, o sistema CRISPR/Cas inclui um gRNA descrito em WO2017/115268, cujo conteúdo é incorporado neste documento por referência na sua totalidade. Em aspectos, o sistema CRISPR/Cas inclui um gRNA compreendendo uma sequência guia complementar a uma sequência alvo dentro do gene BCL11a ou de seus elementos reguladores. Em outros aspectos, o sistema CRISPR/Cas inclui um gRNA compreendendo uma sequência guia complementar a uma sequência alvo no íntron 2 do gene BCL11a (por exemplo, dentro de uma região do íntron 2 do gene BCL11a no ou próximo a um sítio de ligação GATA1. Em aspectos, o sistema CRISPR/Cas inclui um gRNA que compreende uma sequência guia complementar a uma sequência alvo na região do íntron 2 do gene BCL11a de ch2: 60494000 a ch2: 60498000 (de acordo com hg38), por exemplo, dentro de uma região do íntron 2 do gene BCL11a de ch2: 60494250 a ch2: 60496300 (de acordo com hg38). Nas modalidades, o sistema CRISPR/Cas inclui um gRNA compreendendo uma sequência guia listada na Tabela 2 do Pedido Provisório US N°. 62/566.232, depositado em 29 de setembro de 2017, que é incorporado por referência neste documento.

[00103] Sequências guia exemplares de gRNAs que são complementares às sequências alvo dentro do íntron 2 do gene BCL11a. +58, +62 e +55 referem-se aos sítios de hipersensibilidade à DNAse da região intensificadora específica para eritroide, conforme descrito em Bauer et al., *Science* 2013; 342 (6155): 253-257.

[00104] Em outros aspectos, o sistema CRISPR/Cas inclui um gRNA que compreende uma sequência guia complementar a uma sequência alvo dentro do locus da globina no cromossomo 11. Em um aspecto, o sistema CRISPR/Cas inclui um gRNA que compreende

uma sequência guia complementar a uma sequência dentro de uma região HPFH. Conforme usado neste documento, o termo "região HPFH" refere-se a um sítio genômico que, quando modificado (por exemplo, mutado ou deletado), causa aumento da produção de HbF nos glóbulos vermelhos adultos e inclui as regiões HPFH identificadas na literatura (consulte, por exemplo, Herança Mendeliana Online no Homem: http://www.omim.org/entry/141749, incorporada por referência neste documento). Em uma modalidade exemplar, a região HPFH é uma região dentro ou que engloba o aglomerado de genes da beta globina no cromossomo 11p15. Em uma modalidade exemplar, a região HPFH está dentro ou abrange pelo menos parte do gene delta globina e seus elementos reguladores. Numa modalidade exemplar, a região HPFH é uma região do promotor de HBG1. Em uma modalidade exemplar, a região HPFH é uma região do promotor de HBG2. Numa modalidade exemplar, a região HPFH é uma região descrita em Sankaran VG et al. NEJM (2011) 365: 807-814, incorporado por referência neste documento na sua totalidade. Em uma modalidade exemplar, a região HPFH é o HPFH delecional de ponto de interrupção francês, como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o HPFH argelino, como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Numa modalidade exemplar, a região HPFH é o HPFH do Sri Lanka, como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o HPFH-3 como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o HPFH-2 como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade, a região HPFH-1 é o HPFH-3 como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Numa modalidade exemplar, a região HPFH é o HPFH do Sri Lanka (δβ)0-thalassemia, como descrito em Sankaran

VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o HPFH siciliano (δβ)0-talassemia, como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o HPFH Macedônia (δβ)0-talassemia, como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é o Kurdish β0-thalassemia HPFH como descrito em Sankaran VG et al. NEJM (2011) 365: 807-814. Em uma modalidade exemplar, a região HPFH é a região localizada em Chr11: 5213874-5214400 (hg18). Em uma modalidade exemplar, a região HPFH é a região localizada em Chr11: 5215943-5215046 (hg18). Em uma modalidade exemplar, a região HPFH é a região localizada em Chr11: 5234390-5238486 (hg38). Em modalidades, o sistema CRISPR/Cas inclui um gRNA que compreende uma sequência guia que compreende uma sequência como descrita em WO2017/077394, cujo conteúdo é incorporado neste documento por referência na sua totalidade. Nas modalidades, o sistema CRISPR/Cas inclui um gRNA que compreende uma sequência guia que compreende uma sequência selecionada a partir de uma sequência guia do documento WO2017/077394. Nas modalidades, o sistema CRISPR/Cas inclui um gRNA compreendendo uma sequência guia listada na Tabela 3 do Pedido Provisório US N°. 62/566.232, depositado em 29 de setembro de 2017, que é incorporado por referência neste documento.

[00105] Sequências guia exemplares direcionadas ao HPFH francês (HPFH francês; Sankaran VG et al. Um elemento funcional necessário para o silenciamento da hemoglobina fetal. NEJM (2011) 365: 807-814.)

[00106] Nas modalidades, o sistema CRISPR/Cas inclui um gRNA compreendendo uma sequência guia listada na Tabela 4 do Pedido Provisório US N°. 62/566.232, depositado em 29 de setembro de 2017, que é incorporado por referência neste documento.

[00107] Sequências guia exemplares podem ser direcionadas para as regiões promotoras de HBG1 e/ou HBG2.

gRNA quimicamente modificado

[00108] Em algumas modalidades, o gRNA é quimicamente modificado. Um gRNA compreendendo um ou mais nucleosídeos ou nucleotídeos modificados é chamado de gRNA "modificado" ou gRNA "quimicamente modificado" para descrever a presença de um ou mais componentes ou configurações que não são naturais e/ou naturais que são usados em vez de ou em além dos resíduos canônicos A, G, C e U. Em algumas modalidades, um gRNA modificado é sintetizado com um nucleosídeo ou nucleotídeo não canônico, é chamado "modificado" neste documento. Os nucleosídeos e nucleotídeos modificados podem incluir um ou mais dos seguintes: (i) alteração, Por exemplo, substituição, de um ou de ambos os oxigênios de fosfato não-ligados e/ou de um ou mais dos oxigênios de fosfato de ligação na ligação da estrutura principal fosfodiéster (uma modificação da estrutura principal exemplar); (ii) alteração, Por exemplo, substituição, de um constituinte do açúcar ribose, Por exemplo, do hidroxil 2' no açúcar ribose (uma modificação exemplar no açúcar); (iii) substituição por atacado da fração fosfato por ligantes "defosfo" (uma modificação exemplar da estrutura principal); (iv) modificação ou substituição de uma nucleobase de ocorrência natural, incluindo uma nucleobase não canônica (uma modificação de base exemplar); (v) substituição ou modificação da estrutura principal ribose-fosfato (uma modificação exemplar da estrutura principal); (vi) modificação da extremidade 3' ou da extremidade 5' do oligonucleotídeo, Por exemplo, remoção, modificação ou substituição de um grupo fosfato terminal ou conjugação de uma fração, cap ou ligante (tais modificações de cap 3' ou 5' podem compreender uma modificação de açúcar e/ou estrutura principal); e (vii) modificação ou substituição do açúcar (uma modificação exemplar do açúcar).

[00109] Em algumas modalidades, um gRNA compreende uma uridina modificada em algumas ou todas as posições de uridina. Em algumas modalidades, a uridina modificada é uma uridina modificada na posição 5, por exemplo, com um halogênio ou C1-C6 alcóxi. Em algumas modalidades, a uridina modificada é uma pseudouridina modificada na posição 1, por exemplo, com um C1-C6 alquila. A uridina modificada pode ser, por exemplo, pseudouridina, N1-metil-pseudouridina, 5-metoxiuridina, 5-iodouridina ou uma combinação destes. Em algumas modalidades, a uridina modificada é 5-metoxiuridina. Em algumas modalidades, a uridina modificada é 5-iodouridina. Em algumas modalidades, a uridina modificada é pseudouridina. Em algumas modalidades, a uridina modificada é N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de N1-metil pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de 5-iodouridina e N1-metil-pseudouridina. Em algumas modalidades, a uridina modificada é uma combinação de pseudouridina e 5-metoxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de 5-iodouridina e 5-metoxiuridina.

[00110] Em algumas modalidades, pelo menos 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% ou 100% das posições de uridina em um gRNA de acordo com a descrição são uridinas modificadas. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são uridinas modificadas, por exemplo, 5-metoxiuridina, 5-iodouridina, N1-metil pseudouridina, pseu-

douridina ou uma combinação destes. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são 5-metoxiuridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são N1-metil pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são 5-iodouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são 5-metoxiuridina, e o restante é N1-metil pseudouridina. Em algumas modalidades, 10% -25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, ou 90-100% das posições de uridina em um gRNA de acordo com a descrição são 5-iodouridina e o restante são N1-metil pseudouridina.

[00111] Modificações químicas, como as listadas acima, podem ser combinadas para fornecer gRNAs modificados compreendendo nucleosídeos e nucleotídeos (coletivamente "resíduos") que podem ter duas, três, quatro ou mais modificações. Por exemplo, um resíduo modificado pode ter um açúcar modificado e uma nucleobase modificada. Em algumas modalidades, todas as bases de um gRNA são modificadas, Por exemplo, todas as bases têm um grupo fosfato modificado, como um grupo fosforotioato. Em certas modalidades, todos ou substancialmente todos os grupos fosfato de uma molécula de gRNA são substituídos por grupos fosforotioato. Em algumas modalidades,

os gRNAs modificados compreendem pelo menos um resíduo modificado ou próximo da extremidade 5' do RNA. Em algumas modalidades, os gRNAs modificados compreendem pelo menos um resíduo modificado ou próximo da extremidade 3' do RNA.

[00112] Em algumas modalidades, o gRNA compreende um, dois, três ou mais resíduos modificados. Em algumas modalidades, pelo menos 5% (Por exemplo, pelo menos 5%, pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 30%, pelo menos 35%, pelo menos 40%, pelo menos 45%, pelo menos 50%, pelo menos 55%, pelo menos 60%, pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95% ou 100%) das posições em um gRNA modificado são nucleosídeos ou nucleotídeos modificados.

[00113] Os ácidos nucleicos não modificados podem ser propensos à degradação, por exemplo, nucleases intracelulares ou aquelas encontradas no soro. Por exemplo, as nucleases podem hidrolisar as ligações fosfodiéster do ácido nucleico. Por conseguinte, em um aspecto, os gRNAs descritos neste documento podem conter um ou mais nucleosídeos ou nucleotídeos modificados, Por exemplo, para introduzir estabilidade em relação a nucleases intracelulares ou baseadas em soro. Em algumas modalidades, as moléculas de gRNA modificado descritas neste documento podem exibir uma resposta imune inata reduzida quando introduzidas em uma população de células, tanto *in vivo* quanto *ex vivo*. O termo "resposta imune inata" inclui uma resposta celular a ácidos nucleicos exógenos, incluindo ácidos nucleicos de fita simples, que envolvem a indução da expressão e liberação de citocinas, particularmente os interferons, e a morte celular.

[00114] Em algumas modalidades de uma modificação da estrutura principal, o grupo fosfato de um resíduo modificado pode ser modificado substituindo um ou mais dos oxigênios por um substituinte diferen-

te. Além disso, o resíduo modificado, Por exemplo, resíduo modificado presente em um ácido nucleico modificado, pode incluir a substituição por atacado de uma fração fosfato não modificada por um grupo fosfato modificado, como descrito neste documento. Em algumas modalidades, a modificação da estrutura principal da estrutura principal de fosfato pode incluir alterações que resultam em um ligante não carregado ou um ligante carregado com distribuição de carga assimétrica.

[00115] Exemplos de grupos fosfato modificados incluem fosforotioato, fosforosselenatos, fosfatos de borano, ésteres de borano fosfato, fosfonatos de hidrogênio, fosforamidatos, fosfonatos de alquil ou aril e fosfotriésteres. O átomo de fósforo em um grupo fosfato não modificado é aquiral. No entanto, a substituição de um dos oxigênios sem ponte por um dos átomos ou grupos de átomos acima pode tornar o átomo de fósforo quiral. O átomo de fósforo estereogênico pode possuir a configuração "R" (Rp neste documento) ou a configuração "S" (Sp neste documento). A estrutura principal também pode ser modificada pela substituição de um oxigênio em ponte (*ou seja*, o oxigênio que liga o fosfato ao nucleosídeo), por nitrogênio (fosforamamidatos em ponte), enxofre (fosforotioatos em ponte) e carbono (metilenofosfonatos em ponte). A substituição pode ocorrer no oxigênio de ligação ou em ambos os oxigênios de ligação.

[00116] O grupo fosfato pode ser substituído por conectores que não contenham fósforo em certas modificações na estrutura principal. Em algumas modalidades, o grupo fosfato carregado pode ser substituído por uma fração neutra. Exemplos de frações que podem substituir o grupo fosfato podem incluir, sem limitação, Por exemplo, fosfonato de metil, hidroxilamino, siloxano, carbonato, carboximetil, carbamato, amida, tioéter, ligante de óxido de etileno, sulfonato, sulfonamida, tioformacetal, formalcetal, oxima, metilenoimino, metilenometilimino, metileno-hidrazo, metilenodimetil-hidrazo e metileno-oximetilimino.

[00117] Em algumas modalidades, a invenção compreende um sgRNA compreendendo uma ou mais modificações dentro de uma ou mais das seguintes regiões: os nucleotídeos no terminal 5'; a região inferior do caule; a região de abaulamento; a região do tronco superior; a região do nexo; a região hairpin 1; a região hairpin 2; e os nucleotídeos no terminal 3'. Em algumas modalidades, a modificação compreende um nucleotídeo modificado em 2'-O-metil (2'-O-Me). Em algumas modalidades, a modificação compreende um nucleotídeo modificado em 2'-fluoro (2'-F). Em algumas modalidades, a modificação compreende uma ligação de fosforotioato (PS) entre nucleotídeos.

[00118] Em algumas modalidades, os primeiros três ou quatro nucleotídeos no terminal 5' e os últimos três ou quatro nucleotídeos no terminal 3' são modificados. Em algumas modalidades, os quatro primeiros nucleotídeos no terminal 5' e os últimos quatro nucleotídeos no terminal 3' estão ligados a ligações fosforotioato (PS). Em algumas modalidades, a modificação compreende 2'-O-Me. Em algumas modalidades, a modificação compreende 2'-F.

[00119] Em algumas modalidades, os quatro primeiros nucleotídeos no terminal 5' e os quatro últimos nucleotídeos no terminal 3' estão ligados a uma ligação PS, e os três primeiros nucleotídeos no terminal 5' e os últimos três nucleotídeos no terminal 3' compreendem modificações 2'-O-Me.

[00120] Em algumas modalidades, os quatro primeiros nucleotídeos no terminal 5' e os quatro últimos nucleotídeos no terminal 3' estão ligados a uma ligação PS, e os três primeiros nucleotídeos no terminal 5' e os últimos três nucleotídeos no terminal 3' compreendem modificações 2'-F.

[00121] Em algumas modalidades, o sgRNA compreende o padrão de modificação de SEQ ID NO: 74:

(mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAm

GmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAU-CAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCm-GmAmGmUmCmGmGmUmGmCmU*mU*mU*mU), onde N é qualquer nucleotídeo natural ou não natural. Em algumas modalidades, o sgRNA compreende a SEQ ID NO: 74. Em certas modalidades, o sgRNA compreende a modificação 2'O-metil dos três primeiros resíduos na sua extremidade 5', com ligações de fosforotioato entre os resíduos 1-2, 2-3 e 3-4 do RNA.

Ácido Nucleico Modelo

[00122] As composições e métodos divulgados neste documento podem incluir um ácido nucleico modelo. O modelo pode ser usado para alterar ou inserir uma sequência de ácido nucleico no local ou próximo a um sítio alvo para uma nuclease Cas. Em algumas modalidades, os métodos compreendem a introdução de um modelo para a célula. Em algumas modalidades, um único modelo pode ser fornecido. Em outras modalidades, dois ou mais modelos podem ser fornecidos de modo que a edição possa ocorrer em dois ou mais sítios alvo. Por exemplo, modelos diferentes podem ser fornecidos para editar um único gene em uma célula ou dois genes diferentes em uma célula.

[00123] Em algumas modalidades, o modelo pode ser usado na recombinação homóloga. Em algumas modalidades, a recombinação homóloga pode resultar na integração da sequência modelo ou de uma porção da sequência modelo na molécula de ácido nucleico alvo. Em outras modalidades, o modelo pode ser usado no reparo direcionado à homologia, que envolve a invasão da fita de DNA no sítio da clivagem no ácido nucleico. Em algumas modalidades, o reparo direcionado à homologia pode resultar na inclusão da sequência modelo na molécula de ácido nucleico alvo editada. Em ainda outras modalidades, o modelo pode ser usado na edição de genes mediada por junção de extremidade não homóloga. Em algumas modalidades, a sequência

modelo não tem semelhança com a sequência de ácido nucleico próxima ao sítio de clivagem. Em algumas modalidades, o modelo ou uma porção da sequência modelo é incorporado. Em algumas modalidades, o modelo inclui sequências de repetição invertida terminal (ITR) de flanqueamento.

[00124] Em algumas modalidades, o modelo pode compreender um primeiro braço de homologia e um segundo braço de homologia (também chamado de primeira e segunda sequência de nucleotídeos) que são complementares às sequências localizadas a montante e a jusante do local de clivagem, respectivamente. Onde um modelo contém dois braços de homologia, cada braço pode ter o mesmo comprimento ou comprimentos diferentes, e a sequência entre os braços de homologia pode ser substancialmente semelhante ou idêntica à sequência alvo entre os braços de homologia ou pode ser totalmente não relacionada. Em algumas modalidades, o grau de complementaridade ou porcentagem de identidade entre a primeira sequência de nucleotídeos no modelo e a sequência a montante do sítio de clivagem e entre a segunda sequência de nucleotídeos no modelo e a sequência a jusante do sítio de clivagem, pode permitir recombinação homóloga, como, Por exemplo, recombinação homóloga de alta fidelidade, entre o modelo e a molécula de ácido nucleico alvo. Em algumas modalidades, o grau de complementaridade pode ser de cerca de 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, ou 100%. Em algumas modalidades, o grau de complementaridade pode ser de cerca de 95%, 97%, 98%, 99% ou 100%. Em algumas modalidades, o grau de complementaridade pode ser de pelo menos 98%, 99% ou 100%. Em algumas modalidades, o grau de complementaridade pode ser 100%. Em algumas modalidades, a percentagem de identidade pode ser de cerca de 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a

identidade percentual pode ser de cerca de 95%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a percentagem de identidade pode ser de pelo menos 98%, 99% ou 100%. Em algumas modalidades, a percentagem de identidade pode ser 100%.

[00125] Em algumas modalidades, a sequência modelo pode corresponder, compreender ou consistir em uma sequência endógena de uma célula alvo. Pode também ou alternativamente corresponder, compreender ou consistir em uma sequência exógena de uma célula alvo. Como usado neste documento, o termo "sequência endógena" refere-se a uma sequência que é nativa da célula. O termo "sequência exógena" refere-se a uma sequência que não é nativa de uma célula, ou uma sequência cuja localização nativa no genoma da célula está em um local diferente. Em algumas modalidades, a sequência endógena pode ser uma sequência genômica da célula. Em algumas modalidades, a sequência endógena pode ser uma sequência cromossômica ou extracromossômica. Em algumas modalidades, a sequência endógena pode ser uma sequência plasmídica da célula. Em algumas modalidades, a sequência modelo pode ser substancialmente idêntica a uma porção da sequência endógena em uma célula no local ou próximo ao sítio de clivagem, mas compreende pelo menos uma alteração de nucleotídeo. Em algumas modalidades, a edição da molécula de ácido nucleico alvo clivada com o modelo pode resultar em uma mutação compreendendo uma inserção, deleção ou substituição de um ou mais nucleotídeos da molécula de ácido nucleico alvo. Em algumas modalidades, a mutação pode resultar em uma ou mais alterações de aminoácidos em uma proteína expressa a partir de um gene compreendendo a sequência alvo. Em algumas modalidades, a mutação pode resultar em uma ou mais alterações nucleotídicas em um RNA expresso a partir do gene alvo. Em algumas modalidades, a mutação pode alterar o nível de expressão do gene alvo. Em algumas modalidades, a

mutação pode resultar em aumento ou diminuição da expressão do gene alvo. Em algumas modalidades, a mutação pode resultar em knock-down do gene. Em algumas modalidades, a mutação pode resultar em nocaute genético. Em algumas modalidades, a mutação pode resultar na função do gene restaurado. Em algumas modalidades, a edição da molécula de ácido nucleico alvo clivada com o modelo pode resultar em uma alteração em uma sequência de éxons, uma sequência de íntrons, uma sequência reguladora, uma sequência de controle transcricional, uma sequência de controle de tradução, um sítio de splicing ou uma sequência não codificante da molécula de ácido nucleico alvo, como o DNA.

[00126] Em outras modalidades, a sequência modelo pode compreender uma sequência exógena. Em algumas modalidades, a sequência exógena pode compreender uma sequência de codificação de proteína ou RNA operacionalmente ligada a uma sequência promotora exógena, de modo que, após a integração da sequência exógena na molécula de ácido nucleico alvo, a célula é capaz de expressar a proteína ou RNA codificado por a sequência integrada. Em outras modalidades, após a integração da sequência exógena na molécula de ácido nucleico alvo, a expressão da sequência integrada pode ser regulada por uma sequência promotora endógena. Em algumas modalidades, a sequência exógena pode fornecer uma sequência de cDNA que codifica uma proteína ou uma porção da proteína. Em ainda outras modalidades, a sequência exógena pode compreender ou consistir em uma sequência de éxon, uma sequência de íntron, uma sequência reguladora, uma sequência de controle transcricional, uma sequência de controle de tradução, um sítio de splicing ou uma sequência não codificante. Em algumas modalidades, a integração da sequência exógena pode resultar na função do gene restaurado. Em algumas modalidades, a integração da sequência exógena pode resultar em um knock-in

genético. Em algumas modalidades, a integração da sequência exógena pode resultar em um nocaute genético.

[00127] O modelo pode ter qualquer comprimento adequado. Em algumas modalidades, o modelo pode compreender 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, ou mais nucleotídeos de comprimento. O modelo pode ser um ácido nucleico de fita simples. O modelo pode ser ácido nucleico de fita dupla ou parcialmente dupla. Em certas modalidades, o modelo de fita única tem 20, 30, 40, 50, 75, 100, 125, 150, 175 ou 200 nucleotídeos de comprimento. Em algumas modalidades, o modelo pode compreender uma sequência de nucleotídeos que é complementar a uma porção da molécula de ácido nucleico alvo compreendendo a sequência alvo (*isto é*, um "braço de homologia"). Em algumas modalidades, o modelo pode compreender um braço de homologia que é complementar à sequência localizada a montante ou a jusante do sítio de clivagem na molécula de ácido nucleico alvo.

[00128] Em algumas modalidades, o modelo contém ssDNA ou dsDNA contendo sequências flanqueantes de repetição invertida-terminal (ITR). Em algumas modalidades, o modelo é fornecido como um vetor, plasmídeo, minicírculo, nanocírculo ou produto de PCR.

Purificação de Ácidos Nucleicos

[00129] Em algumas modalidades, o ácido nucleico é purificado. Em algumas modalidades, o ácido nucleico é purificado usando um método de precipitação (Por exemplo, precipitação com LiCl, precipitação com álcool ou um método equivalente, Por exemplo, como descrito neste documento). Em algumas modalidades, o ácido nucleico é purificado usando um método baseado em cromatografia, como um método baseado em HPLC ou um método equivalente (Por exemplo, como descrito neste documento). Em algumas modalidades, o nucleico é purificado usando um método de precipitação (Por exemplo, pre-

cipitação com LiCl) e um método baseado em HPLC.

Sequências alvo

[00130] Em algumas modalidades, um sistema CRISPR/Cas da presente invenção pode ser direcionado e clivar uma sequência alvo em uma molécula de ácido nucleico alvo. Por exemplo, a sequência alvo pode ser reconhecida e clivada pela nuclease Cas. Em certas modalidades, uma sequência alvo para uma nuclease Cas está localizada perto da sequência de PAM cognato da nuclease. Em algumas modalidades, uma nuclease Cas de Classe 2 pode ser direcionada por um gRNA para uma sequência alvo de uma molécula de ácido nucleico alvo, em que o gRNA hibrida com e a proteína Cas de Classe 2 cliva a sequência alvo. Em algumas modalidades, o RNA guia hibrida com uma nuclease Cas de Classe 2 e cliva a sequência alvo adjacente ou compreendendo seu PAM cognato. Em algumas modalidades, a sequência alvo pode ser complementar à sequência de direcionamento do RNA guia. Em algumas modalidades, o grau de complementaridade entre uma sequência alvo de um RNA guia e a porção da sequência alvo correspondente que hibrida com o RNA guia pode ser de cerca de 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a percentagem de identidade entre uma sequência alvo de um RNA guia e a porção da sequência alvo correspondente que hibrida com o RNA guia pode ser de cerca de 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a região de homologia do alvo é adjacente a uma sequência PAM cognata. Em algumas modalidades, a sequência alvo pode compreender uma sequência 100% complementar à sequência de direcionamento do RNA guia. Em outras modalidades, a sequência alvo pode compreender pelo menos uma incompatibilidade, deleção ou inserção, em comparação com a sequência de direcionamento do RNA guia.

[00131] O comprimento da sequência alvo pode depender do sistema de nuclease utilizado. Por exemplo, a sequência de direcionamento de um RNA guia para um sistema CRISPR/Cas pode compreender 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 ou mais de 50 nucleotídeos de comprimento e a sequência alvo é um comprimento correspondente, opcionalmente adjacente a uma sequência PAM. Em algumas modalidades, a sequência alvo pode compreender 15-24 nucleotídeos de comprimento. Em algumas modalidades, a sequência alvo pode compreender 17-21 nucleotídeos de comprimento. Em algumas modalidades, a sequência alvo pode compreender 20 nucleotídeos de comprimento. Quando são utilizadas nickases, a sequência alvo pode compreender um par de sequências alvo reconhecidas por um par de nickases que clivam as cadeias opostas da molécula de DNA. Em algumas modalidades, a sequência alvo pode compreender um par de sequências alvo reconhecidas por um par de nickases que clivam as mesmas cadeias da molécula de DNA. Em algumas modalidades, a sequência alvo pode compreender uma parte das sequências alvo reconhecidas por uma ou mais nucleases Cas.

[00132] A molécula de ácido nucleico alvo pode ser qualquer molécula de DNA ou RNA que seja endógena ou exógena a uma célula. Em algumas modalidades, a molécula de ácido nucleico alvo pode ser um DNA episssomal, um plasmídeo, um DNA genômico, genoma viral, DNA mitocondrial ou DNA cromossômico de uma célula ou na célula. Em algumas modalidades, a sequência alvo da molécula de ácido nucleico alvo pode ser uma sequência genômica de uma célula ou de uma célula, incluindo uma célula humana.

[00133] Em outras modalidades, a sequência alvo pode ser uma sequência viral. Em outras modalidades, a sequência alvo pode ser uma sequência de patógenos. Em ainda outras modalidades, a se-

quência alvo pode ser uma sequência sintetizada. Em outras modalidades, a sequência alvo pode ser uma sequência cromossômica. Em certas modalidades, a sequência alvo pode compreender uma junção de translocação, Por exemplo, uma translocação associada a um câncer. Em algumas modalidades, a sequência alvo pode estar em um cromossomo eucariótico, como um cromossomo humano. Em certas modalidades, a sequência alvo é uma sequência específica do fígado, na medida em que é expressa nas células do fígado.

[00134] Em algumas modalidades, a sequência alvo pode estar localizada em uma sequência de codificação de um gene, uma sequência de íntron de um gene, uma sequência reguladora, uma sequência de controle transcricional de um gene, uma sequência de controle de tradução de um gene, um sítio de splicing ou uma sequência não codificante entre genes. Em algumas modalidades, o gene pode ser um gene codificador de proteína. Em outras modalidades, o gene pode ser um gene de RNA não codificante. Em algumas modalidades, a sequência alvo pode compreender todo ou uma porção de um gene associado à doença. Em algumas modalidades, a sequência alvo pode estar localizada em um sítio funcional não gênico no genoma, por exemplo, um sítio que controla aspectos da organização da cromatina, como um sítio de estrutura em andaime ou região de controle de locus.

[00135] Em modalidades que envolvem uma nuclease Cas, como uma nuclease Cas de Classe 2, a sequência alvo pode ser adjacente a um motivo adjacente do protospacer ("PAM"). Em algumas modalidades, o PAM pode ser adjacente a ou dentro de 1, 2, 3 ou 4 nucleotídeos da extremidade 3' da sequência alvo. O comprimento e a sequência do PAM podem depender da proteína Cas usada. Por exemplo, o PAM pode ser selecionado a partir de um consenso ou uma sequência PAM específica para uma proteína Cas9 ou ortólogo Cas9 específico, incluindo aqueles divulgados na Figura 1 de Ran et al., *Natu-*

re, 520: 186-191 (2015) e Figura S5 da Zetsche 2015, cuja descrição relevante de cada uma delas é incorporada por referência neste documento. Em algumas modalidades, o PAM pode ter 2, 3, 4, 5, 6, 7, 8, 9 ou 10 nucleotídeos de comprimento. As sequências de PAM exemplificativas não limitativas incluem NGG, NGGNG, NG, NAAAAN, NNAAAAW, NNNNACA, GNNNCNNA, TTN e NNNNGATT (em que N é definido como qualquer nucleotídeo e W é definido como A ou T). Em algumas modalidades, a sequência de PAM pode ser NGG. Em algumas modalidades, a sequência de PAM pode ser NGGNG. Em algumas modalidades, a sequência de PAM pode ser TTN. Em algumas modalidades, a sequência de PAM pode ser NNAAAAW.

Formulação Lipídica

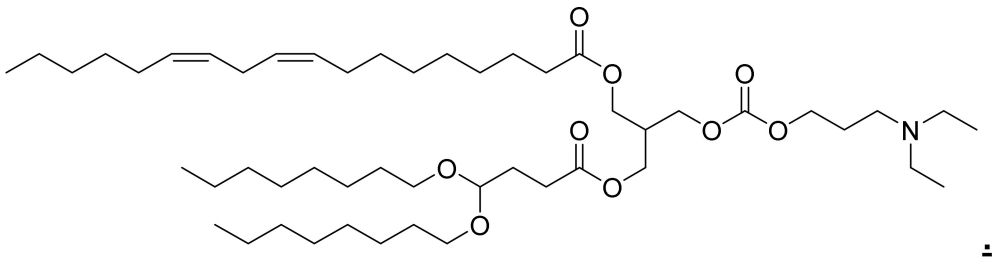
[00136] São divulgadas neste documento várias modalidades de formulações de LNP para agentes biologicamente ativos, como RNAs, incluindo cargas CRISPR/Cas. Tais formulações de LNP incluem um "lipídio de amina" ou "lipídio biodegradável", opcionalmente junto com um ou mais lipídios auxiliares, lipídios neutros e lipídios dissimulados, como lipídios com PEG. Por "nanopartícula lipídica" entende-se uma partícula que compreende uma pluralidade de (*isto é*, mais de uma) moléculas lipídicas fisicamente associadas entre si por forças intermoleculares.

Lipídios com amina

[00137] Em algumas modalidades, composições de LNP para a entrega de agentes biologicamente ativos compreendem um "lipídio de amina", que é definido como lipídio A ou seus equivalentes, incluindo análogos de acetal do Lipídio A.

[00138] Em algumas modalidades, o lipídio de amina é o lipídio A, que é (9Z, 12Z)-3-((4,4-bis(octilóxi)butanoil)óxi)-2-((((3-(dietilamino) propóxi)carbonil)óxi)metil)propil octadeca-9,12-dienoato, também chamado de 3-((4,4-bis(octilóxi)butanoil)óxi)-2- ((((3-(dietilamino)propóxi)

carbonil)óxi)metil)propil (9Z, 12Z)-octadeca-9,12-dienoato. O lipídio A pode ser descrito como:



[00139] O lipídio A pode ser sintetizado de acordo com o documento WO2015/095340 (Por exemplo, pp. 84-86). Em certas modalidades, o lipídio de amina é equivalente ao lipídio A.

[00140] Em certas modalidades, um lipídio de amina é um análogo do lipídio A. Em certas modalidades, um análogo de Lipídio A é um análogo de acetal do Lipídio A. Em determinadas composições de LNP, o análogo de acetal é um análogo de acetal C4-C12. Em algumas modalidades, o análogo de acetal é um análogo de acetal C5-C12. Em modalidades adicionais, o análogo de acetal é um análogo de acetal C5-C10. Em outras modalidades, o análogo de acetal é escolhido dentre um análogo de acetal de C4, C5, C6, C7, C9, C10, C11 e C12.

[00141] Os lipídios com amina e outros "lipídios biodegradáveis" adequados para uso nos LNPs descritos neste documento são biodegradáveis *in vivo*. Os lipídios com aminas têm baixa toxicidade (Por exemplo, são tolerados nos modelos animais sem efeito adverso em quantidades maiores ou iguais a 10 mg/kg). Em certas modalidades, os LNPs que compreendem um lipídio de amina incluem aqueles em que pelo menos 75% do lipídio de amina é removido do plasma dentro de 8, 10, 12, 24 ou 48 horas ou 3, 4, 5, 6, 7 ou 10 dias. Em certas modalidades, os LNPs que compreendem um lipídio de amina incluem aqueles em que pelo menos 50% do mRNA ou do gRNA são eliminados do plasma dentro de 8, 10, 12, 24 ou 48 horas ou 3, 4, 5, 6, 7, ou

10 dias. Em certas modalidades, os LNPs compreendendo um lipídio de amina incluem aqueles em que pelo menos 50% do LNP é eliminado do plasma dentro de 8, 10, 12, 24 ou 48 horas ou 3, 4, 5, 6, 7 ou 10 dias, por exemplo, medindo um lipídio (por exemplo, um lipídio de amina), RNA (por exemplo, mRNA) ou outro componente. Em certas modalidades, o componente lipídio encapsulado em lipídio versus lipídio livre, RNA ou ácido nucleico do LNP é medido.

[00142] Os lipídios biodegradáveis incluem, por exemplo, os lipídios biodegradáveis dos documentos WO/2017/173054, WO2015/095340 e WO2014/136086.

[00143] A depuração lipídica pode ser medida como descrito na literatura. *Veja* Maier, M.A., *et al*. Biodegradable Lipids Enabling Rapidly Eliminated Lipid Nanoparticles for Systemic Delivery of RNAi Therapeutics. *Mol. Ther.* 2013, 21(8), 1570-78 ("*Maier*"). Por exemplo, em *Maier*, os sistemas LNP-siRNA contendo siRNA direcionado às luciferases foram administradas aos camundongos machos C57Bl/6 de seis a oito semanas de idade a 0,3 mg/kg por injeção intravenosa em bolus *via* veia da tail lateral. Amostras de sangue, fígado e baço foram coletadas em 0,083, 0,25, 0,5, 1, 2, 4, 8, 24, 48, 96 e 168 horas após a dose. Os camundongos foram perfundidos com solução salina antes da coleta de tecido e as amostras de sangue foram processadas para obter plasma. Todas as amostras foram processadas e analisadas por LC-MS. Além disso, *Maier* descreve um procedimento para avaliar a toxicidade após a administração de formulações de LNP-siRNA. Por exemplo, um siRNA direcionado à luciferase foi administrado em 0, 1, 3, 5 e 10 mg/kg (5 animais/grupo) via injeção intravenosa em bolus intravenosa a um volume de dose de 5 mL/kg em ratos Sprague-Dawley machos. Após 24 horas, cerca de 1 mL de sangue foi obtido da veia jugular de animais conscientes e o soro foi isolado. 72 horas após a dose, todos os animais foram sacrificados para necropsia. Fo-

ram realizadas avaliações de sinais clínicos, peso corporal, química sérica, pesos de órgãos e histopatologia. Embora *Maier* descreva métodos para avaliar formulações de siRNA-LNP, esses métodos podem ser aplicados para avaliar a depuração, farmacocinética e toxicidade da administração de composições de LNP da presente invenção.

[00144] Os lipídios podem levar a um aumento da taxa de depuração. Em algumas modalidades, a taxa de depuração é uma taxa de depuração lipídica, por exemplo, a taxa na qual um lipídio é eliminado do sangue, soro ou plasma. Em algumas modalidades, a taxa de depuração é uma taxa de depuração de RNA, por exemplo, a taxa na qual um mRNA ou um gRNA é eliminado do sangue, soro ou plasma. Em algumas modalidades, a taxa de depuração é a taxa na qual o LNP é eliminado do sangue, soro ou plasma. Em algumas modalidades, a taxa de depuração é a taxa na qual o LNP é eliminado de um tecido, como tecido do fígado ou tecido do baço. Em certas modalidades, uma alta taxa de depuração leva a um perfil de segurança sem efeitos adversos substanciais. Os lipídios com amina e lipídios biodegradáveis podem reduzir o acúmulo de LNP na circulação e nos tecidos. Em algumas modalidades, uma redução no acúmulo de LNP na circulação e nos tecidos leva a um perfil de segurança sem efeitos adversos substanciais.

[00145] Os lipídios podem ser ionizáveis, dependendo do pH do meio em que estão. Por exemplo, em um meio levemente ácido, o lipídio, tal como os lipídios com amina podem ser protonados e, assim, carregar uma carga positiva. Por outro lado, em um meio ligeiramente básico, como, por exemplo, sangue, onde o pH é de aproximadamente 7,35, o lipídio, tal como um lipídio de amina pode não ser protonado e, portanto, não tem carga.

[00146] A capacidade de um lipídio suportar uma carga está relacionada ao seu pKa intrínseco. Em algumas modalidades, os lipídios

com amina da presente invenção podem cada um, independentemente, ter um pKa na faixa de cerca de 5,1 a cerca de 7,4. Em algumas modalidades, os lipídios biodisponíveis da presente invenção podem cada um, independentemente, ter um pKa na faixa de cerca de 5,1 a cerca de 7,4. Por exemplo, os lipídios com amina da presente invenção podem cada um, independentemente, ter um pKa na faixa de cerca de 5,8 a cerca de 6,5. Os lípidos com um pKa variando entre cerca de 5,1 e cerca de 7,4 são eficazes para a entrega de carga in vivo, por exemplo, para o fígado. Além disso, verificou-se que lipídios com um pKa variando entre cerca de 5,3 e cerca de 6,4 são eficazes para entrega in vivo, por exemplo, para tumores. *Ver*, Por exemplo, WO2014/136086.

Lipídios adicionais

[00147] "Lipídios neutros" adequados para uso em uma composição lipídica da descrição incluem, por exemplo, uma variedade de lipídios neutros, não carregados ou zwitteriônicos. Exemplos de fosfolipídios neutros adequados para uso na presente invenção incluem, mas não estão limitados a, 5-heptadecilbenzeno-1,3-diol (resorcinol), dipalmitoilfosfatidilcolina (DPPC), distearoilfosfatidilcolina (DSPC), fosfocolina (DOPC), dimiristoilfosfatidilcolina (DMPC), fosfatidilcolina (PLPC), 1,2-distearoil-sn-glicero-3-fosfocolina (DAPC), fosfatidiletanolamina (PE), fosfatidilcolina do ovo (EPC), dilaurililfosfatidilcolina (DLPC), dimiristoilfosfatidilcolina (DMPC), 1-miristoil-2-palmitoilfosfatidilcolina (MPPC), 1-palmitoil-2-miristoilfosfatidilcolina (PMPC), 1-palmitoil-2-estearoilfosfatidilcolina (PSPC), 1,2-diarachidoyl-sn-glicero-3-fosfocolina (DBPC), 1-estearoil-2-palmitoilfosfatidilcolina (SPPC), 1,2-dieicosenoil-sn-glicero-3-fosfocololina (DEPC), palmitoiloloilfosfatidilcolina (POPC), lisofosfatidilcolina, dioleoilfosfatidiletanolamina (DOPE), dilinoleoilfosfatidilcolina distearoilfosfatidiletanolamina (DSPE), dimiristoilfosfatidiletanolamina (DMPE), dipalmitoil fosfatidiletanolamina (DPPE), palmitoiloleoil fosfatidiletanolamina (POPE), lisofosfatidiletanilamina e combinações des-

tes. Numa modalidade, o fosfolipídio neutro pode ser selecionado a partir do grupo que consiste em distearoilfosfatidilcolina (DSPC) e dimiristoil fosfatidil etanolamina (DMPE). Noutra modalidade, o fosfolipídio neutro pode ser distearoilfosfatidilcolina (DSPC).

[00148] Os "lipídios auxiliares" incluem esteroides, esteróis e alquil resorcinóis. Os lipídios auxiliares adequados para uso na presente invenção incluem, mas não estão limitados ao colesterol, 5-heptadecilresorcinol e hemisuccinato de colesterol. Numa modalidade, o lipídio auxiliar pode ser colesterol. Numa modalidade, o lipídio auxiliar pode ser hemissuccinato de colesterol.

[00149] "Lipídios furtivos" são lipídios dissimulados que alteram o tempo em que as nanopartículas podem existir *in vivo* (Por exemplo, no sangue). Os lipídios furtivos podem auxiliar no processo de formulação, por exemplo, reduzindo a agregação de partículas e controlando o tamanho das partículas. Os lipídios furtivos utilizados neste documento podem modular as propriedades farmacocinéticas dos LNP. Os lipídios furtivos adequados para uso em uma composição lipídica da descrição incluem, mas não estão limitados aos lipídios furtivos com um grupo principal hidrofílico ligado a uma fração lipídica. Lípidos furtivos adequados para uso em uma composição lipídica da presente invenção e informações sobre a bioquímica de tais lipídios podem ser encontrados em Romberg *et al.*, Pharmaceutical Research, vol. 25, Nº 1, 2008, pág. 55-71 e Hoekstra *et al.*, Biochimica et Biophysica Acta 1660 (2004) 41-52. Lipídios com PEG adequados adicionais são divulgados, Por exemplo, no documento WO 2006/007712.

[00150] Em uma modalidade, o grupo principal hidrofílico de lipídio furtivo compreende uma fração de polímero selecionada a partir de polímeros baseados em PEG. Os lipídios furtivos podem compreender uma fração lipídica. Em algumas modalidades, o lipídio furtivo é um lipídio com PEG.

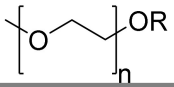
[00151] Em uma modalidade, um lipídio furtivo compreende uma fração de polímero selecionada a partir de polímeros baseados em PEG (às vezes chamado de poli(óxido de etileno)), poli(oxazolina), poli(álcool vinílico), poli(glicerol), poli(N-vinilpirrolidona), poliaminoácidos e poli[N-(2-hidroxipropil)metacrilamida].

[00152] Em uma modalidade, o lipídio com PEG compreende uma fração de polímero à base de PEG (às vezes referido como poli(óxido de etileno)).

[00153] O lipídio com PEG compreende ainda uma fração lipídica. Em algumas modalidades, a fração lipídica pode ser derivada de diacilglicerol ou diacilglicamida, incluindo aqueles que compreendem um grupo dialquilglicerol ou dialquilglicamida com comprimento de cadeia alquil independentemente compreendendo de cerca de C4 a cerca de C40 átomos de carbono saturados ou insaturados, em que a cadeia pode compreender um ou mais átomos de carbono funcionais grupos como, por exemplo, uma amida ou éster. Em algumas modalidades, o comprimento da cadeia alquil compreende cerca de C10 a C20. O grupo dialquilglicerol ou dialquilglicamida pode ainda compreender um ou mais grupos alquil substituídos. Os comprimentos da cadeia podem ser simétricos ou assimétricos.

[00154] Salvo indicação em contrário, o termo "PEG", conforme usado neste documento, significa qualquer polietileno glicol ou outro polímero de éter polialquileno. Numa modalidade, PEG é um polímero linear ou ramificado opcionalmente substituído de etileno glicol ou óxido de etileno. Numa modalidade, o PEG não é substituído. Em uma modalidade, o PEG é substituído, Por exemplo, por um ou mais grupos alquil, alcóxi, acil, hidróxi ou arila. Em uma modalidade, o termo inclui copolímeros de PEG, como PEG-poliuretano ou PEG-polipropileno (ver, Por exemplo, J. Milton Harris, Química de poli(etileno glicol): aplicações biotecnológicas e biomédicas (1992)); em outra modalidade, o

termo não inclui copolímeros de PEG. Em uma modalidade, o PEG tem um peso molecular de cerca de 130 a cerca de 50.000, em uma sub-modalidade, cerca de 150 a cerca de 30.000, em uma sub-modalidade, cerca de 150 a cerca de 20.000, em uma sub-modalidade de cerca de 150 a cerca de 15.000, em uma sub-modalidade, cerca de 150 a cerca de 10.000, em uma sub-modalidade, cerca de 150 a cerca de 6.000, em uma sub-modalidade, cerca de 150 a cerca de 5.000, em uma sub-modalidade, cerca de 150 a cerca de 4.000, em uma sub-modalidade, cerca de 150 a cerca de 3.000, em uma sub-modalidade, cerca de 300 a cerca de 3.000, em uma sub-modalidade, cerca de 1.000 a cerca de 3.000 e em uma sub-modalidade, cerca de 1.500 a cerca de 2.500.

[00155] Em certas modalidades, o PEG (Por exemplo, conjugado a uma porção lipídica ou lipídica, como um lipídio furtivo), é um "PEG-2K", também denominado "PEG 2000", que tem um peso molecular médio de cerca de 2.000 daltons. O PEG-2K é representado neste documento pela seguinte fórmula (I), em que n é 45, significando que o grau médio de polimerização do número compreende cerca de 45 subunidades  (I) No entanto, outras modalidades de PEG

dalidades, n pode variar de cerca de 30 a cerca de 60. Em algumas modalidades, n pode variar de cerca de 35 a cerca de 55. Em algumas modalidades, n pode variar de cerca de 40 a cerca de 50. Em algumas modalidades, n pode variar de cerca de 42 a cerca de 48. Em algumas modalidades, n pode ser 45. Em algumas modalidades, R pode ser selecionado dentre H, alquil substituído e alquil não substituído. Em algumas modalidades, R pode ser alquil não substituído. Em algumas

modalidades, R pode ser metila.

[00156] Em qualquer uma das modalidades descritas neste documento, o lipídio com PEG pode ser selecionado a partir de PEG-dilauroilglicerol, PEG-dimiristoilglicerol (PEG-DMG) (# de catálogo GM-020 da NOF, Tóquio, Japão), dipalmitoilglicerol PEG, distearoilglicerol PEG (PEG-DSPE) (# de catálogo DSPE-020CN, NOF, Tóquio, Japão), PEG-dilaurilglicamida, PEG-dimiristilglicamida, PEG-dipalmitoilglicamida e PEG-distearoililglicamida, PEG-colesterol (1-[8'-(Colest-5-en -3[beta]-óxi)carboxamido-3',6'-dioxaoctanil]carbamoil-[ômega]-metil-poli (etileno glicol), DMB PEG (3,4-ditetradecoxilbenzil- [ômega]-metil-poli(etileno glicol)éter), 1,2-dimiristoil-sn-glicero-3-fosfoetanolamina-N-[metóxi (polietileno glicol)-2000] (PEG2k-DMG) (cat. #880150P de Avanti Polar Lipids, Alabaster, Alabama, EUA), 1,2-distearoil-sn-glicero-3-fosfoetanolamina-N-[metóxi (polietileno glicol)-2000] (PEG2k-DSPE) (cat. #880120C de Avanti Polar Lipids, Alabaster, Alabama, EUA), 1,2-distearoil-sn-glicerol, metoxipolietilenoglicol (PEG2k-DSG; GS-020, NOF Tóquio, Japão), poli(etileno glicol)-2000-dimetacrilato (PEG2k-DMA) e 1,2-distеariloxipropil-3-amina-N-[metóxi (polietileno glicol)-2000] (PEG2k-DSA). Numa modalidade, o lipídio com PEG pode ser PEG2k-DMG. Em algumas modalidades, o lipídio do PEG pode ser PEG2k-DSG. Numa modalidade, o lipídio com PEG pode ser PEG2k-DSPE. Numa modalidade, o lipídio com PEG pode ser PEG2k-DMA. Numa modalidade, o lipídio com PEG pode ser PEG2k-DMA. Em uma modalidade, o lipídio com PEG pode ser o composto S027, divulgado no documento WO2016/010840 (parágrafos [00240] a [00244]). Numa modalidade, o lipídio com PEG pode ser PEG2k-DSA. Numa modalidade, o lipídio com PEG pode ser PEG2k-C11. Em algumas modalidades, o lipídio com PEG pode ser PEG2k-C14. Em algumas modalidades, o lipídio com PEG pode ser PEG2k-C16. Em algumas modalidades, o lipídio com PEG pode ser PEG2k-C18.

Formulações de LNP

[00157] O LNP pode conter (i) um lipído biodegradável, (ii) um lipídio neutro opcional, (iii) um lipídio auxiliar e (iv) um lipídio furtivo, tal como um lipidio com PEG. O LNP pode conter um lipídio biodegradável e um ou mais lipídios neutros, lipídios auxiliares e lipídios dissimulados, como lipídios com PEG.

[00158] O LNP pode conter (i) um lipídio de amina para encapsulamento e escape endossômico, (ii) um lipídio neutro para estabilização, (iii) um lipídio auxiliar, também para estabilização e (iv) um lipídio furtivo, como um lipídio com PEG. O LNP pode conter um lipídio de amina e um ou mais lipídios neutros, lipídios auxiliares, também para estabilização, e lipídios dissimulados, como lipídios com PEG.

[00159] Em algumas modalidades, uma composição de LNP pode compreender um componente de RNA que inclui um ou mais de um agente de ligação a DNA guiado a RNA, um mRNA da nuclease Cas, um mRNA da nuclease Cas de Classe 2, um mRNA de Cas9 e um gRNA. Em algumas modalidades, uma composição de LNP pode incluir uma nuclease Cas de Classe 2 e um gRNA como componente de RNA. Em certas modalidades, uma composição de LNP pode compreender o componente de RNA, um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio furtivo. Em certas composições de LNP, o lipídio auxiliar é o colesterol. Noutras composições, o lipídio neutro é DSPC. Em modalidades adicionais, o lipídio furtivo é PEG2k-DMG ou PEG2k-C11. Em certas modalidades, a composição de LNP compreende lipídio A ou um equivalente de lipídio A; um lipídio auxiliar; um lipídio neutro; um lipídio furtivo; e um RNA guia. Em certas composições, o lipídio de amina é o lipídio A. Em certas composições, o lipídio de amina é o lipídio A ou um análogo de acetal deste; o lipídio auxiliar é o colesterol; o lipídio neutro é DSPC; e o lipídio furtivo é o PEG2k-DMG.

[00160] Em certas modalidades, as composições lipídicas são descritas de acordo com as respectivas razões molares dos lipídios componentes na formulação. Modalidades da presente invenção fornecem composições lipídicas descritas de acordo com as respectivas proporções molares dos lipídios componentes na formulação. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 30% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 40% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 45% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 50% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 55% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 50% molar a cerca de 55% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 50% molar. Numa modalidade, a% molar do lipídio de amina pode ser de cerca de 55% molar. Em algumas modalidades, a% molar do lipídio de amina do lote de LNP será de ± 30%, ± 25%, ± 20%, ± 15%, ± 10%, ± 5% ou ± 2,5% da% molar alvo. Em algumas modalidades, a% molar do lipídio de amina do lote de LNP será de ± 4% molar, ± 3% molar, ± 2% molar, ± 1,5% molar, ± 1% molar, ± 0,5% molar, ou ± 0,25% molar da% alvo molar. Todos os números% molares são dados como uma fração do componente lipídico das composições de LNP. Em certas modalidades, a variabilidade entre lotes de LNP da% molar do lipídio de amina será menor que 15%, menor que 10% ou menor que 5%.

[00161] Numa modalidade, a % molar do lipídio neutro pode ser de cerca de 5% molar a cerca de 15% molar. Numa modalidade, a% molar do lipídio neutro pode ser de cerca de 7% molar a cerca de 12% molar. Numa modalidade, a% molar do lipídio natural pode ser de cerca de 9% molar. Em algumas modalidades, a% molar do lipídio neutro

do lote de LNP será de ± 30%, ± 25%, ± 20%, ± 15%, ± 10%, ± 5% ou ± 2,5% da% molar do alvo lipídico neutro. Em certas modalidades, a variabilidade entre lotes de LNP será menor que 15%, menor que 10% ou menor que 5%.

[00162] Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 20% molar a cerca de 60% molar. Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 25% molar a cerca de 55% molar. Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 25% molar a cerca de 50% molar. Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 25% molar a cerca de 40% molar. Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 30% molar a cerca de 50% molar. Numa modalidade, a% molar do lipídio auxiliar pode ser de cerca de 30% molar a cerca de 40% molar. Em uma modalidade, a% molar do lipídio auxiliar é ajustada com base nas concentrações de lipídio de amina, lipídio neutro e lipídio com PEG para levar o componente lipídico a 100% molar. Em algumas modalidades, a% auxiliar em mol do lote de LNP será de ± 30%, ± 25%, ± 20%, ± 15%, ± 10%, ± 5% ou ± 2,5% da% molar alvo. Em certas modalidades, a variabilidade entre lotes de LNP será menor que 15%, menor que 10% ou menor que 5%.

[00163] Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 1% molar a cerca de 10% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 2% molar a cerca de 10% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 2% molar a cerca de 8% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 2% molar a cerca de 4% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 2,5% molar a cerca de 4% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 3% molar. Numa modalidade, a% molar do lipídio com PEG pode ser de cerca de 2,5%

molar. Em algumas modalidades, a% molar do lipídio com PEG do lote de LNP será de ± 30%, ± 25%, ± 20%, ± 15%, ± 10%, ± 5% ou ± 2,5% da% molar do alvo lipídico PEG. Em certas modalidades, a variabilidade entre lotes de LNP será menor que 15%, menor que 10% ou menor que 5%.

[00164] Em certas modalidades, a carga inclui um mRNA que codifica um agente de ligação a DNA guiado por RNA (por exemplo, uma nuclease Cas, uma nuclease Cas ou Cas9 de Classe 2) e um gRNA ou um ácido nucleico que codifica um gRNA ou uma combinação de mRNA e gRNA. Numa modalidade, uma composição de LNP pode compreender um lipídio A ou seus equivalentes. Em alguns aspectos, o lipídio de amina é o lipídio A. Em alguns aspectos, o lipídio de amina é um equivalente do lipídio A, por exemplo, um análogo do lipídio A. Em certos aspectos, o lipídio de amina é um análogo acetal do lipídio A. Em várias modalidades, uma composição de LNP compreende um lipídio de amina, um lipídio neutro, um lipídio auxiliar e um lipídio com PEG. Em certas modalidades, o lipídio auxiliar é o colesterol. Em certas modalidades, o lipídio neutro é DSPC. Em modalidades específicas, o lipídio com PEG é PEG2k-DMG. Em algumas modalidades, uma composição de LNP pode compreender um lipídio A, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em algumas modalidades, uma composição de LNP compreende um lipídio de amina, DSPC, colesterol e um lipídio com PEG. Em algumas modalidades, a composição de LNP compreende um lipídio com PEG compreendendo DMG. Em certas modalidades, o lipídio de amina é selecionado a partir do lipídio A e um equivalente do lipídio A, incluindo um análogo de acetal do lipídio A. Em modalidades adicionais, uma composição de LNP compreende lipídio A, colesterol, DSPC e PEG2k-DMG.

[00165] Modalidades da presente invenção também fornecem composições lipídicas descritas de acordo com a razão molar entre os gru-

pos amina com carga positiva do lipídio de amina (N) e os grupos fosfato com carga negativa (P) do ácido nucleico a ser encapsulado. Isso pode ser matematicamente representado pela equação N/P. Em algumas modalidades, uma composição de LNP pode compreender um componente lipídico que compreende um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio auxiliar; e um componente de ácido nucleico, em que a razão N/P é de cerca de 3 a 10. Em algumas modalidades, uma composição de LNP pode compreender um componente lipídico que compreende um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio auxiliar; e um componente de RNA, em que a razão N/P é de cerca de 3 a 10. Numa modalidade, a razão N/P pode ser de cerca de 5 a 7. Numa modalidade, a razão N/P pode ser de cerca de 4,5 a 8. Numa modalidade, a razão N/P pode ser de cerca de 6. Numa modalidade, a razão N/P pode ser 6 ± 1. Numa modalidade, a razão N/P pode ser de cerca de 6 ± 0,5. Em algumas modalidades, a razão N/P será de ± 30%, ± 25%, ± 20%, ± 15%, ± 10%, ± 5% ou ± 2,5% da razão N/P alvo. Em certas modalidades, a variabilidade entre lotes de LNP será menor que 15%, menor que 10% ou menor que 5%.

[00166] Em algumas modalidades, o componente de RNA pode compreender um mRNA, como um mRNA que codifica uma nuclease Cas. Numa modalidade, o componente de RNA pode compreender um mRNA de Cas9. Em algumas composições compreendendo um mRNA que codifica uma nuclease Cas, o LNP compreende ainda um ácido nucleico de gRNA, como um gRNA. Em algumas modalidades, o componente de RNA compreende um mRNA da nuclease Cas e um gRNA. Em algumas modalidades, o componente de RNA compreende um mRNA de nuclease Cas de Classe 2 e um gRNA.

[00167] Em certas modalidades, uma composição de LNP pode compreender um mRNA que codifica uma nuclease Cas, como uma

nuclease Cas de Classe 2, uma lipina com amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em certas composições de LNP compreendendo um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2, o lipídio auxiliar é o colesterol. Noutras composições compreendendo um mRNA que codifica uma nuclease Cas tal como uma nuclease Cas de Classe 2, o lipídio neutro é DSPC. Em modalidades adicionais compreendendo um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2, o lipídio com PEG é PEG2k-DMG ou PEG2k-C11. Em composições específicas que compreendem um mRNA que codifica uma nuclease Cas, como uma nuclease Cas de Classe 2, o lipídio de amina é selecionado a partir do lipídio A e seus equivalentes, como um análogo de acetal do lipídio A.

[00168] Em algumas modalidades, uma composição de LNP pode compreender um gRNA. Em certas modalidades, uma composição de LNP pode compreender um lipídio de amina, um gRNA, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em certas composições de LNP compreendendo um gRNA, o lipídio auxiliar é o colesterol. Em algumas composições compreendendo um gRNA, o lipídio neutro é DSPC. Em modalidades adicionais compreendendo um gRNA, o lipídio com PEG é PEG2k-DMG ou PEG2k-C11. Em certas modalidades, o lipídio de amina é selecionado a partir do lipídio A e seus equivalentes, como um análogo de acetal do lipídio A.

[00169] Numa modalidade, uma composição de LNP pode compreender um sgRNA. Numa modalidade, uma composição de LNP pode compreender um sgRNA de Cas9. Numa modalidade, uma composição de LNP pode compreender um sgRNA de Cpf1. Em algumas composições compreendendo um sgRNA, o LNP inclui um lipídio de amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em certas composições compreendendo um sgRNA, o lipídio auxiliar é o colesterol. Noutras composições compreendendo um sgRNA, o lipídio

neutro é DSPC. Em modalidades adicionais compreendendo um sgRNA, o lipídio com PEG é PEG2k-DMG ou PEG2k-C11. Em certas modalidades, o lipídio de amina é selecionado a partir do lipídio A e seus equivalentes, como um análogo de acetal do lipídio A.

[00170] Em certas modalidades, uma composição de LNP compreende um mRNA que codifica uma nuclease Cas e um gRNA, que pode ser um sgRNA. Em uma modalidade, uma composição de LNP pode compreender um lipídio de amina, um mRNA que codifica uma nuclease Cas, um gRNA, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG. Em certas composições compreendendo um mRNA que codifica uma nuclease Cas e um gRNA, o lipídio auxiliar é o colesterol. Em algumas composições compreendendo um mRNA que codifica uma nuclease Cas e um gRNA, o lipídio neutro é DSPC. Em modalidades adicionais compreendendo um mRNA que codifica uma nuclease Cas e um gRNA, o lipídio com PEG é PEG2k-DMG ou PEG2k-C11. Em certas modalidades, o lipídio de amina é selecionado a partir do lipídio A e seus equivalentes, como um análogo de acetal do lipídio A.

[00171] Em certas modalidades, as composições de LNP incluem um mRNA de nuclease Cas, como um mRNA Cas de Classe 2 e pelo menos um gRNA. Em certas modalidades, a composição de LNP inclui uma relação de gRNA para mRNA de nuclease Cas, como mRNA de nuclease Cas de Classe 2 de cerca de 25:1 a cerca de 1:25. Em certas modalidades, a formulação de LNP inclui uma relação de gRNA para mRNA de nuclease Cas, como mRNA de nuclease de Cas de Classe 2 de cerca de 10:1 a cerca de 1:10. Em certas modalidades, a formulação de LNP inclui uma relação de gRNA para mRNA de nuclease Cas, como o mRNA de nuclease Cas de Classe 2 de cerca de 8:1 a cerca de 1:8. Conforme medido neste documento, as relações são em peso. Em algumas modalidades, a formulação de LNP inclui uma relação de

gRNA para mRNA de nuclease Cas, como mRNA de Cas de Classe 2 de cerca de 5:1 a cerca de 1:5. Em algumas modalidades, o intervalo de relações é de cerca de 3:1 a 1:3, cerca de 2:1 a 1:2, cerca de 5:1 a 1:2, cerca de 5:1 a 1:1, cerca de 3:1 a 1:2, cerca de 3:1 a 1:1, cerca de 3:1, cerca de 2:1 a 1:1. Em algumas modalidades, a relação de gRNA para mRNA é de cerca de 3:1 ou cerca de 2:1 Em algumas modalidades, a relação de gRNA para mRNA de nuclease Cas, como a nuclease Cas de Classe 2 é de cerca de 1:1. A relação pode ser de 25:1, 10:1, 5:1, 3:1, 1:1, 1:3, 1:5, 1:10 ou 1:25.

[00172] As composições de LNP divulgadas neste documento podem incluir um ácido nucleico modelo. O ácido nucleico modelo pode ser co-formulado com um mRNA que codifica uma nuclease Cas, tal como um mRNA de nuclease Cas de Classe 2. Em algumas modalidades, o ácido nucleico modelo pode ser co-formulado com um RNA guia. Em algumas modalidades, o ácido nucleico modelo pode ser co-formulado com um mRNA que codifica uma nuclease Cas e um RNA guia. Em algumas modalidades, o ácido nucleico modelo pode ser formulado separadamente de um mRNA que codifica uma nuclease Cas ou um RNA guia. O ácido nucleico modelo pode ser entregue com ou separadamente das composições de LNP. Em algumas modalidades, o ácido nucleico modelo pode ser de fita simples ou dupla, dependendo do mecanismo de reparo desejado. O modelo pode ter regiões de homologia para o DNA alvo ou para sequências adjacentes ao DNA alvo.

[00173] Em algumas modalidades, os LNPs são formados misturando uma solução aquosa de RNA com uma solução lipídica à base de solvente orgânico, por exemplo, 100% de etanol. Soluções ou solventes adequados incluem ou podem conter: água, PBS, tampão Tris, NaCl, tampão citrato, etanol, clorofórmio, éter dietílico, ciclo-hexano, tetra-hidrofurano, metanol, isopropanol. Um tampão farmaceuticamen-

te aceitável, por exemplo, para administração *in vivo* de LNPs, pode ser usado. Em certas modalidades, um tampão é usado para manter o pH da composição compreendendo LNPs em pH 6,5 ou acima. Em certas modalidades, um tampão é usado para manter o pH da composição compreendendo LNPs em pH 7,0 ou acima. Em certas modalidades, a composição tem um pH variando de cerca de 7,2 a cerca de 7,7. Em modalidades adicionais, a composição tem um pH variando de cerca de 7,3 a cerca de 7,7 ou variando de cerca de 7,4 a cerca de 7,6. Em outras modalidades, a composição tem um pH de cerca de 7,2, 7,3, 7,4, 7,5, 7,6 ou 7,7. O pH de uma composição pode ser medido com uma sonda de micro pH. Em certas modalidades, um crioprotetor é incluído na composição. Exemplos não limitativos de crioprotetores incluem sacarose, trealose, glicerol, DMSO e etileno glicol. Composições exemplares podem incluir até 10% de crioprotetor, como, por exemplo, sacarose. Em certas modalidades, a composição de LNP pode incluir cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9 ou 10% de crioprotetor. Em certas modalidades, a composição de LNP pode incluir cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9 ou 10% de sacarose. Em algumas modalidades, a composição de LNP pode incluir um tampão. Em algumas modalidades, o tampão pode compreender um tampão fosfato (PBS), um tampão Tris, um tampão citrato e suas misturas. Em certas modalidades exemplares, o tampão compreende NaCl. Em certas modalidades, o NaCl é omitido. Quantidades exemplares de NaCl podem variar de cerca de 20 mM a cerca de 45 mM. Quantidades exemplares de NaCl podem variar de cerca de 40 mM a cerca de 50 mM. Em algumas modalidades, a quantidade de NaCl é de cerca de 45 mM. Em algumas modalidades, o tampão é um tampão Tris. Quantidades exemplares de Tris podem variar de cerca de 20 mM a cerca de 60 mM. Quantidades exemplares de Tris podem variar de cerca de 40 mM a cerca de 60 mM. Em algumas modalidades, a quantidade de Tris é de cerca de 50

mM. Em algumas modalidades, o tampão compreende NaCl e Tris. Certas modalidades exemplares das composições de LNP contêm sacarose a 5% e 45 mM de NaCl em tampão Tris. Em outras modalidades exemplares, as composições contêm sacarose em uma quantidade de cerca de 5% p/v, cerca de 45 mM de NaCl e cerca de 50 mM de Tris a pH 7,5. As quantidades de sal, tampão e crioprotetor podem variar de tal modo que a osmolalidade da formulação geral seja mantida. Por exemplo, a osmolalidade final pode ser mantida em menos de 450 mOsm/L. Em outras modalidades, a osmolalidade está entre 350 e 250 mOsm/L. Certas modalidades têm uma osmolalidade final de 300 +/- 20 mOsm/L.

[00174] Em algumas modalidades, mistura microfluídica, mistura T ou mistura cruzada é usada. Em certos aspectos, as taxas de fluxo, tamanho da junção, geometria da junção, formato da junção, diâmetro do tubo, soluções e/ou concentrações de RNA e lipídios podem variar. As composições de LNPs ou LNP podem ser concentradas ou purificadas, Por exemplo, através da diálise, filtração de fluxo tangencial ou cromatografia. Os LNPs podem ser armazenados como uma suspensão, uma emulsão ou um pó liofilizado, por exemplo. Em algumas modalidades, uma composição de LNP é armazenada a 2-8°C, em certos aspectos, as composições de LNP são armazenadas à temperatura ambiente. Em modalidades adicionais, uma composição de LNP é armazenada congelada, por exemplo, a -20°C ou -80°C. Em outras modalidades, uma composição de LNP é armazenada a uma temperatura variando de cerca de 0°C a cerca de -80°C. As composições de LNP congeladas podem ser descongeladas antes do uso, por exemplo no gelo a 4°C à temperatura ambiente ou a 25°C. As composições de LNP congeladas podem ser mantidas a várias temperaturas, por exemplo no gelo a 4°C à temperatura ambiente a 25°C ou a 37°C.

Métodos de manipulação de células-tronco, por exemplo, HSPCs; Células-tronco manipuladas, por exemplo, HSPCs

[00175] As composições de LNP divulgadas neste documento podem ser usadas em métodos para manipulação de células-tronco, por exemplo, HSPCs, por exemplo, pela edição de genes do sistema CRISPR/Cas *in vitro*. Em algumas modalidades, a população de células geneticamente manipuladas é uma população de células CD34+. Em algumas modalidades, é fornecido um método para produzir uma população de células HSPC ou CD34+ geneticamente manipuladas*in vitro*, o método compreendendo (a) pré-incubação de um fator sérico com uma composição de LNP para a entrega de um mRNA da nuclease Cas e um gRNA; (b) entrando em contato com a população de células HSPC ou CD34+ com a composição pré-incubada de LNP *in vitro*; e (c) cultivando a população de células HSPC ou CD34+ *in vitro*, produzindo assim um HSPC geneticamente manipulado. Em algumas modalidades, os métodos envolvem o contato de uma célula HSPC ou CD34+ com uma composição de LNP descrita neste documento de acordo com os métodos de entrega descritos nesta descrição.

[00176] Em algumas modalidades, as células-tronco manipuladas, por exemplo, HSPCs, são fornecidas, por exemplo, uma população de HSPC ou HSPC manipulada. Tais células manipuladas são produzidas de acordo com os métodos descritos neste documento. Em algumas modalidades, o HSPC manipulado reside dentro de um tecido ou órgão, por exemplo, medula óssea, sangue ou outro tecido dentro de um sujeito, por exemplo, após o transplante de um HSPC manipulado.

[00177] Em alguns dos métodos e células descritos neste documento, uma célula compreende uma modificação, por exemplo, uma inserção ou deleção("indel") ou substituição de nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma inserção de 1, 2, 3, 4 ou 5 ou mais nucleotídeos em uma sequên-

cia alvo. Em algumas modalidades, a modificação compreende uma inserção de 1 ou 2 nucleotídeos em uma sequência alvo. Em outras modalidades, a modificação compreende uma deleção de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 ou 25 ou mais nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma deleção de 1 ou 2 nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende um indel que resulta em uma mutação de deslocamento de frameshift em uma sequência alvo. Em algumas modalidades, a modificação compreende uma substituição de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 ou 25 ou mais nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma substituição de 1 ou 2 nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende um ou mais de uma inserção, deleção ou substituição de nucleotídeos resultantes da incorporação de um ácido nucleico padrão, por exemplo, qualquer um dos ácidos nucleicos padrão descritos neste documento.

[00178] Em algumas modalidades, é fornecida uma população de células compreendendo células manipuladas, por exemplo, uma população de células compreendendo células manipuladas de acordo com os métodos descritos neste documento. Em algumas modalidades, a população compreende células manipuladas cultivadas *in vitro*. Em algumas modalidades, a população reside dentro de um tecido ou órgão, por exemplo, um fígado dentro de um sujeito. Em algumas modalidades, pelo menos 5%, pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 30%, pelo menos 35%, pelo menos 40%, pelo menos 45%, pelo menos 50%, pelo menos 55%, pelo menos 60%, pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90% ou pelo menos 95% ou mais das células da população é manipulada. Em certas modalidades, um método divulgado neste documento resulta em pelo

menos 5%, pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 30%, pelo menos 35%, pelo menos 40%, pelo menos 45%, pelo menos 50%, pelo menos 55%, pelo menos 60%, pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90% ou pelo menos 95% de eficiência de edição (ou "percentagem de edição"), definida pela detecção de indels. Em outras modalidades, um método divulgado neste documento resulta em pelo menos 5%, pelo menos 10%, pelo menos 15%, pelo menos 20%, pelo menos 25%, pelo menos 30%, pelo menos 35%, pelo menos 40%, pelo menos 45%, pelo menos 50%, pelo menos 55%, pelo menos 60%, pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90% ou pelo menos 95% de eficiência na modificação do DNA, definida pela detecção de uma mudança na sequência, seja por inserção, deleção, substituição ou de outra forma. Em certas modalidades, um método neste documento divulgado resulta em um nível de eficiência de edição ou um nível de eficiência de modificação de DNA entre cerca de 5% a cerca de 100%, cerca de 10% a cerca de 50%, cerca de 20 a cerca de 100%, cerca de 20 a cerca de 80%, cerca de 40 a cerca de 100% ou cerca de 40 a cerca de 80% em uma população de células.

[00179] Em alguns dos métodos e células descritos neste documento, as células dentro da população compreendem uma modificação, por exemplo, um indel ou substituição em uma sequência alvo. Em algumas modalidades, a modificação compreende uma inserção de 1, 2, 3, 4 ou 5 ou mais nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma inserção de 1 ou 2 nucleotídeos em uma sequência alvo. Em outras modalidades, a modificação compreende uma deleção de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 ou 25 ou mais nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma deleção de 1 ou 2 nucleo-

tídeos em uma sequência alvo. Em algumas modalidades, a modificação resulta em uma mutação de deslocamento de frameshift em uma sequência alvo. Em algumas modalidades, a modificação compreende um indel que resulta em uma mutação de deslocamento de frameshift em uma sequência alvo. Em algumas modalidades, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos pelo menos 97%, pelo menos 98% ou pelo menos 99% ou mais das células manipuladas na população compreendem uma mutação de deslocamento de frameshift. Em algumas modalidades, a modificação compreende uma substituição de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 ou 25 ou mais nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende uma substituição de 1 ou 2 nucleotídeos em uma sequência alvo. Em algumas modalidades, a modificação compreende um ou mais de uma inserção, deleção ou substituição de nucleotídeos resultantes da incorporação de um ácido nucleico padrão, por exemplo, qualquer um dos ácidos nucleicos padrão descritos neste documento.

Métodos de edição de genes

[00180] Os métodos divulgados neste documento podem ser utilizados para edição de genes em uma célula-tronco, uma população de HSPC ou HSPC *in vitro*. Em uma modalidade, uma ou mais composições de LNP descritas neste documento podem ser administradas a uma célula-tronco, uma HSPC ou uma população de HSPC. Em uma modalidade, uma ou mais composições de LNP descritas neste documento podem entrar em contato com uma célula-tronco, um HSPC e HSC ou um HPC. Em uma modalidade, uma célula geneticamente modificada pode ser produzida entrando em contato com uma célula com uma composição de LNP de acordo com os métodos descritos neste documento. Em alguns métodos de edição de genes, a popula-

ção de HSPC ou HSPC é mantida em cultura. Em alguns métodos de edição de genes, a população de HSPC ou HSPC é transplantada para um paciente. Em algumas modalidades, o HSPC geneticamente manipulado reside dentro de um tecido ou órgão, por exemplo, medula óssea, sangue ou outro tecido dentro de um paciente, por exemplo, após o transplante de um HSPC manipulado.

[00181] Em algumas modalidades, o método compreende uma população de células-tronco, uma HSPC ou HSPC autóloga em relação a um paciente a ser administrado à célula. Em algumas modalidades, o método compreende uma população de HSPC ou HSPC que é alogênica com relação a um paciente para ser administrada a referida célula.

[00182] Em várias modalidades, os métodos descritos neste documento atingem a edição do gene CRISPR-Cas na população de células-tronco, HSPC ou HSPC. Em algumas modalidades, os métodos compreendem ainda detectar a edição de genes na população de HSPC ou HSPC. Em algumas modalidades, a edição de genes é medida como percentagem de edição. Em algumas modalidades, as edições dos genes são medidas como percentagem de modificação do DNA. Os métodos podem atingir pelo menos 40, 50, 60, 70, 80, 90 ou 95% de edição. Os métodos podem alcançar pelo menos 40, 50, 60, 70, 80, 90 ou 95% de modificação do DNA.

[00183] Em uma modalidade, uma composição de LNP compreendendo um mRNA que codifica uma nuclease Cas de Classe 2 e um gRNA pode ser administrado a uma célula-tronco, uma HSPC ou uma população de HSPC. Em modalidades adicionais, um ácido nucleico modelo também é introduzido na célula. Em certos casos, uma composição de LNP compreendendo uma nuclease Cas de Classe 2 e um sgRNA pode ser administrada a uma célula.

[00184] Numa modalidade, as composições de LNP podem ser

usadas para editar um gene em uma célula-tronco, uma população de HSPC ou HSPC, resultando em um nocaute genético. Em uma modalidade, as composições de LNP podem ser usadas para editar um gene em uma população de HSPC ou HSPC resultando em knockdown de genes, por exemplo, na população de células. O knockdown ou nocaute pode ser detectado medindo os níveis de proteína alvo. O knockdown ou nocaute pode ser detectado através da detecção do DNA alvo. Em outra modalidade, as composições de LNP podem ser usadas para editar um gene em uma população de HSPC ou HSPC, resultando em uma correção de gene. Em uma modalidade adicional, as composições de LNP podem ser usadas para editar uma célula que resulta na inserção de genes.

[00185] As composições de LNP podem ser administradas como uma formulação em associação com um ou mais excipientes farmaceuticamente aceitáveis. O termo "excipiente" inclui qualquer ingrediente que não seja (m) o (s) composto (s) da descrição, o (s) outro (s) componente (s) lipídico (s) e o agente biologicamente ativo. Um excipiente pode conferir uma característica funcional (Por exemplo, controle da taxa de liberação do medicamento) e/ou uma característica não funcional (Por exemplo, auxiliar de processamento ou diluente) às formulações. A escolha do excipiente dependerá, em grande medida, de fatores como o modo particular de administração, o efeito do excipiente na célula-tronco ou na cultura de HSPC, e na solubilidade e estabilidade, e na natureza da forma de dosagem.

[00186] Nos casos em que a formulação é aquosa, excipientes como açúcares (incluindo, entre outros, glicose, manitol, sorbitol, *etc.)* sais, carboidratos e agentes tamponantes (de preferência a um pH de 3 a 9), mas, para alguns pedidos, eles podem ser formulados de maneira mais adequada com uma solução não aquosa estéril ou como uma forma seca para ser usada em conjunto com um veículo adequa-

do, como água estéril, livre de pirogênio (WFI).

[00187] Embora a invenção seja descrita em conjunto com as modalidades ilustradas, é entendido que não se destina a limitar a invenção às modalidades. Pelo contrário, a invenção pretende abranger todas as alternativas, modificações e equivalentes, incluindo equivalentes de características específicas, que podem ser incluídas na invenção, conforme definido pelas reivindicações anexas.

[00188] Tanto a descrição geral acima como a descrição detalhada, bem como os exemplos a seguir, são exemplificativos e explicativos, e não são restritivos aos ensinamentos. Os títulos das seções usados neste documento são para fins organizacionais apenas e não devem ser interpretados como limitando o assunto desejado de qualquer modo. No caso de qualquer literatura incorporada por referência contradizer qualquer termo definido nesta especificação, esta especificação controla. Todos os intervalos fornecidos no pedido abrangem os pontos de extremidade, salvo indicação em contrário.

[00189] Deve-se notar que, conforme usado neste pedido, a forma singular "um", "uma" e "o/a" incluem referências plurais, a menos que o contexto indique claramente o contrário. Assim, por exemplo, a referência a "uma composição" inclui uma pluralidade de composições e a referência a "uma célula" inclui uma pluralidade de células e similares. O uso de "ou" é inclusivo e significa "e/ou", salvo indicação em contrário.

[00190] Os intervalos numéricos incluem os números que definem o intervalo. Os valores medidos e mensuráveis são entendidos como aproximados, levando em consideração dígitos significativos e o erro associado à medição. O termo "cerca" ou "aproximadamente" significa um erro aceitável para um valor particular determinado por aquele versado na técnica, o que depende em parte da forma como o valor é medido ou determinado. O uso de um modificador como "cerca de"

antes de um intervalo ou antes de uma lista de valores que modifica cada ponto final do intervalo ou cada valor na lista. Por exemplo, "cerca de 50-55" abrange "cerca de 50 a cerca de 55". Além disso, o uso de "compreendem", "compreende", "compreendendo", "contém", "contendo", "contêm", "incluem", "inclui", "incluindo" não são limitantes.

[00191] A menos que especificamente indicado na especificação acima, modalidades na especificação que recitam "compreendendo" vários componentes também são contemplados como "consistindo em" ou "consistindo essencialmente em" os componentes recitados; modalidades na especificação que recitam "consistindo em" vários componentes também são contemplados como "compreendendo" ou "consistindo essencialmente em" os componentes recitados; modalidades na especificação que recitam "cerca de" vários componentes também são contemplados como "nos" componentes recitados; e modalidades na especificação que recitam "consistindo essencialmente em" vários componentes também são contemplados como "consistindo em" ou "compreendendo" os componentes recitados (essa intercambiabilidade não se aplica ao uso desses termos nas reivindicações).

EXEMPLOS

Exemplo 1 - Métodos

Cultura Celular

[00192] As células da medula óssea CD34+ humanas criopreservadas foram obtidas de AllCells (cat. ABM017F) ou StemCell Technologies (cat. Nº 70008). Após descongelar e lavar duas vezes em 20 ml de StemSpan SFEM (tecnologias de células-tronco, cat. Nº 09650), as células foram cultivadas por 48 horas em StemSpan SFEM (StemCell Technologies, cat. Nº 09650) contendo trombopoietina (TPO, 50ng/ml, StemCell Technologies, Nº de cat. 02922), ligando Flt3 humano (Flt3l, 50ng/ml, StemCell Technologies, Nº de cat. 78137,2), interleucina-6 humana (Il-6, 50ng/mL, StemCell Technologies, Mº de cat. 78148,2),

fator de células-tronco humanas (SCF, 50ng/ml, tecnologias StemCell, código Nº 78155,2) e StemRegenin-1 (SR1, 0,75uM), bem como Penicilina/Estreptomicina (P/S, 100U/ml de penicilina e Estreptomicina 100ug/ml, Life Technologies, Nº de cat. 15140122).

Formulação de nanopartículas lipídicas ("LNP")

[00193] Os LNPs foram formulados dissolvendo componentes de nanopartículas lipídicas em etanol 100% com as seguintes razões molares: 45% molar (12,7 mM) de lipídio de amina (Por exemplo, lipídio A); Lipídio auxiliar 44% molar (12,4 mM) (Por exemplo, colesterol); 9% em mol (2,53 mM) de lipídio neutro (Por exemplo, DSPC); e 2% molar (.563 mM) de lipídio com PEG (Por exemplo, PEG2k-DMG ou PEG2k-C11), exceto conforme especificado abaixo. A razão N/P (mol de amina lipídica para mol de RNA) foi de 4,5. Os números de identificação para formulações de LNP são os seguintes: LNP522, LNP525 (mRNA GFP) e LNP670, LNP926 (guia único B2M, mRNA Cas9) e LNP899 (guia único AAVS1, mRNA Cas9). As cargas de RNA foram dissolvidas em tampão acetato 50 mM, pH 4,5 ou citrato de sódio 25 mM, NaCl 100 mM, pH 5,0, resultando em uma concentração de carga de RNA de aproximadamente 0,45 mg/mL.

[00194] Os LNPs foram formados por mistura microfluídica das soluções de lipídios e RNA usando um instrumento de bancada Precision Nanosystems NanoAssemblrTM, de acordo com o protocolo do fabricante. Uma razão de 2:1 de solvente aquoso para orgânico foi mantida durante a mistura usando taxas de fluxo diferenciais. Após mistura, os LNPs foram coletados e diluídos em solução salina tamponada com fosfato, pH 7,4 (PBS) ou 50 mM tris, pH 7,5 (Tris) (aproximadamente 1: 1) para reduzir o teor de etanol antes do processamento posterior. A troca final de tampão foi concluída por diálise em PBS ou Tris (excesso de 100 vezes o volume da amostra), durante a noite a 4°C sob agitação suave usando um cassete de diálise Slide-a-Lyzer™ G2 10 kDa

(ThermoFisher Scientific). As formulações processadas com Tris foram diluídas 1:1 em tris 100 mM, solução salina 90 mM, sacarose a 5% (p/v), pH 7,5 (2x TSS). Alternativamente, os LNPs foram coletados após a mistura, diluídos em água, mantidos em temperatura ambiente por 1 hora e diluídos uma segunda vez 1:1 com água. A troca final de tampão no TSS foi concluída com as colunas de dessalinização PD-10 (GE). Se necessário, as formulações por qualquer método de processamento foram concentradas por centrifugação com filtros centrífugos Amicon 100 kDa (Millipore). A mistura resultante foi então filtrada usando um filtro estéril de 0,2 µm. O filtrado resultante foi armazenado a 2-8°C se o tampão final fosse PBS ou -80°C se o tampão final fosse TSS.

Transcrição in vitro ("IVT") de mRNA da Nuclease e RNA de guia único (sgRNA)

[00195] O mRNA de Cas9 tampado e poliadenilado foi gerado por transcrição *in vitro* usando um modelo de DNA de plasmídeo linearizado e polimerase de RNA T7. O DNA do plasmídeo contendo um promotor T7 e uma região poli(A/T) de 100 resíduos foi linearizado por incubação a 37°C por 2 horas com XbaI com as seguintes condições: plasmídeo de 200 ng/µL, 2 U/µL de XbaI (NEB), e 1x tampão de reação. O XbaI foi inativado por aquecimento da reação a 65°C por 20 min. O plasmídeo linearizado foi purificado a partir de sais de enzima e tampão usando uma coluna de sílica maxi spin (Epoch Life Sciences) e analisado por gel de agarose para confirmar a linearização. A reação de IVT para gerar mRNA modificado por Cas9 foi incubada a 37°C por 4 horas nas seguintes condições: 50 ng/µL de plasmídeo linearizado; 2 mM cada de GTP, ATP, CTP e N1-metil pseudo-UTP (Trilink); ARCA 10 mM (Trilink); 5 U/µL de RNA polimerase T7 (NEB); 1 U/µL de inibidor da RNase murina (NEB); 0,004 U/µL de *E. coli* pirofosfatase inorgânica (NEB); e 1x tampão de reação. Após as 4 horas de incubação,

TURBO DNase (ThermoFisher) foi adicionada a uma concentração final de 0,01 U/µL e a reação foi incubada por mais 30 minutos para remover o molde de DNA. O mRNA de Cas9 foi purificado usando um método de precipitação de LiCl.

[00196] Para todos os métodos, a concentração do transcrito foi determinada medindo a absorvância da luz a 260 nm (Nanodrop), e o transcrito foi analisado por eletroforese capilar pela Bioanlayzer (Agilent). Os SgRNAs foram sintetizados quimicamente.

Transfecção por LNP de células CD34+ da medula óssea humana

[00197] LNPs contendo mRNA de GFP ou mRNA de Cas9 e guia único direcionado a beta2-microgobulina (B2M) foram adicionados em várias concentrações variando de 50,0ng a 800,0ng a 30.000 células CD34+ humanas da medula óssea em um volume total de 100,0ul. A sequência do sgRNA, que tem como alvo a sequência alvo GGCCACGGAGCGAGACATCT B2M (SEQ ID NO: 75), é a seguinte: mG*mG*mC*CACGGAGCGAGACAUCUGUUUUAGAmGmCmUmAm GmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmA mGmUmCmGmGmUmGmCmU*mU*mU*mU (SEQ ID NO:76). Nesta sequência de ácido nucleico, A, U, G e C denotam adenina, uracil, citosina e guanina, respectivamente; m" indica nucleotídeos 2'-O-metil; e "*" indica ligações fosforotioato.

[00198] Para estudos séricos específicos de espécie (estudos Triple S), os LNPs foram incubados em soro a 6,0% de M. musculus (BioreclamationIVT, cat. N°. MSESRM, lote n° MSE245821), M. fascicularis (BioreclamationIVT, cat. N° CYNSRM, lote n° CYN197451) e H. sapiens (Sigma, agrupado, H4522-20ml, lote n° SLBR7629V; BioreclamationIVT, cat. n° HMSRM, lote n° BRH1278638; BioreclamationIVT, cat. n° HMSRM, lote n° BRH1227947) a 37°C por 5 minutos antes da transfecção celular. A apolipoproteína E3 recombinante humana

(ApoE3, R&D Systems, cat. Nº 4144-AE) foi usada em uma variedade de concentrações (01.ug/ml, 1,0ug/ml, 10,0ug/ml e 50,0ug/ml) no tampão nas mesmas condições de incubação descritas acima.

Leitura por Citometria de Fluxo de Células da Medula Óssea Humanas Transfectadas com LNP CD34+

[00199] As células foram coletadas para coloração de anticorpos 24 horas após a transfecção com LNP-GFP ou 5 dias após a transfecção com LNP-B2M. Após lavagem das células no meio de amostra (PBS + 2% de FBS + 2mM EDTA), as células foram bloqueadas com Human TruStain FcX (Biolegend, cat. Nº 422302) à temperatura ambiente (TA) por 5 min.

[00200] Foram marcadas células com os seguintes anticorpos e marcadores, como mostrado na Tabela 2.

Tabela 2. Anticorpos e marcadores.

Anticorpo	**Fluoróforo**	**Fabricante**	**Número de Catálogo**
hCD34	Violeta brilhante 421™ (BV421)	Biolegend	343610
hCD38	R-fitoeritrina (PE)	Biolegend	356604
hCD90	PE-Cy™7 (fluoróforo em tandem corante PE-cianina)	Biolegend	328124
hCD45RA	Alexa Fluor® 700 (AF700)	Biolegend	304120
hB2M	Aloficocianina (APC)	Biolegend	316312
7-AAD	n/a	Biolegend	420404

[00201] As células foram corridas em um Beckman Coulter CytoflexS e analisadas usando o pacote de software FlowJo.

[00202] A sobrevivência celular foi avaliada com coloração com 7-AAD. Normalização para células vivas com base na intercalação de 7-AAD em regiões de DNA ricas em GC e subsequente detecção em ensaios de citometria de fluxo. A sobrevivência celular foi calculada usando a seguinte fórmula:

$[(\text{Amostra1}_{\text{número de eventos da célula}}/\text{Amostra1}_{\text{número de eventos do grânulo}})^*$ Amos-

tra1$_{\text{Número total de grânulos adicionados}}$]/Média{[(Controle1$_{\text{número de de eventos da célula}}$/Controle1$_{\text{Número do de eventos do grânulo}}$)*Controle1$_{\text{Número total de grânulos adicionados}}$], [(Controle2$_{\text{Número de eventos da célula}}$/ controle1$_{\text{Número de eventos de grânulos}}$)*Controle2$_{\text{Número total de grânulos adicionados}}$],..., ControlN}

[00203] Nesta fórmula, "amostra" é definida como qualquer população de HSPCs CD34+ humanos que, durante o curso do experimento, receberam tratamento com LNPs, mRNA, gRNA ou qualquer combinação dos primeiros e "controle" definidos em qualquer população de HSPCs CD34+ humanos que, durante o curso do experimento, não receberam nenhum tratamento com LNPs, mRNA, gRNA ou qualquer combinação dos primeiros.

Sequenciamento de próxima geração ("NGS") e análise para eficiência de clivagem

[00204] Para determinar quantitativamente a eficiência da edição no local alvo no genoma, o sequenciamento profundo foi utilizado para identificar a presença de inserções e deleções introduzidas pela edição de genes.

[00205] As células foram coletadas no dia 5 após a transfecção e o DNA extraído usando o PureLink Genomic DNA Mini Kit (ThermoFisher Scientific, cat. No. K182002). Iniciadores para locus alvo B2M contendo as sequências adaptadoras Illumina P5 e P7 foram utilizados para amplificar o sítio genômico de interesse em uma reação de PCR padrão.

[00206] Os iniciadores de PCR foram projetados em torno do sítio alvo B2M e a área genômica de interesse foi amplificada. As amostras foram submetidas para a preparação da amostra (Kit de Reagente Illumina MiSeq v2, 300 ciclos, nº de cat. 15033624) e o sequenciamento realizado em um instrumento Illumina MiSeq. A frequência de edição no locus de interesse foi analisada usando um duto sob medida. Em resumo, PCR adicional foi realizado de acordo com os protocolos

do fabricante (Illumina) para adicionar a química necessária para o sequenciamento. Os amplicons foram sequenciados em um instrumento Illumina MiSeq. As leituras foram alinhadas ao genoma de referência humano (Por exemplo, hg38) após a eliminação daqueles com baixos índices de qualidade. Os arquivos resultantes que contêm as leituras foram mapeados para o genoma de referência (arquivos BAM), onde foram selecionadas leituras que se sobrepunham à região de interesse alvo e o número de leituras de tipo selvagem versus o número de leituras que contêm uma inserção, substituição ou deleção foi calculado.

[00207] A percentagem de edição (Por exemplo, "eficiência de edição" ou "percentagem de edição") é fornecida como o número total de leituras de sequência com inserções ou deleções sobre o número total de leituras de sequência, incluindo o tipo selvagem.

Análise de Formulação

[00208] As formulações de LNP são analisadas quanto ao tamanho médio de partícula, polidispersidade(pdi), conteúdo total de RNA e eficiência de encapsulação do RNA. O tamanho médio das partículas e a polidispersidade são medidos por espalhamento dinâmico de luz (DLS) usando um instrumento Malvern Zetasizer DLS. As amostras de LNP são diluídas 30X em PBS antes de serem medidas por DLS. O diâmetro médio Z, que é uma medida baseada na intensidade do tamanho médio das partículas, foi relatado juntamente com o número de diâmetro médio e pdi.

[00209] Um ensaio baseado em fluorescência (Ribogreen®, ThermoFisher Scientific) é usado para determinar a concentração total de RNA e o RNA livre. A eficiência da encapsulação é calculada como (RNA total - RNA livre)/RNA total. As amostras de LNP são diluídas adequadamente com 1x tampão TE contendo 0,2% de Triton-X 100 para determinar o RNA total ou 1x tampão TE para determinar o RNA livre. As curvas padrão são preparadas utilizando a solução de RNA

inicial usada para fazer as formulações e diluídas em 1x tampão TE +/- 0,2% de Triton-X 100.O corante RiboGreen® diluído (100X em 1xTE tampão, de acordo com as instruções do fabricante) é então adicionado a cada um dos padrões e amostras e deixado incubar por 10 minutos à temperatura ambiente, na ausência de luz. Um leitor de microplacas SpectraMax M5 (Molecular Devices) é usado para ler as amostras com excitação, corte automático e comprimentos de onda de emissão definidos para 488 nm, 515 nm e 525 nm, respectivamente. O RNA total e o RNA livre são determinados a partir das curvas padrão apropriadas. A eficiência da encapsulação é calculada como (RNA total - RNA livre)/RNA total. O mesmo procedimento pode ser usado para determinar a eficiência de encapsulamento de um componente de carga baseado em DNA. Para o DNA de fita simples, pode-se usar o Oligreen Dye e, para o DNA de fita dupla, o Picogreen Dye.

Exemplo 2 - Entrega de GFP às células da medula óssea CD34+

[00210] Os LNPs foram formulados conforme descrito no Exemplo 1 com mRNA de GFP em um tampão final de PBS e adicionados a 30.000 células humanas da medula óssea CD34+ em um volume total de 100,0ul, fornecendo 0, 50,0ng, 100,0ng e 200,0ng de mRNA de GFP em várias reações. Antes da administração às células, os LNPs foram pré-incubados com soro a 6% (v/v) de M. musculus (BioreclamationIVT, cat. N°. MSESRM, lote n° MSE245821) a 37°C por 5 minutos. As células foram cultivadas como descrito no Exemplo 1.

[00211] As células GFP+ foram quantificadas 24 horas após a adição de LNP às células CD34+ humanas por citometria de fluxo. A população de células GFP+ foi determinada no canal FITC (excitação máxima 490, emissão máxima 525, linha laser 488) em relação a um controle GFP- (denominado "controle" na Fig. 1). A percentagem de células GFP+ em todas as células vivas da medula óssea CD34+ humana 24 horas após a entrega do mRNA de GFP mediada por LNP é

mostrada na Fig. 1. As composições de LNP são as seguintes, Fig. 1 (A): 45% de lipídio A, 44% de colesterol, 9% de DSPC, 2% de PEG; Fig. 1 (B): 45% de lipídio A, 45% de colesterol, 9% de DSPC, 1% de PEG. Tamanho da amostra biológica n=3.

[00212] As composições de LNP demonstram a entrega dependente da dose de mRNA para células da medula óssea CD34+ *in vitro*.

Exemplo 3 - A pré-incubação de LNPs facilita a entrega

[00213] Os testes demonstram que os LNPs requerem incubação com 6% de soro de camundongo (v/v) antes da transfecção, a fim de fornecer eficientemente mRNA de GFP para células da medula óssea $CD34^+$ humanas. As células foram cultivadas e transfectadas com composições de LNP, como descrito no Exemplo 2, com as seguintes modificações:

[00214] A Fig. 2A mostra a percentagem de células GFP^+ em todas as células vivas de amostras de medula óssea $CD34^+$ humanas com aplicação de LNP no dia 0 imediatamente após o descongelamento de um frasco de células criopreservadas. LNPs com mRNA de 50,0 ng, 100,0 ng ou 200,0 ng de GFP foram adicionados às células com e sem incubação sérica antes da transfecção.

[00215] A Fig. 2B mostra a percentagem de células GFP^+ em todas as células vivas de amostras de medula óssea $CD34^+$ humanas com aplicação de LNP no dia 2 após o descongelamento de um frasco de células criopreservadas. LNPs com 50,0 ng, 100,0 ng ou 200,0 ng de mRNA de GFP foram adicionados às células com e sem incubação sérica antes da transfecção. Tamanho da amostra biológica n=3.

Exemplo 4 - Entrega de RNAs Cas9 e Guia via LNPs; Edição de genes em células CD34+ da medula óssea

[00216] Os LNPs foram formulados como descrito no Exemplo 1 com sgRNA (G529) e mRNA de Cas9 como descrito no Exemplo 1 e com uma razão de peso de 1: 1 em um tampão final de TSS. A com-

posição de LNP era 45% de lipídio A, 44% de colesterol, 9% de DSPC, 2% de PEG com uma razão N/P de 4,5.

[00217] Usando os métodos de entrega de LNP para transfectar células $CD34^+$ da medula óssea humana, um RNAm de Cas9 e sgRNA de B2M foram entregues eficientemente às células usando pré-incubação com percentagens crescentes (v/v) de soro de M. musculus ou M. fascicularis. Os complexos Cas9-sgRNA ativos são entregues por LNPs pré-incubados com vários soros.

[00218] A Fig. 3A mostra a análise FACS de células transfectadas, mostrando a percentagem de células negativas B2M após a aplicação de LNPs (a 400,0 ng de mRNA de Cas9 e sgRNA (1: 1 em peso)) incubados com soro de camundongo ("Camundongo-S") em 6%, 30% e 60% (v/v) ou soro de primata não humano ("Cyno-S") a 6%, 30% e 60% (v/v). Os LNPs sem pré-incubação sérica ("somente LNP") e as células sem tratamento ("Ctrl") servem não mostram uma distribuição eficiente (medindo a queda da expressão de B2M). A pré-incubação com soro de camundongo ou primata facilita a eliminação eficiente da expressão de B2M nas células da medula óssea $CD34^+$.

[00219] A Fig. 3B mostra a frequência de edição no nível genômico conforme determinado por NGS para células da medula óssea $CD34^+$ humanas transfectadas com LNPs (a 400,0 ng de mRNA de Cas9 e sgRNA (1:1 em peso)) incubadas com soro de camundongo a 6%, 30% e 60% (v/v) ou soro de primata não humano a 6%, 30% e 60% (v/v). Como na Fig. 3 (A), a aplicação de LNP sem pré-incubação sérica e as células sem tratamento não apresentam distribuição eficiente (medindo a% de edição). Inserções ("In", cinza claro) e deleções ("Del", preto) são representadas graficamente no eixo Y, mostrando mais de 60%, mais de 70%, mais de 80% e mais de 90% de edição eficiência nas células $CD34^+$. As amostras "apenas LNP" e "Cntrl" não exibem níveis detectáveis de indels no locus B2M. Tamanho da amos-

tra biológica n=3.

Exemplo 5 - Pré-incubação com fator sérico isolado ApoE3

[00220] Para investigar se a etapa de pré-incubação sérica pode ser substituída por proteína recombinante, os LNPs, como descrito no Exemplo 4, entregando Cas9 e o sgRNA B2M foram pré-incubados com Apolipoproteína E3 recombinante humana (ApoE3), soro de camundongo ou não humano soro de primata durante a etapa de incubação do LNP antes da transfecção celular.

[00221] Na Fig. 4A, as percentagens de células B2M negativas após a aplicação de LNPs (a 400 ng de mRNA de Cas9 e sgRNA (1:1)) incubadas com soro de camundongo ("camundongo-S") a 6% (v/v), soro de primata não humano ("cyno-S") a 6% (v/v) ou ApoE3 a 0,1 μg/ml, 1,0 μg/ml, 10,0 μg/ml e 50,0 μg/ml. As células $CD34^+$ da medula óssea humana sem tratamento servem como controle negativo ("Ctrl"). ApoE3 mostra um aumento dependente da dose na entrega para as células CD34+, e pode ser usado na etapa de pré-incubação.

[00222] Da mesma forma, a edição de genes mostra uma resposta dependente da dose ao ApoE3. A Fig. 4B mostra a percentagem de edição do alvo B2M conforme determinado por NGS para células da medula óssea $CD34^+$ humanas transfectadas com 400,0 ng de LNPs incubadas com soro de camundongo a 6% (v/v), soro de primata não humano a 6% (v/v) ou ApoE3 a 0,1 μg/ml, 1,0 μg/ml, 10,0 μg/ml e 50,0 μg/ml. As células da medula óssea $CD34^+$ humanas sem tratamento servem como controle negativo, que não exibe níveis detectáveis de indels no locus B2M. Tamanho da amostra biológica n=3.

Exemplo 6 - Pré-incubação com fatores séricos

[00223] Esta experiência testou a pré-incubação de LNP com uma variedade de apolipoproteínas diferentes, mostrando a captação *in vitro* de LNP com isoformas de ApoE, medidas como nível de knockdown de B2M e frequência de edição em uma população de HSPC.

Antes da transfecção, os LNPs (ID LNP926) foram incubados a 37C por 5 minutos com soro de M. fascicularis a 6% (v/v) ou as seguintes apolipoproteínas em várias concentrações: ApoA-I humana recombinante (Millipore Sigma, cat # SRP4693), ApoB do plasma humano (Millipore Sigma, cat# A5353), ApoC-I do plasma humano (Millipore Sigma, cat# A7785), ApoE2 recombinante humano (Millipore Sigma, cat# SRP4760), ApoE3 recombinante humano (Millipore Sigma, cat# SRP4696), ApoE4 recombinante humano (Millipore Sigma, cat# A3234). Os LNPs foram adicionados às células humanas da medula óssea CD34+ a uma concentração de 200ng de carga total de RNA (razão 1:1 p/p de mRNA de Cas9 e guia único). A expressão de B2M no nível da proteína foi determinada por citometria de fluxo, utilizando os mesmos anticorpos descritos acima no dia 5 após a transfecção. A análise dos dados foi realizada utilizando o pacote de software FlowJo. Os dados representam uma amostra biológica (N=1), média +/- DP de duplicatas técnicas.

[00224] A Fig. 5 mostra o knockdown de B2M em uma população de HSPCs CD34+ após a transfecção de LNPs pré-incubados com ApoE2, ApoE3 e ApoE4. Nenhum tratamento, nenhuma pré-incubação e pré-incubação com ApoA-I, ApoB e ApoC-I não resultaram em knockdown de B2M. Nesta experiência, a pré-incubação de LNP com ApoE2 mostrou menos knockdown de B2M em comparação com as outras duas isoformas de ApoE.

Exemplo 7 - Curso de tempo da exposição ao LNP

[00225] Esta experiência testou a duração da exposição ao LNP quanto ao seu impacto na viabilidade e nas taxas de edição. O LNP899, entregando mRNA de Cas9 e G562 direcionado ao AAVS1, foi pré-incubado a 37C por cerca de 5 minutos com soro de primata não humano a 6% (v/v). Os LNPs foram adicionados às células humanas da medula óssea CD34+ a uma concentração de 300ng de carga

total de RNA (razão 1:1 p/p de mRNA de Cas9 e guia único). Às 2 horas, 6 horas ou 24 horas após a transfecção, as células foram centrifugadas e ressuspensas em meio fresco sem LNP. A viabilidade celular foi avaliada em 3 e 8 dias usando as Contas de Contagem Absoluta CountBright™ (Invitrogen, Cat. C36950) medido em um citômetro de fluxo CytoFLEXS (Beckman Coulter). A edição foi medida por NGS como descrito no Exemplo 1. A Tabela 3 e a Fig. 6B mostram a frequência de edição 8 dias após a transdução. A Tabela 3 e a Fig. 6A mostram a viabilidade celular aos 3 dias e 8 dias após a transdução.

Tabela 3 - Viabilidade e edição em diferentes exposições ao LNP

Exposição ao LNP	Viabilidade 3d (células/ml)	Viabilidade SD 3d	Viabilidade 8d (células/ml)	Viabilidade SD 8d	Freq Indel	Indels SD
Controle Negativo	640.341	31.810	4.453.259	965.751	0,00	0,00
2h	583.279	44.284	3.400.929	525.282	0,60	0,00
6h	515.561	74.690	2.762.987	375.265	0,88	0,01
24h	447.350	56.802	2.435.103	240.977	0,97	0,00

BREVE DESCRIÇÃO DAS SEQUÊNCIAS DIVULGADAS

SEQ ID NO	Descrição
1	Sequência de codificação do DNA de Cas9 usando o análogo da timidina dos códons mínimos de uridina listados na Tabela 1, com códons de iniciação e de parada
2	Sequência de codificação de DNA de Cas9 usando códons com expressão geralmente alta em seres humanos
3	Sequência de aminoácidos de Cas9 com um sinal de localização nuclear (1xNLS) como os 7 aminoácidos C-terminais
4	ORF de mRNA de Cas9 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
5	ORF de mRNA de Cas9 usando códons com expressão geralmente alta em seres humanos, com códons de iniciação e de parada

SEQ ID NO	Descrição
10	Sequência de codificação de mRNA de Cas9 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
13	Sequência de aminoácidos de Cas9 (sem NLS)
14	ORF de mRNA de Cas9 que codifica SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
15	Sequência de codificação de Cas9 codificando a SEQ ID NO: 13 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
16	Sequência de aminoácidos de Cas9 nickase (sem NLS)
17	ORF de mRNA de Cas9 nickase que codifica SEQ ID NO: 16 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
18	Sequência de codificação de Cas9 nickase codificando a SEQ ID NO: 16 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
19	Sequência de aminoácidos de dCas9 (sem NLS)
20	ORF de mRNA de dCas9 que codifica SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
21	Sequência de codificação de dCas9 codificando a SEQ ID NO: 13 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)

SEQ ID NO	Descrição
22	Sequência de aminoácidos de Cas9 com dois sinais de localização nuclear (2xNLS) como aminoácidos C-terminais
23	ORF de mRNA de Cas9 que codifica SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
24	Sequência de codificação de Cas9 codificando a SEQ ID NO: 13 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
25	Sequência de aminoácidos de Cas9 nickase com dois sinais de localização nuclear como aminoácidos C-terminais
26	ORF de mRNA de Cas9 nickase que codifica SEQ ID NO: 16 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
27	Sequência de codificação de Cas9 nickase codificando a SEQ ID NO: 16 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
28	Sequência de aminoácidos de dCas9 com dois sinais de localização nuclear como aminoácidos C-terminais
29	ORF de mRNA de dCas9 que codifica SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
30	Sequência de codificação de dCas9 codificando a SEQ ID NO: 13 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)
31	Promotor T7

SEQ ID NO	Descrição
32	Beta-globina humana 5' UTR
33	Beta-globina humana 3' UTR
34	Alfa-globina humana 5' UTR
35	Alfa-globina humana 3' UTR
36	*Xenopus laevis*beta-globina 5' UTR
37	*Xenopus laevis*beta-globina 3' UTR
38	Hormônio do Crescimento Bovino 5' UTR
39	Hormônio do Crescimento Bovino 3' UTR
40	Mus musculus hemoglobina alfa, cadeia adulta 1 (Hba-a1), 3'UTR
41	HSD17B4 5' UTR
42	RNA guia único G282 direcionado ao gene TTR de camundongo
43	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 4, sequência de Kozak e 3' UTR de ALB
45	ORF Cas9 alternativa com conteúdo U de 19,36%
46	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 45, sequência de Kozak e 3' UTR de ALB
47	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 45 e 3' UTR de ALB
48	Transcrição Cas9 compreendendo Cas9 ORF usando códons com expressão geralmente alta em seres humanos
49	Transcrição Cas9 compreendendo a sequência Kozak com Cas9 ORF usando códons com expressão geralmente alta em seres humanos
50	ORF Cas9 com junções de emenda removidas; 12,75% de conteúdo em U
51	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 50, sequência de Kozak e 3' UTR de ALB
52	ORF Cas9 com códons mínimos de uridina frequentemente usados em seres humanos em geral; 12,75% de conteúdo

SEQ ID NO	Descrição
	em U
53	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 52, sequência de Kozak e 3' UTR de ALB
54	ORF Cas9 com códons mínimos de uridina raramente usados em seres humanos em geral; 12,75% de conteúdo em U
55	Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 54, sequência de Kozak e 3' UTR de ALB
63	sequência poli-A 100
64	RNA guia único G209 direcionado ao gene TTR de camundongo
65	ORF que codifica Neisseria meningitidis Cas9 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de parada
66	ORF que codifica Neisseria meningitidis Cas9 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação das proteínas de fusão)
67	Transcrição compreendendo SEQ ID NO: 65 (codificando Neisseria meningitidis Cas9)
68	Sequência de aminoácidos de Neisseria meningitidis Cas9
69	RNA guia único G390 direcionando o gene TTR de rato
70	RNA guia único G502 direcionando o gene TTR do macaco cynomolgus
71	RNA guia único G509 direcionando o gene TTR do macaco cynomolgus
72	RNA guia único G534 direcionando o gene TTR de rato

[00226] Veja a Tabela de Sequências abaixo para as próprias sequências. As sequências de transcrição geralmente incluem GGG como os três primeiros nucleotídeos para uso com ARCA ou AGG como os três primeiros nucleotídeos para uso com CleanCap™. Por conseguinte, os três primeiros nucleotídeos podem ser modificados para uso

com outras abordagens de capeamento, como a enzima de captação de Vaccinia. Promotores e sequências poli-A não estão incluídas nas sequências de transcrição. Um promotor como um promotor T7 (SEQ ID NO: 31) e uma sequência poli-A como SEQ ID NO: 63 podem ser anexados às sequências de transcrição divulgadas nas extremidades 5' e 3', respectivamente. A maioria das sequências de nucleotídeos é fornecida como DNA, mas pode ser facilmente convertida em RNA alterando Ts para Us.

Tabela de Sequência

[00227] A tabela de sequência a seguir fornece uma lista de sequências divulgadas neste documento. Entende-se que se uma sequência de DNA (compreendendo Ts) for referenciada em relação a um RNA, Ts deverá ser substituído por Nós (que pode ser modificado ou não modificado dependendo do contexto) e vice-versa.

Descrição	Sequência	SEQ ID No.
Sequência de codificação 2 do DNA Cas9	ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGATGGGCAGTCATCA CAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGGAAACACAGACAGACACAGCATC AAGAAGAACCTGATCGGAGCACTGCTGTTCGACAGCGGAGAAACAGCAGAAGCAACAAGACTGA AGAGAACAGCAAGAAGAAGATACACAAGAAGAAAGAACAGAATCTGCTACCTGCAGGAAATCTTC AGCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACAGACTGGAAGAAAGCTTCCTGGTCGA AGAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAGTCGCATACCACG AAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGACAAGGCAGACCTG AGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGACACTTCCTGATCGAAGGAGA CCTGAACCCGGACAACAGCGACGTCGACAAGCTGTTCATCCAGCTGGTCCAGACATACAACCAGC TGTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAAGGCAATCCTGAGCGCAAGACT GAGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCCGGGAGAAAAGAAGAACGGACTG TTCGGAAACCTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAACTTCGACCTGGC AGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGACAACCTGCTGGCA CAGATCGGAGACCAGTACGCAGACCTGTTCCTGGCAGCAAAGAACCTGAGCGACGCAATCCTGC TGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGGCACCGCTGAGCGCAAGCATGATCAAG AGATACGACGAACACCACCAGGACCTGACACTGCTGAAGGCACTGGTCAGACAGCAGCTGCCGG AAAAGTACAAGGAAATCTTCTTCGACCAGAGCAAGAACGGATACGCAGGATACATCGACGGAGGA GCAAGCCAGGAAGAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAAGA ACTGCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTCGACAACGGAAGCA TCCCGCACCAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAAGACAGGAAGACTTCTACCC GTTCCTGAAGGACAACAGAGAAAAGATCGAAAAGATCCTGACATTCAGAATCCCGTACTACGTCG GACCGCTGGCAAGAGGAAACAGCAGATTCGCATGGATGACAAGAAAGAGCGAAGAAACAATCAC ACCGTGGAACTTCGAAGAAGTCGTCGACAAGGGAGCAAGCGCACAGAGCTTCATCGAAAGAATG ACAAACTTCGACAAGAACCTGCCGAACGAAAAGGTCCTGCCGAAGCACAGCCTGCTGTACGAATA CTTCACAGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGAATGAGAAAGCCGGCAT TCCTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGACAAACAGAAAGGTCACA GTCAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCTTCGACAGCGTCGAAATCAGCGG	1

	AGTCGAAGACAGATTCAACGCAAGCCTGGGAACATACCACGACCTGCTGAAGATCATCAAGGACA AGGACTTCCTGGACAACGAAGAAAACGAAGACATCCTGGAAGACATCGTCCTGACACTGACACTG TTCGAAGACAGAGAAATGATCGAAGAAAGACTGAAGACATACGCACACCTGTTCGACGACAAGGT CATGAAGCAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTGAGCAGAAAGCTGATCAAC GGAATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGACGGATTCGCAA ACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCAAGGAAGACATCCAGAAGGCA CAGGTCAGCGGACAGGGAGACAGCCTGCACGAACACATCGCAAACCTGGCAGGAAGCCCGGCA ATCAAGAAGGGAATCCTGCAGACAGTCAAGGTCGTCGACGAACTGGTCAAGGTCATGGGAAGAC ACAAGCCGGAAAACATCGTCATCGAAATGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAG AACAGCAGAGAAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGCCAGATCCTGAA GGAACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTACCTGCAGAACG GAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGCGACTACGACGTCGACCA CATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTCCTGACAAGAAGCGAC AAGAACAGAGGAAAGAGCGACAACGTCCCGAGCGAAGAAGTCGTCAAGAAGATGAAGAACTACT GGAGACAGCTGCTGAACGCAAAGCTGATCACACAGAGAAAGTTCGACAACCTGACAAAGGCAGA GAGAGGAGGACTGAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTCGAAACAAGA CAGATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAAAGTACGACGAAAACGA CAAGCTGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTCAGCGACTTCAGAAAGG ACTTCCAGTTCTACAAGGTCAGAGAAATCAACAACTACCACCACGCACACGACGCATACCTGAAC GCAGTCGTCGGAACAGCACTGATCAAGAAGTACCCGAAGCTGGAAAGCGAATTCGTCTACGGAG ACTACAAGGTCTACGACGTCAGAAAGATGATCGCAAAGAGCGAACAGGAAATCGGAAAGGCAACA GCAAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGACAGAAATCACACTGGCAAACGGA GAAATCAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAAATCGTCTGGGACAAGG GAAGAGACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCAACATCGTCAAGAAGACA GAAGTCCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCGAAGAGAAACAGCGACAAGCTGA TCGCAAGAAAGAAGGACTGGGACCCGAAGAAGTACGGAGGATTCGACAGCCCGACAGTCGCATA CAGCGTCCTGGTCGTCGCAAAGGTCGAAAAGGGAAAGAGCAAGAAGCTGAAGAGCGTCAAGGAA CTGCTGGGAATCACAATCATGGAAAGAAGCAGCTTCGAAAAGAACCCGATCGACTTCCTGGAAGC AAAGGGATACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCTGTTCGAAC	

	TGGAAAACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGAAACGAACTGGC ACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTACGAAAAGCTGAAGGGAAGCC CGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAACAGCACAAGCACTACCTGGACGAAATCATC GAACAGATCAGCGAATTCAGCAAGAGAGTCATCCTGGCAGACGCAAACCTGGACAAGGTCCTGA GCGCATACAACAAGCACAGAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCCACCTGTTC ACACTGACAAACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGACAGAAAGAG ATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGAGCATCACAGGACTGTACG AAACAAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAA GGTCTAG	
Sequência de codificação 1 do DNA Cas9	ATGGATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGGGTTGGGCAGTGATCAC GGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGGAACACCGATAGACACAGCATCA AGAAAAATCTCATCGGAGCCCTGCTGTTTGACTCCGGCGAAACCGCAGAAGCGACCCGGCTCAAA CGTACCGCGAGGCGACGCTACACCCGGCGGAAGAATCGCATCTGCTATCTGCAAGAGATCTTTTC GAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTGGTGGAGG AGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTCGACGAAGTGGCGTACCACGAA AAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAGGCCGACCTCAG ATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCCGCGGACACTTCCTGATCGAAGGCGATCT GAACCCTGATAACTCCGACGTGGATAAGCTTTTCATTCAACTGGTGCAGACCTACAACCAACTGTT CGAAGAAAACCCAATCAATGCTAGCGGCGTCGATGCCAAGGCCATCCTGTCCGCCCGGCTGTCG AAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGGGAGAGAAAAAGAACGGACTTTTCG GCAACTTGATCGCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGG ACGCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGCACAAATT GGCGATCAGTACGCGGATCTGTTCCTTGCCGCTAAGAACCTTTCGGACGCAATCTTGCTGTCCGA TATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAGCGCCTCGATGATTAAGCGGTACG ACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTCGTGAGACAGCAACTGCCTGAAAAGTAC AAGGAGATCTTCTTCGACCAGTCCAAGAATGGGTACGCAGGGTACATCGATGGAGGCGCTAGCC AGGAAGAGTTCTATAAGTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGAAGAACTGCTG GTCAAGCTGAACAGGGAGGATCTGCTCCGGAAACAGAGAACCTTTGACAACGGATCCATTCCCCA CCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAGGACTTTTACCCATTCCTCA	2

	AGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCCGTATTACGTGGGCCCACTG GCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAAAATCAGAGGAAACCATCACTCCTTGGAA TTTCGAGGAAGTTGTGGATAAGGGAGCTTCGGCACAAAGCTTCATCGAACGAATGACCAACTTCG ACAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCT ACAACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGA GAACAGAAGAAAGCAATTGTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCGTCAAGCAGCT TAAAGAGGACTACTTCAAGAAGATCGAGTGTTTCGACTCAGTGGAAATCAGCGGGGTGGAGGACA GATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGAAGATCATCAAGGACAAGGACTTCCTTG ACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGACCCTTTTCGAGGATCGC GAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAACT CAAGCGCCGCCGGTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGAT AAACAGAGCGGTAAAACTATCCTGGATTTCCTCAAATCGGATGGCTTCGCTAATCGTAACTTCATG CAATTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAAAAAGCACAAGTGTCCGGACA GGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTCGCCGGCGATTAAGAAGGGAATT CTGCAAACTGTGAAGGTGGTCGACGAGCTGGTGAAGGTCATGGGACGGCACAAACCGGAGAATA TCGTGATTGAAATGGCCCGAGAAAACCAGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAG GATGAAGCGGATCGAAGAAGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTG GAAAACACGCAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGACATGTA CGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACCACATCGTTCCACAGT CCTTTCTGAAGGATGACTCGATCGATAACAAGGTGTTGACTCGCAGCGACAAGAACAGAGGGAAG TCAGATAATGTGCCATCGGAGGAGGTCGTGAAGAAGATGAAGAATTACTGGCGGCAGCTCCTGAA TGCGAAGCTGATTACCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAG AGCTGGATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACGT GGCGCAGATCTTGGACTCCCGCATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAAG TGAAGGTGATTACCCTGAAAAGCAAACTTGTGTCGGACTTTCGGAAGGACTTTCAGTTTTACAAAG TGAGAGAAATCAACAACTACCATCACGCGCATGACGCATACCTCAACGCTGTGGTCGGTACCGCC CTGATCAAAAAGTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTACAAGGTCTACGACGTG AGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGAAATACTTCTTTTACTCA AACATCATGAACTTTTTCAAGACTGAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTG	

	ATCGAAACTAACGGAGAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTC GCAAAGTGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGATTT TCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAAGAAAGACTGGGA CCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTCCTCGTGGTGGCCAAG GTGGAGAAGGGAAAGAGCAAAAAGCTCAAATCCGTCAAAGAGCTGCTGGGGATTACCATCATGGA ACGATCCTCGTTCGAGAAGAACCCGATTGATTTCCTCGAGGCGAAGGGTTACAAGGAGGTGAAGA AGGATCTGATCATCAAACTCCCCAAGTACTCACTGTTCGAACTGGAAAATGGTCGGAAGCGCATG CTGGCTTCGGCCGGAGAACTCCAAAAAGGAAATGAGCTGGCCTTGCCTAGCAAGTACGTCAACTT CCTCTATCTTGCTTCGCACTACGAAAAACTCAAAGGGTCACCGGAAGATAACGAACAGAAGCAGC TTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAATCTCCGAGTTTTCAAAGC GCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGTCGGCCTACAATAAGCATAGAGATAAG CCGATCAGAGAACAGGCCGAGAACATTATCCACTTGTTCACCCTGACTAACCTGGGAGCCCCAGC CGCCTTCAAGTACTTCGATACTACTATCGATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGA CGCGACCCTGATCCACCAAAGCATCACTGGACTCTACGAAACTAGGATCGATCTGTCGCAGCTGG GTGGCGATGGCGGTGGATCTCCGAAAAAGAAGAGAAAGGTGTAATGA	
Sequência de aminoácidos Cas9	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG	3

	LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV	
Quadro de leitura aberta (ORF) de mRNA Cas9 2	AUGGACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU	4

	CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC	

	GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAG GAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAG	
ORF de mRNA de Cas9 1	AUGGAUAAGAAGUACUCAAUCGGGCUGGAUAUCGGAACUAAUUCCGUGGGUUGGGCAGUGAUC ACGGAUGAAUACAAAGUGCCGUCCAAGAAGUUCAAGGUCCUGGGGAACACCGAUAGACACAGC AUCAAGAAAAAUCUCAUCGGAGCCCUGCUGUUUGACUCCGGCGAAACCGCAGAAGCGACCCGG CUCAAACGUACCGCGAGGCGACGCUACACCCGGCGGAAGAAUCGCAUCUGCUAUCUGCAAGAG AUCUUUUCGAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACCGCCUGGAAGAAUCUUUC CUGGUGGAGGAGGACAAGAAGCAUGAACGGCAUCCUAUCUUUGGAAACAUCGUCGACGAAGUG GCGUACCACGAAAAGUACCCGACCAUCUACCAUCUGCGGAAGAAGUUGGUUGACUCAACUGAC AAGGCCGACCUCAGAUUGAUCUACUUGGCCCUCGCCCAUAUGAUCAAAUUCCGCGGACACUUC CUGAUCGAAGGCGAUCUGAACCCUGAUAACUCCGACGUGGAUAAGCUUUUCAUUCAACUGGUG CAGACCUACAACCAACUGUUCGAAGAAAACCCAAUCAAUGCUAGCGGCGUCGAUGCCAAGGCC AUCCUGUCCGCCCGGCUGUCGAAGUCGCGGCGCCUCGAAAACCUGAUCGCACAGCUGCCGGG AGAGAAAAAGAACGGACUUUUCGGCAACUUGAUCGCUCUCUCACUGGGACUCACUCCCAAUUU	5

	CAAGUCCAAUUUUGACCUGGCCGAGGACGCGAAGCUGCAACUCUCAAAGGACACCUACGACGA CGACUUGGACAAUUUGCUGGCACAAAUUGGCGAUCAGUACGCGGAUCUGUUCCUUGCCGCUAA GAACCUUUCGGACGCAAUCUUGCUGUCCGAUAUCCUGCGCGUGAACACCGAAAUAACCAAAGC GCCGCUUAGCGCCUCGAUGAUUAAGCGGUACGACGAGCAUCACCAGGAUCUCACGCUGCUCAA AGCGCUCGUGAGACAGCAACUGCCUGAAAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAA UGGGUACGCAGGGUACAUCGAUGGAGGCGCUAGCCAGGAAGAGUUCUAUAAGUUCAUCAAGCC AAUCCUGGAAAAGAUGGACGGAACCGAAGAACUGCUGGUCAAGCUGAACAGGGAGGAUCUGCU CCGGAAACAGAGAACCUUUGACAACGGAUCCAUUCCCCACCAGAUCCAUCUGGGUGAGCUGCA CGCCAUCUUGCGGCGCCAGGAGGACUUUUACCCAUUCCUCAAGGACAACCGGGAAAAGAUCGA GAAAAUUCUGACGUUCCGCAUCCCGUAUUACGUGGGCCCACUGGCGCGCGGCAAUUCGCGCU UCGCGUGGAUGACUAGAAAAUCAGAGGAAACCAUCACUCCUUGGAAUUUCGAGGAAGUUGUGG AUAAGGGAGCUUCGGCACAAAGCUUCAUCGAACGAAUGACCAACUUCGACAAGAAUCUCCCAAA CGAGAAGGUGCUUCCUAAGCACAGCCUCCUUUACGAAUACUUCACUGUCUACAACGAACUGAC UAAAGUGAAAUACGUUACUGAAGGAAUGAGGAAGCCGGCCUUUCUGUCCGGAGAACAGAAGAA AGCAAUUGUCGAUCUGCUGUUCAAGACCAACCGCAAGGUGACCGUCAAGCAGCUUAAAGAGGA CUACUUCAAGAAGAUCGAGUGUUUCGACUCAGUGGAAAUCAGCGGGGUGGAGGACAGAUUCAA CGCUUCGCUGGGAACCUAUCAUGAUCUCCUGAAGAUCAUCAAGGACAAGGACUUCCUUGACAA CGAGGAGAACGAGGACAUCCUGGAAGAUAUCGUCCUGACCUUGACCCUUUUCGAGGAUCGCGA GAUGAUCGAGGAGAGGCUUAAGACCUACGCUCAUCUCUUCGACGAUAAGGUCAUGAAACAACU CAAGCGCCGCCGGUACACUGGUUGGGGCCGCCUCUCCCGCAAGCUGAUCAACGGUAUUCGCG AUAAACAGAGCGGUAAAACUAUCCUGGAUUUCCUCAAAUCGGAUGGCUUCGCUAAUCGUAACU UCAUGCAAUUGAUCCACGACGACAGCCUGACCUUUAAGGAGGACAUCCAAAAAGCACAAGUGU CCGGACAGGGAGACUCACUCCAUGAACACAUCGCGAAUCUGGCCGGUUCGCCGGCGAUUAAGA AGGGAAUUCUGCAAACUGUGAAGGUGGUCGACGAGCUGGUGAAGGUCAUGGGACGGCACAAA CCGGAGAAUAUCGUGAUUGAAAUGGCCCGAGAAAACCAGACUACCCAGAAGGGCCAGAAAAAC UCCCGCGAAAGGAUGAAGCGGAUCGAAGAAGGAAUCAAGGAGCUGGGCAGCCAGAUCCUGAAA GAGCACCCGGUGGAAAACACGCAGCUGCAGAACGAGAAGCUCUACCUGUACUAUUUGCAAAAU GGACGGGACAUGUACGUGGACCAAGAGCUGGACAUCAAUCGGUUGUCUGAUUACGACGUGGA CCACAUCGUUCCACAGUCCUUUCUGAAGGAUGACUCGAUCGAUAACAAGGUGUUGACUCGCAG	

	CGACAAGAACAGAGGGAAGUCAGAUAAUGUGCCAUCGGAGGAGGUCGUGAAGAAGAUGAAGAA UUACUGGCGGCAGCUCCUGAAUGCGAAGCUGAUUACCCAGAGAAAGUUUGACAAUCUCACUAA AGCCGAGCGCGGCGGACUCUCAGAGCUGGAUAAGGCUGGAUUCAUCAAACGGCAGCUGGUCG AGACUCGGCAGAUUACCAAGCACGUGGCGCAGAUCUUGGACUCCCGCAUGAACACUAAAUACG ACGAGAACGAUAAGCUCAUCCGGGAAGUGAAGGUGAUUACCCUGAAAAGCAAACUUGUGUCGG ACUUUCGGAAGGACUUUCAGUUUUACAAAGUGAGAGAAAUCAACAACUACCAUCACGCGCAUGA CGCAUACCUCAACGCUGUGGUCGGUACCGCCCUGAUCAAAAAGUACCCUAAACUUGAAUCGGA GUUUGUGUACGGAGACUACAAGGUCUACGACGUGAGGAAGAUGAUAGCCAAGUCCGAACAGGA AAUCGGGAAAGCAACUGCGAAAUACUUCUUUUACUCAAACAUCAUGAACUUUUUCAAGACUGAA AUUACGCUGGCCAAUGGAGAAAUCAGGAAGAGGCCACUGAUCGAAACUAACGGAGAAACGGGC GAAAUCGUGUGGGACAAGGGCAGGGACUUCGCAACUGUUCGCAAAGUGCUCUCUAUGCCGCAA GUCAAUAUUGUGAAGAAAACCGAAGUGCAAACCGGCGGAUUUUCAAAGGAAUCGAUCCUCCCA AAGAGAAAUAGCGACAAGCUCAUUGCACGCAAGAAAGACUGGGACCCGAAGAAGUACGGAGGA UUCGAUUCGCCGACUGUCGCAUACUCCGUCCUCGUGGUGGCCAAGGUGGAGAAGGGAAAGAG CAAAAAGCUCAAAUCCGUCAAAGAGCUGCUGGGGAUUACCAUCAUGGAACGAUCCUCGUUCGA GAAGAACCCGAUUGAUUUCCUCGAGGCGAAGGGUUACAAGGAGGUGAAGAAGGAUCUGAUCAU CAAACUCCCCAAGUACUCACUGUUCGAACUGGAAAAUGGUCGGAAGCGCAUGCUGGCUUCGGC CGGAGAACUCCAAAAAGGAAAUGAGCUGGCCUUGCCUAGCAAGUACGUCAACUUCCUCUAUCU UGCUUCGCACUACGAAAAACUCAAAGGGUCACCGGAAGAUAACGAACAGAAGCAGCUUUUCGU GGAGCAGCACAAGCAUUAUCUGGAUGAAAUCAUCGAACAAAUCUCCGAGUUUUCAAAGCGCGU GAUCCUCGCCGACGCCAACCUCGACAAAGUCCUGUCGGCCUACAAUAAGCAUAGAGAUAAGCC GAUCAGAGAACAGGCCGAGAACAUUAUCCACUUGUUCACCCUGACUAACCUGGGAGCCCCAGC CGCCUUCAAGUACUUCGAUACUACUAUCGAUCGCAAAAGAUACACGUCCACCAAGGAAGUUCU GGACGCGACCCUGAUCCACCAAAGCAUCACUGGACUCUACGAAACUAGGAUCGAUCUGUCGCA GCUGGGUGGCGAUGGCGGUGGAUCUCCGAAAAAGAAGAGAAAGGUGUAAUGA	
Sequência de codificação bare Cas9	GACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG	10

	AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC	

	GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC	

	CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAG GAAGCCCGAAGAAGAAGAGAAAGGUC	
Sequência de aminoácidos de Cas9 (sem NLS)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	**13**
ORF de mRNA Cas9 que codifica	AUGGACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC	**14**

SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e parada	ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC	

	GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA	

	AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACUAG	
Sequência de codificação Cas9 que codifica SEQ ID NO: 13 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação da proteína de fusão)	GACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA	**15**

	UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG	

	GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC	
Sequência de aminoácidos da Cas9 nickase (sem NLS)	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI	**16**

	NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
ORF de mRNA de nickase Cas9 que codifica SEQ ID NO: 16 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e parada	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG	**17**

	GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC	

	AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACUAG	
Sequência de codificação de nickase Cas9 que codifica SEQ ID NO: 16 usando códons de uridina mínimos, conforme listado na Tabela 1 (sem códons de iniciação ou de parada; adequado para inclusão na sequência de codificação de proteínas de fusão)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG	**18**

	AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC	

	AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC	
Sequência de aminoácidos de dCas9 (sem NLS)	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV	**19**

	KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD AIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
ORF de mRNA dCas9 que codifica SEQ ID NO: 19 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e parada	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC	**20**

	UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACGCAAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG	

	CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACUAG	
sequência de codificação dCas9 que codifica SEQ ID NO: 19 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação de proteínas de fusão)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA	**21**

	UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG	

	AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACGCAAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAG GAAGC	
Sequência de aminoácidos de	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR	**22**

Cas9 com dois sinais de localização nuclear como aminoácidos C-terminal	RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD GSGSPKKKRKVDGSPKKKRKVDSG	
ORF de mRNA Cas9 que codifica SEQ ID NO: 22 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e parada	AUGGACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG	**23**

	ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG	

	UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAA GCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUCGACA GCGGAUAG	

Sequência de codificação Cas9 que codifica SEQ ID NO: 23 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação da proteína de fusão)	GACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU	**24**

	GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC	

	CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAAGCG GAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUCGACAGCG GA	
Sequência de aminoácidos de Cas9 nickase com dois sinais de localização nuclear como aminoácidos C-terminal	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ	**25**

	LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGSGSPKKKRKVDGSPKKKRKVDSG	
ORF de mRNA de nickase Cas9 que codifica SEQ ID NO: 25 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e de fim	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC	**26**

	GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC	

	GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAA GCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUCGACA GCGGAUAG	
Sequência de codificação de nickase Cas9 que codifica SEQ ID NO: 25 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação de proteínas de fusão)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC	**27**

	GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC	

	CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC GGAAGCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUC GACAGCGGA	
Sequência de aminoácidos de dCas9 com dois sinais de localização nuclear como aminoácidos C-terminal	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV	**28**

	KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD AIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD TTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGSGSPKKKRKVDGSPKKKRKVDSG	
ORF de mRNA de dCas9 que codifica SEQ ID NO: 28 usando códons mínimos de uridina, conforme listado na Tabela 1, com códons de iniciação e parada	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUC ACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCA UCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGAC UGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAU CUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCU GGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGC AUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAG GCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUG AUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAG ACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCC UGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAA AGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGA GCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACC UGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGC UGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC	**29**

	UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCC UGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAA AGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAU CCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUG GAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAG GUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUC AAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUC GUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUC AAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGC CUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAA AACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUC GAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGA AGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAG AGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAG CUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAAC AUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAA GAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGG UCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAU GUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACGCAAUCGUCCC GCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAG AGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACA GCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGG AGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAU CACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAG	

	CUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGAC UUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAG ACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAAC AGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAAC GGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAG AAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGC GUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGAC UUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUAC AGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCA AACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCA GAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCC ACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC GGAAGCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUC GACAGCGGAUAG	
Sequência de codificação dCas9 que codifica SEQ ID NO: 28 usando códons mínimos de uridina, conforme listado na Tabela 1 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação de proteí-	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACA GACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUC AAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGG UCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUA CCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCA	**30**

nas de fusão)	GACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACA UACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUG AGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAG AAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUG GACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUG AGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUG AGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUAC GCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAG CAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUC CUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCC UGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGA UGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAG CAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGU CCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAA GUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGU CGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAA GAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCU GGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAAC GAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGA AGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGC GGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGA GACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUG CAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUC	

	GUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA UGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCG AAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUA CGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACGCAAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGG AAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCU GCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGG ACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCAC AAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUG AUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAG UCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACU ACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAG CAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAG GGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAG ACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUC GCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUC CUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGC CUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAG CUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUAC CUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAAC CUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAA AACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGAC ACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACC AGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC	

	GGAAGCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUC GACAGCGGA	
Promotor T7	TAATACGACTCACTATA	**31**
Beta-globina humana 5' UTR	ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC	**32**
Beta-globina humana 3' UTR	GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTG GGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGC	**33**
Alfa-globina humana 5' UTR	CATAAACCCTGGCGCGCTCGCGGCCCGGCACTCTTCTGGTCCCCACAGACTCAGAGAGAACCCA CC	**34**
Alfa-globina humana 3' UTR	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCCCTTCC TGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC	**35**
Beta-globina Xenopus laevis 5' UTR	AAGCTCAGAATAAACGCTCAACTTTGGCC	**36**
Beta-globina Xenopus laevis 3' UTR	ACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAAT GTTGTCCCCCAAAATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACATTCT	**37**
Hormônio de Crescimento Bovino 5' UTR	CAGGGTCCTGTGGACAGCTCACCAGCT	**38**
Hormônio de Crescimento Bovino 3' UTR	TTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCA CTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCA	**39**
Hemoglobina alfa Mus musculus, cadeia adulta 1 (Hba-a1), 3'UTR	GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTCTTCTCTCCCTTGCACCTGTACCTCTT GGTCTTTGAATAAAGCCTGAGTAGGAAG	**40**
HSD17B4 5' UTR	TCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC	**41**
RNA guia G282 visando o TTR	mU*mU*mA*CAGCCACGUCUACAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**42**
Transcrição Cas9 com 5' UTR de	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC	**43**

HSD, ORF correspondente à SEQ ID NO: 4, sequência de Kozak e 3' UTR de ALB	GGATCCGCCACCATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGAT GGGCAGTCATCACAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGGAAACACAGA CAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGCTGTTCGACAGCGGAGAAACAGCAGAA GCAACAAGACTGAAGAGAACAGCAAGAAGAAGATACACAAGAAGAAAGAACAGAATCTGCTACCT GCAGGAAATCTTCAGCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACAGACTGGAAGAAA GCTTCCTGGTCGAAGAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAA GTCGCATACCACGAAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGA CAAGGCAGACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGACACTTCC TGATCGAAGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTTCATCCAGCTGGTCCA GACATACAACCAGCTGTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAAGGCAATCC TGAGCGCAAGACTGAGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCCGGGAGAAAA GAAGAACGGACTGTTCGGAAACCTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGC AACTTCGACCTGGCAGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGG ACAACCTGCTGGCACAGATCGGAGACCAGTACGCAGACCTGTTCCTGGCAGCAAAGAACCTGAG CGACGCAATCCTGCTGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGGCACCGCTGAGC GCAAGCATGATCAAGAGATACGACGAACACCACCAGGACCTGACACTGCTGAAGGCACTGGTCA GACAGCAGCTGCCGGAAAAGTACAAGGAAATCTTCTTCGACCAGAGCAAGAACGGATACGCAGG ATACATCGACGGAGGAGCAAGCCAGGAAGAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGA TGGACGGAACAGAAGAACTGCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAAC ATTCGACAACGGAAGCATCCCGCACCAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAAGA CAGGAAGACTTCTACCCGTTCCTGAAGGACAACAGAGAAAAGATCGAAAAGATCCTGACATTCAG AATCCCGTACTACGTCGGACCGCTGGCAAGAGGAAACAGCAGATTCGCATGGATGACAAGAAAGA GCGAAGAAACAATCACACCGTGGAACTTCGAAGAAGTCGTCGACAAGGGAGCAAGCGCACAGAG CTTCATCGAAAGAATGACAAACTTCGACAAGAACCTGCCGAACGAAAAGGTCCTGCCGAAGCACA GCCTGCTGTACGAATACTTCACAGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGA ATGAGAAAGCCGGCATTCCTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGA CAAACAGAAAGGTCACAGTCAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCTTCGAC AGCGTCGAAATCAGCGGAGTCGAAGACAGATTCAACGCAAGCCTGGGAACATACCACGACCTGC TGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAAGAAAACGAAGACATCCTGGAAGACATC	

	GTCCTGACACTGACACTGTTCGAAGACAGAGAAATGATCGAAGAAAGACTGAAGACATACGCACA CCTGTTCGACGACAAGGTCATGAAGCAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTG AGCAGAAAGCTGATCAACGGAATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAA GAGCGACGGATTCGCAAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCAAGG AAGACATCCAGAAGGCACAGGTCAGCGGACAGGGAGACAGCCTGCACGAACACATCGCAAACCT GGCAGGAAGCCCGGCAATCAAGAAGGGAATCCTGCAGACAGTCAAGGTCGTCGACGAACTGGTC AAGGTCATGGGAAGACACAAGCCGGAAAACATCGTCATCGAAATGGCAAGAGAAAACCAGACAAC ACAGAAGGGACAGAAGAACAGCAGAGAAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTG GGAAGCCAGATCCTGAAGGAACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACC TGTACTACCTGCAGAACGGAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGC GACTACGACGTCGACCACATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGG TCCTGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGTCCCGAGCGAAGAAGTCGTCAA GAAGATGAAGAACTACTGGAGACAGCTGCTGAACGCAAAGCTGATCACACAGAGAAAGTTCGACA ACCTGACAAAGGCAGAGAGAGGAGGACTGAGCGAACTGGACAAGGCAGGATTCATCAAGAGACA GCTGGTCGAAACAAGACAGATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAA AGTACGACGAAAACGACAAGCTGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTC AGCGACTTCAGAAAGGACTTCCAGTTCTACAAGGTCAGAGAAATCAACAACTACCACCACGCACA CGACGCATACCTGAACGCAGTCGTCGGAACAGCACTGATCAAGAAGTACCCGAAGCTGGAAAGC GAATTCGTCTACGGAGACTACAAGGTCTACGACGTCAGAAAGATGATCGCAAAGAGCGAACAGGA AATCGGAAAGGCAACAGCAAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGACAGAAAT CACACTGGCAAACGGAGAAATCAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAA ATCGTCTGGGACAAGGGAAGAGACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCA ACATCGTCAAGAAGACAGAAGTCCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCGAAGAG AAACAGCGACAAGCTGATCGCAAGAAAGAAGGACTGGGACCCGAAGAAGTACGGAGGATTCGAC AGCCCGACAGTCGCATACAGCGTCCTGGTCGTCGCAAAGGTCGAAAAGGGAAAGAGCAAGAAGC TGAAGAGCGTCAAGGAACTGCTGGGAATCACAATCATGGAAAGAAGCAGCTTCGAAAAGAACCCG ATCGACTTCCTGGAAGCAAAGGGATACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAA GTACAGCCTGTTCGAACTGGAAAACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAG AAGGGAAACGAACTGGCACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTACGA	

	AAAGCTGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAACAGCACAAGCAC TACCTGGACGAAATCATCGAACAGATCAGCGAATTCAGCAAGAGAGTCATCCTGGCAGACGCAAA CCTGGACAAGGTCCTGAGCGCATACAACAAGCACAGAGACAAGCCGATCAGAGAACAGGCAGAA AACATCATCCACCTGTTCACACTGACAAACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACAC AACAATCGACAGAAAGAGATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGA GCATCACAGGACTGTACGAAACAAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAG CCCGAAGAAGAAGAGAAAGGTCTAGCTAGCCATCACATTTAAAAGCATCTCAGCCTACCATGAGA ATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCATCTCTTTTTCTTTTTCGTTGGTGTAAAGCCA ACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAA AAAATGGAAAGAACCTCGAG	
ORF Cas9 alternativa com conteúdo 19,36% U	ATGGATAAGAAGTACTCGATCGGGCTGGATATCGGAACTAATTCCGTGGGTTGGGCAGTGATCAC GGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGGAACACCGATAGACACAGCATCA AGAAGAATCTCATCGGAGCCCTGCTGTTTGACTCCGGCGAAACCGCAGAAGCGACCCGGCTCAA ACGTACCGCGAGGCGACGCTACACCCGGCGGAAGAATCGCATCTGCTATCTGCAAGAAATCTTTT CGAACGAAATGGCAAAGGTGGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTGGTGGAG GAGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTGGACGAAGTGGCGTACCACGA AAAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAGGCCGACCTCA GATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCCGCGGACACTTCCTGATCGAAGGCGAT CTGAACCCTGATAACTCCGACGTGGATAAGCTGTTCATTCAACTGGTGCAGACCTACAACCAACTG TTCGAAGAAAACCCAATCAATGCCAGCGGCGTCGATGCCAAGGCCATCCTGTCCGCCCGGCTGT CGAAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGGGAGAGAAGAAGAACGGACTTTT CGGCAACTTGATCGCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGA GGACGCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGCACAAA TTGGCGATCAGTACGCGGATCTGTTCCTTGCCGCTAAGAACCTTTCGGACGCAATCTTGCTGTCC GATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAGCGCCTCGATGATTAAGCGGTA CGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTCGTGAGACAGCAACTGCCTGAAAAGT ACAAGGAGATTTTCTTCGACCAGTCCAAGAATGGGTACGCAGGGTACATCGATGGAGGCGCCAG CCAGGAAGAGTTCTATAAGTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGAAGAACTGC	**45**

	TGGTCAAGCTGAACAGGGAGGATCTGCTCCGCAAACAGAGAACCTTTGACAACGGAAGCATTCCA CACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAGGACTTTTACCCATTCCT CAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCCGTATTACGTGGGCCCAC TGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAAAATCAGAGGAAACCATCACTCCTTGG AATTTCGAGGAAGTTGTGGATAAGGGAGCTTCGGCACAATCCTTCATCGAACGAATGACCAACTTC GACAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTACGAATACTTCACTGTC TACAACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGAGCGG AGAACAGAAGAAAGCGATTGTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCGTCAAGCAGC TTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTCGACTCAGTGGAAATCAGCGGAGTGGAGGAC AGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGAAGATCATCAAGGACAAGGACTTCCTT GACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGACCCTTTTCGAGGATCG CGAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAAC TCAAGCGCCGCCGGTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGA TAAACAGAGCGGTAAAACTATCCTGGATTTCCTCAAATCGGATGGCTTCGCTAATCGTAACTTCAT GCAGTTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAGAAAGCACAAGTGAGCGGA CAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTCGCCGGCGATTAAGAAGGGAA TCCTGCAAACTGTGAAGGTGGTGGACGAGCTGGTGAAGGTCATGGGACGGCACAAACCGGAGAA TATCGTGATTGAAATGGCCCGAGAAAACCAGACTACCCAGAAGGGCCAGAAGAACTCCCGCGAAA GGATGAAGCGGATCGAAGAAGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGT GGAAAACACGCAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGACATGT ACGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACCACATCGTTCCACAG TCCTTTCTGAAGGATGACTCCATCGATAACAAGGTGTTGACTCGCAGCGACAAGAACAGAGGGAA GTCAGATAATGTGCCATCGGAGGAGGTCGTGAAGAAGATGAAGAATTACTGGCGGCAGCTCCTGA ATGCGAAGCTGATTACCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCA GAGCTGGATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACG TGGCGCAGATCCTGGACTCCCGCATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAA GTGAAGGTGATTACCCTGAAAAGCAAACTTGTGTCGGACTTTCGGAAGGACTTTCAGTTTTACAAA GTGAGAGAAATCAACAACTACCATCACGCGCATGACGCATACCTCAACGCTGTGGTCGGCACCGC CCTGATCAAGAAGTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTACAAGGTCTACGACGT	

	GAGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGAAATACTTCTTTTACT CAAACATCATGAACTTCTTCAAGACTGAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCA CTGATCGAAACTAACGGAGAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTG TTCGCAAAGTGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGA TTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAAGAAAGACTG GGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTCCTCGTGGTGGCC AAGGTGGAGAAGGGAAAGAGCAAGAAGCTCAAATCCGTCAAAGAGCTGCTGGGGATTACCATCAT GGAACGATCCTCGTTCGAGAAGAACCCGATTGATTTCCTGGAGGCGAAGGGTTACAAGGAGGTG AAGAAGGATCTGATCATCAAACTGCCCAAGTACTCACTGTTCGAACTGGAAAATGGTCGGAAGCG CATGCTGGCTTCGGCCGGAGAACTCCAGAAAGGAAATGAGCTGGCCTTGCCTAGCAAGTACGTC AACTTCCTCTATCTTGCTTCGCACTACGAGAAACTCAAAGGGTCACCGGAAGATAACGAACAGAAG CAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAATCTCCGAGTTTTCA AAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGTCGGCCTACAATAAGCATAGAGA TAAGCCGATCAGAGAACAGGCCGAGAACATTATCCACTTGTTCACCCTGACTAACCTGGGAGCTC CAGCCGCCTTCAAGTACTTCGATACTACTATCGACCGCAAAAGATACACGTCCACCAAGGAAGTTC TGGACGCGACCCTGATCCACCAAAGCATCACTGGACTCTACGAAACTAGGATCGATCTGTCGCAG CTGGGTGGCGATGGTGGCGGTGGATCCTACCCATACGACGTGCCTGACTACGCCTCCGGAGGTG GTGGCCCCAAGAAGAAACGGAAGGTGTGATAG	
Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 45, sequência de Kozak e 3' UTR de ALB	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCTGCCACCATGGATAAGAAGTACTCGATCGGGCTGGATATCGGAACTAATTCCGTGGGTTG GGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGGAACACCGATA GACACAGCATCAAGAAGAATCTCATCGGAGCCCTGCTGTTTGACTCCGGCGAAACCGCAGAAGC GACCCGGCTCAAACGTACCGCGAGGCGACGCTACACCCGGCGGAAGAATCGCATCTGCTATCTG CAAGAAATCTTTTCGAACGAAATGGCAAAGGTGGACGACAGCTTCTTCCACCGCCTGGAAGAATC TTTCCTGGTGGAGGAGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTGGACGAAG TGGCGTACCACGAAAAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGAC AAGGCCGACCTCAGATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCCGCGGACACTTCCT GATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGGATAAGCTGTTCATTCAACTGGTGCAGA	**46**

	CCTACAACCAACTGTTCGAAGAAAACCCAATCAATGCCAGCGGCGTCGATGCCAAGGCCATCCTG TCCGCCCGGCTGTCGAAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGGGAGAGAAG AAGAACGGACTTTTCGGCAACTTGATCGCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAAT TTTGACCTGGCCGAGGACGCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACA ATTTGCTGGCACAAATTGGCGATCAGTACGCGGATCTGTTCCTTGCCGCTAAGAACCTTTCGGAC GCAATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAGCGCCTC GATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTCGTGAGACAG CAACTGCCTGAAAAGTACAAGGAGATTTTCTTCGACCAGTCCAAGAATGGGTACGCAGGGTACAT CGATGGAGGCGCCAGCCAGGAAGAGTTCTATAAGTTCATCAAGCCAATCCTGGAAAAGATGGACG GAACCGAAGAACTGCTGGTCAAGCTGAACAGGGAGGATCTGCTCCGCAAACAGAGAACCTTTGAC AACGGAAGCATTCCACACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAGG ACTTTTACCCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCCGT ATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAAAATCAGAGGA AACCATCACTCCTTGGAATTTCGAGGAAGTTGTGGATAAGGGAGCTTCGGCACAATCCTTCATCGA ACGAATGACCAACTTCGACAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTA CGAATACTTCACTGTCTACAACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCC GGCCTTTCTGAGCGGAGAACAGAAGAAAGCGATTGTCGATCTGCTGTTCAAGACCAACCGCAAGG TGACCGTCAAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTCGACTCAGTGGAAATCA GCGGAGTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGAAGATCATCAAG GACAAGGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGAC CCTTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCTTCGACGATA AGGTCATGAAACAACTCAAGCGCCGCCGGTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGAT CAACGGTATTCGCGATAAACAGAGCGGTAAAACTATCCTGGATTTCCTCAAATCGGATGGCTTCGC TAATCGTAACTTCATGCAGTTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAGAAAG CACAAGTGAGCGGACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTCGCCGGC GATTAAGAAGGGAATCCTGCAAACTGTGAAGGTGGTGGACGAGCTGGTGAAGGTCATGGGACGG CACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACCAGACTACCCAGAAGGGCCAGAA GAACTCCCGCGAAAGGATGAAGCGGATCGAAGAAGGAATCAAGGAGCTGGGCAGCCAGATCCTG AAAGAGCACCCGGTGGAAAACACGCAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAA	

	TGGACGGGACATGTACGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACC ACATCGTTCCACAGTCCTTTCTGAAGGATGACTCCATCGATAACAAGGTGTTGACTCGCAGCGACA AGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGGTCGTGAAGAAGATGAAGAATTACTGG CGGCAGCTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCG CGGCGGACTCTCAGAGCTGGATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAG ATTACCAAGCACGTGGCGCAGATCCTGGACTCCCGCATGAACACTAAATACGACGAGAACGATAA GCTCATCCGGGAAGTGAAGGTGATTACCCTGAAAAGCAAACTTGTGTCGGACTTTCGGAAGGACT TTCAGTTTTACAAAGTGAGAGAAATCAACAACTACCATCACGCGCATGACGCATACCTCAACGCTG TGGTCGGCACCGCCCTGATCAAGAAGTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTAC AAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGA AATACTTCTTTTACTCAAACATCATGAACTTCTTCAAGACTGAAATTACGCTGGCCAATGGAGAAAT CAGGAAGAGGCCACTGATCGAAACTAACGGAGAAACGGGCGAAATCGTGTGGGACAAGGGCAGG GACTTCGCAACTGTTCGCAAAGTGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTG CAAACCGGCGGATTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACG CAAGAAAGACTGGGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTC CTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAGAAGCTCAAATCCGTCAAAGAGCTGCTGG GGATTACCATCATGGAACGATCCTCGTTCGAGAAGAACCCGATTGATTTCCTGGAGGCGAAGGGT TACAAGGAGGTGAAGAAGGATCTGATCATCAAACTGCCCAAGTACTCACTGTTCGAACTGGAAAAT GGTCGGAAGCGCATGCTGGCTTCGGCCGGAGAACTCCAGAAAGGAAATGAGCTGGCCTTGCCTA GCAAGTACGTCAACTTCCTCTATCTTGCTTCGCACTACGAGAAACTCAAAGGGTCACCGGAAGATA ACGAACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAATCT CCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGTCGGCCTACAAT AAGCATAGAGATAAGCCGATCAGAGAACAGGCCGAGAACATTATCCACTTGTTCACCCTGACTAA CCTGGGAGCTCCAGCCGCCTTCAAGTACTTCGATACTACTATCGACCGCAAAAGATACACGTCCA CCAAGGAAGTTCTGGACGCGACCCTGATCCACCAAAGCATCACTGGACTCTACGAAACTAGGATC GATCTGTCGCAGCTGGGTGGCGATGGTGGCGGTGGATCCTACCCATACGACGTGCCTGACTACG CCTCCGGAGGTGGTGGCCCCAAGAAGAAACGGAAGGTGTGATAGCTAGCCATCACATTTAAAAGC ATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCATCTCTTTTTCTT TTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTC	

	TCTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 45 e 3' UTR de ALB	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCTATGGATAAGAAGTACTCGATCGGGCTGGATATCGGAACTAATTCCGTGGGTTGGGCAGT GATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGGAACACCGATAGACACA GCATCAAGAAGAATCTCATCGGAGCCCTGCTGTTTGACTCCGGCGAAACCGCAGAAGCGACCCG GCTCAAACGTACCGCGAGGCGACGCTACACCCGGCGGAAGAATCGCATCTGCTATCTGCAAGAA ATCTTTTCGAACGAAATGGCAAAGGTGGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTG GTGGAGGAGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTGGACGAAGTGGCGT ACCACGAAAAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAGGCC GACCTCAGATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCCGCGGACACTTCCTGATCGAA GGCGATCTGAACCCTGATAACTCCGACGTGGATAAGCTGTTCATTCAACTGGTGCAGACCTACAA CCAACTGTTCGAAGAAAACCCAATCAATGCCAGCGGCGTCGATGCCAAGGCCATCCTGTCCGCCC GGCTGTCGAAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGGGAGAGAAGAAGAACG GACTTTTCGGCAACTTGATCGCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACC TGGCCGAGGACGCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCT GGCACAAATTGGCGATCAGTACGCGGATCTGTTCCTTGCCGCTAAGAACCTTTCGGACGCAATCT TGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAGCGCCTCGATGATT AAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTCGTGAGACAGCAACTGC CTGAAAAGTACAAGGAGATTTTCTTCGACCAGTCCAAGAATGGGTACGCAGGGTACATCGATGGA GGCGCCAGCCAGGAAGAGTTCTATAAGTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGA AGAACTGCTGGTCAAGCTGAACAGGGAGGATCTGCTCCGCAAACAGAGAACCTTTGACAACGGAA GCATTCCACACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAGGACTTTTAC CCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCCGTATTACGT GGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAAAATCAGAGGAAACCATC ACTCCTTGGAATTTCGAGGAAGTTGTGGATAAGGGAGCTTCGGCACAATCCTTCATCGAACGAAT GACCAACTTCGACAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTACGAATA CTTCACTGTCTACAACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTT TCTGAGCGGAGAACAGAAGAAAGCGATTGTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCG	**47**

	TCAAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTCGACTCAGTGGAAATCAGCGGA GTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGAAGATCATCAAGGACAA GGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGACCCTTT TCGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCTTCGACGATAAGGTC ATGAAACAACTCAAGCGCCGCCGGTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACG GTATTCGCGATAAACAGAGCGGTAAAACTATCCTGGATTTCCTCAAATCGGATGGCTTCGCTAATC GTAACTTCATGCAGTTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAGAAAGCACAA GTGAGCGGACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTCGCCGGCGATTA AGAAGGGAATCCTGCAAACTGTGAAGGTGGTGGACGAGCTGGTGAAGGTCATGGGACGGCACAA ACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACCAGACTACCCAGAAGGGCCAGAAGAACT CCCGCGAAAGGATGAAGCGGATCGAAGAAGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGA GCACCCGGTGGAAAACACGCAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAATGGAC GGGACATGTACGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACCACATC GTTCCACAGTCCTTTCTGAAGGATGACTCCATCGATAACAAGGTGTTGACTCGCAGCGACAAGAA CAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGGTCGTGAAGAAGATGAAGAATTACTGGCGG CAGCTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCGCGG CGGACTCTCAGAGCTGGATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAGATTA CCAAGCACGTGGCGCAGATCCTGGACTCCCGCATGAACACTAAATACGACGAGAACGATAAGCTC ATCCGGGAAGTGAAGGTGATTACCCTGAAAAGCAAACTTGTGTCGGACTTTCGGAAGGACTTTCA GTTTTACAAAGTGAGAGAAATCAACAACTACCATCACGCGCATGACGCATACCTCAACGCTGTGGT CGGCACCGCCCTGATCAAGAAGTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTACAAGG TCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGAAATAC TTCTTTTACTCAAACATCATGAACTTCTTCAAGACTGAAATTACGCTGGCCAATGGAGAAATCAGGA AGAGGCCACTGATCGAAACTAACGGAGAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTT CGCAACTGTTCGCAAAGTGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAAC CGGCGGATTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAAGA AAGACTGGGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTCCTCGT GGTGGCCAAGGTGGAGAAGGGAAAGAGCAAGAAGCTCAAATCCGTCAAAGAGCTGCTGGGGATT ACCATCATGGAACGATCCTCGTTCGAGAAGAACCCGATTGATTTCCTGGAGGCGAAGGGTTACAA	

	GGAGGTGAAGAAGGATCTGATCATCAAACTGCCCAAGTACTCACTGTTCGAACTGGAAAATGGTC GGAAGCGCATGCTGGCTTCGGCCGGAGAACTCCAGAAAGGAAATGAGCTGGCCTTGCCTAGCAA GTACGTCAACTTCCTCTATCTTGCTTCGCACTACGAGAAACTCAAAGGGTCACCGGAAGATAACGA ACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAATCTCCGA GTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGTCGGCCTACAATAAGC ATAGAGATAAGCCGATCAGAGAACAGGCCGAGAACATTATCCACTTGTTCACCCTGACTAACCTG GGAGCTCCAGCCGCCTTCAAGTACTTCGATACTACTATCGACCGCAAAAGATACACGTCCACCAA GGAAGTTCTGGACGCGACCCTGATCCACCAAAGCATCACTGGACTCTACGAAACTAGGATCGATC TGTCGCAGCTGGGTGGCGATGGTGGCGGTGGATCCTACCCATACGACGTGCCTGACTACGCCTC CGGAGGTGGTGGCCCCAAGAAGAAACGGAAGGTGTGATAGCTAGCCATCACATTTAAAAGCATCT CAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCATCTCTTTTTCTTTTTC GTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCTCTG TGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
Transcrição Cas9 compreendendo ORF de Cas9 usando códons com expressão geralmente alta em seres humanos	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCCATGCCTAAGAAAAAGCGGAAGGTCGACGGGGATAAGAAGTACTCAATCGGGCTGGATAT CGGAACTAATTCCGTGGGTTGGGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCA AGGTCCTGGGGAACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGTTTGAC TCCGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGCGACGCTACACCCGGCGG AAGAATCGCATCTGCTATCTGCAAGAGATCTTTTCGAACGAAATGGCAAAGGTCGACGACAGCTTC TTCCACCGCCTGGAAGAATCTTTCCTGGTGGAGGAGGACAAGAAGCATGAACGGCATCCTATCTT TGGAAACATCGTCGACGAAGTGGCGTACCACGAAAAGTACCCGACCATCTACCATCTGCGGAAGA AGTTGGTTGACTCAACTGACAAGGCCGACCTCAGATTGATCTACTTGGCCCTCGCCCATATGATCA AATTCCGCGGACACTTCCTGATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGGATAAGCTT TTCATTCAACTGGTGCAGACCTACAACCAACTGTTCGAAGAAAACCCAATCAATGCTAGCGGCGTC GATGCCAAGGCCATCCTGTCCGCCCGGCTGTCGAAGTCGCGGCGCCTCGAAAACCTGATCGCAC AGCTGCCGGGAGAGAAAAAGAACGGACTTTTCGGCAACTTGATCGCTCTCTCACTGGGACTCACT CCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGACGCGAAGCTGCAACTCTCAAAGGACACCTA CGACGACGACTTGGACAATTTGCTGGCACAAATTGGCGATCAGTACGCGGATCTGTTCCTTGCCG	**48**

	CTAAGAACCTTTCGGACGCAATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACCAAA GCGCCGCTTAGCGCCTCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCA AAGCGCTCGTGAGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAAT GGGTACGCAGGGTACATCGATGGAGGCGCTAGCCAGGAAGAGTTCTATAAGTTCATCAAGCCAAT CCTGGAAAAGATGGACGGAACCGAAGAACTGCTGGTCAAGCTGAACAGGGAGGATCTGCTCCGG AAACAGAGAACCTTTGACAACGGATCCATTCCCCACCAGATCCATCTGGGTGAGCTGCACGCCAT CTTGCGGCGCCAGGAGGACTTTTACCCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTC TGACGTTCCGCATCCCGTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGAT GACTAGAAAATCAGAGGAAACCATCACTCCTTGGAATTTCGAGGAAGTTGTGGATAAGGGAGCTT CGGCACAAAGCTTCATCGAACGAATGACCAACTTCGACAAGAATCTCCCAAACGAGAAGGTGCTT CCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCTACAACGAACTGACTAAAGTGAAATACGTT ACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGAGAACAGAAGAAAGCAATTGTCGATCTGCT GTTCAAGACCAACCGCAAGGTGACCGTCAAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAGT GTTTCGACTCAGTGGAAATCAGCGGGGTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCAT GATCTCCTGAAGATCATCAAGGACAAGGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGA AGATATCGTCCTGACCTTGACCCTTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACCT ACGCTCATCTCTTCGACGATAAGGTCATGAAACAACTCAAGCGCCGCCGGTACACTGGTTGGGGC CGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGATAAACAGAGCGGTAAAACTATCCTGGATTT CCTCAAATCGGATGGCTTCGCTAATCGTAACTTCATGCAATTGATCCACGACGACAGCCTGACCTT TAAGGAGGACATCCAAAAAGCACAAGTGTCCGGACAGGGAGACTCACTCCATGAACACATCGCGA ATCTGGCCGGTTCGCCGGCGATTAAGAAGGGAATTCTGCAAACTGTGAAGGTGGTCGACGAGCT GGTGAAGGTCATGGGACGGCACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACCAG ACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATGAAGCGGATCGAAGAAGGAATCAAGG AGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACGCAGCTGCAGAACGAGAAGCT CTACCTGTACTATTTGCAAAATGGACGGGACATGTACGTGGACCAAGAGCTGGACATCAATCGGT TGTCTGATTACGACGTGGACCACATCGTTCCACAGTCCTTTCTGAAGGATGACTCGATCGATAACA AGGTGTTGACTCGCAGCGACAAGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGGTCGT GAAGAAGATGAAGAATTACTGGCGGCAGCTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTG ACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAGCTGGATAAGGCTGGATTCATCAAACGG	

	CAGCTGGTCGAGACTCGGCAGATTACCAAGCACGTGGCGCAGATCTTGGACTCCCGCATGAACA CTAAATACGACGAGAACGATAAGCTCATCCGGGAAGTGAAGGTGATTACCCTGAAAAGCAAACTT GTGTCGGACTTTCGGAAGGACTTTCAGTTTTACAAAGTGAGAGAAATCAACAACTACCATCACGCG CATGACGCATACCTCAACGCTGTGGTCGGTACCGCCCTGATCAAAAAGTACCCTAAACTTGAATC GGAGTTTGTGTACGGAGACTACAAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAG GAAATCGGGAAAGCAACTGCGAAATACTTCTTTTACTCAAACATCATGAACTTTTTCAAGACTGAAA TTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGGAGAAACGGGCGA AATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTCGCAAAGTGCTCTCTATGCCGCAAGTCA ATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGATTTTCAAAGGAATCGATCCTCCCAAAGAGA AATAGCGACAAGCTCATTGCACGCAAGAAAGACTGGGACCCGAAGAAGTACGGAGGATTCGATTC GCCGACTGTCGCATACTCCGTCCTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAAGCTC AAATCCGTCAAAGAGCTGCTGGGGATTACCATCATGGAACGATCCTCGTTCGAGAAGAACCCGAT TGATTTCCTCGAGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAAACTCCCCAAGT ACTCACTGTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGCTTCGGCCGGAGAACTCCAAAAA GGAAATGAGCTGGCCTTGCCTAGCAAGTACGTCAACTTCCTCTATCTTGCTTCGCACTACGAAAAA CTCAAAGGGTCACCGGAAGATAACGAACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCT GGATGAAATCATCGAACAAATCTCCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCG ACAAAGTCCTGTCGGCCTACAATAAGCATAGAGATAAGCCGATCAGAGAACAGGCCGAGAACATT ATCCACTTGTTCACCCTGACTAACCTGGGAGCCCCAGCCGCCTTCAAGTACTTCGATACTACTATC GATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGACGCGACCCTGATCCACCAAAGCATCAC TGGACTCTACGAAACTAGGATCGATCTGTCGCAGCTGGGTGGCGATTGATAGTCTAGCCATCACA TTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCATCT CTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATTTTG CCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
Transcrição Cas9 compreendendo a sequência Kozak com ORF de Cas9 usando códons com expressão geralmente alta em seres humanos	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCCGCCACCATGCCTAAGAAAAAGCGGAAGGTCGACGGGGATAAGAAGTACTCAATCGGGC TGGATATCGGAACTAATTCCGTGGGTTGGGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAG AAGTTCAAGGTCCTGGGGAACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCT	**49**

	GTTTGACTCCGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGCGACGCTACACC CGGCGGAAGAATCGCATCTGCTATCTGCAAGAGATCTTTTCGAACGAAATGGCAAAGGTCGACGA CAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTGGTGGAGGAGGACAAGAAGCATGAACGGCATC CTATCTTTGGAAACATCGTCGACGAAGTGGCGTACCACGAAAAGTACCCGACCATCTACCATCTG CGGAAGAAGTTGGTTGACTCAACTGACAAGGCCGACCTCAGATTGATCTACTTGGCCCTCGCCCA TATGATCAAATTCCGCGGACACTTCCTGATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGG ATAAGCTTTTCATTCAACTGGTGCAGACCTACAACCAACTGTTCGAAGAAAACCCAATCAATGCTA GCGGCGTCGATGCCAAGGCCATCCTGTCCGCCCGGCTGTCGAAGTCGCGGCGCCTCGAAAACCT GATCGCACAGCTGCCGGGAGAGAAAAAGAACGGACTTTTCGGCAACTTGATCGCTCTCTCACTGG GACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGACGCGAAGCTGCAACTCTCAAAG GACACCTACGACGACGACTTGGACAATTTGCTGGCACAAATTGGCGATCAGTACGCGGATCTGTT CCTTGCCGCTAAGAACCTTTCGGACGCAATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAA TAACCAAAGCGCCGCTTAGCGCCTCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACG CTGCTCAAAGCGCTCGTGAGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTC CAAGAATGGGTACGCAGGGTACATCGATGGAGGCGCTAGCCAGGAAGAGTTCTATAAGTTCATCA AGCCAATCCTGGAAAAGATGGACGGAACCGAAGAACTGCTGGTCAAGCTGAACAGGGAGGATCT GCTCCGGAAACAGAGAACCTTTGACAACGGATCCATTCCCCACCAGATCCATCTGGGTGAGCTGC ACGCCATCTTGCGGCGCCAGGAGGACTTTTACCCATTCCTCAAGGACAACCGGGAAAAGATCGAG AAAATTCTGACGTTCCGCATCCCGTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGC GTGGATGACTAGAAAATCAGAGGAAACCATCACTCCTTGGAATTTCGAGGAAGTTGTGGATAAGG GAGCTTCGGCACAAAGCTTCATCGAACGAATGACCAACTTCGACAAGAATCTCCCAAACGAGAAG GTGCTTCCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCTACAACGAACTGACTAAAGTGAAA TACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGAGAACAGAAGAAAGCAATTGTCGA TCTGCTGTTCAAGACCAACCGCAAGGTGACCGTCAAGCAGCTTAAAGAGGACTACTTCAAGAAGA TCGAGTGTTTCGACTCAGTGGAAATCAGCGGGGTGGAGGACAGATTCAACGCTTCGCTGGGAAC CTATCATGATCTCCTGAAGATCATCAAGGACAAGGACTTCCTTGACAACGAGGAGAACGAGGACA TCCTGGAAGATATCGTCCTGACCTTGACCCTTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTT AAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAACTCAAGCGCCGCCGGTACACTGG TTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGATAAACAGAGCGGTAAAACTATCC	

	TGGATTTCCTCAAATCGGATGGCTTCGCTAATCGTAACTTCATGCAATTGATCCACGACGACAGCC TGACCTTTAAGGAGGACATCCAAAAAGCACAAGTGTCCGGACAGGGAGACTCACTCCATGAACAC ATCGCGAATCTGGCCGGTTCGCCGGCGATTAAGAAGGGAATTCTGCAAACTGTGAAGGTGGTCG ACGAGCTGGTGAAGGTCATGGGACGGCACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGA AAACCAGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATGAAGCGGATCGAAGAAGGA ATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACGCAGCTGCAGAACG AGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGACATGTACGTGGACCAAGAGCTGGACATC AATCGGTTGTCTGATTACGACGTGGACCACATCGTTCCACAGTCCTTTCTGAAGGATGACTCGATC GATAACAAGGTGTTGACTCGCAGCGACAAGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGG AGGTCGTGAAGAAGATGAAGAATTACTGGCGGCAGCTCCTGAATGCGAAGCTGATTACCCAGAGA AAGTTTGACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAGCTGGATAAGGCTGGATTCAT CAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACGTGGCGCAGATCTTGGACTCCCGC ATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAAGTGAAGGTGATTACCCTGAAAAG CAAACTTGTGTCGGACTTTCGGAAGGACTTTCAGTTTTACAAAGTGAGAGAAATCAACAACTACCA TCACGCGCATGACGCATACCTCAACGCTGTGGTCGGTACCGCCCTGATCAAAAAGTACCCTAAAC TTGAATCGGAGTTTGTGTACGGAGACTACAAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCC GAACAGGAAATCGGGAAAGCAACTGCGAAATACTTCTTTTACTCAAACATCATGAACTTTTTCAAGA CTGAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGGAGAAACG GGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTCGCAAAGTGCTCTCTATGCCGC AAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGATTTTCAAAGGAATCGATCCTCCCAA AGAGAAATAGCGACAAGCTCATTGCACGCAAGAAAGACTGGGACCCGAAGAAGTACGGAGGATT CGATTCGCCGACTGTCGCATACTCCGTCCTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAA AAGCTCAAATCCGTCAAAGAGCTGCTGGGGATTACCATCATGGAACGATCCTCGTTCGAGAAGAA CCCGATTGATTTCCTCGAGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAAACTCC CCAAGTACTCACTGTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGCTTCGGCCGGAGAACTC CAAAAAGGAAATGAGCTGGCCTTGCCTAGCAAGTACGTCAACTTCCTCTATCTTGCTTCGCACTAC GAAAAACTCAAAGGGTCACCGGAAGATAACGAACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCA TTATCTGGATGAAATCATCGAACAAATCTCCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAA CCTCGACAAAGTCCTGTCGGCCTACAATAAGCATAGAGATAAGCCGATCAGAGAACAGGCCGAGA	

	ACATTATCCACTTGTTCACCCTGACTAACCTGGGAGCCCCAGCCGCCTTCAAGTACTTCGATACTA CTATCGATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGACGCGACCCTGATCCACCAAAGC ATCACTGGACTCTACGAAACTAGGATCGATCTGTCGCAGCTGGGTGGCGATTGATAGTCTAGCCA TCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATT CATCTCTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCA TTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
ORF de Cas9 com junções de emenda removidas; 12,75% de conteúdo em U	ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGATGGGCAGTCATCA CAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGGAAACACAGACAGACACAGCATC AAGAAGAACCTGATCGGAGCACTGCTGTTCGACAGCGGAGAAACAGCAGAAGCAACAAGACTGA AGAGAACAGCAAGAAGAAGATACACAAGAAGAAAGAACAGAATCTGCTACCTGCAGGAAATCTTC AGCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACcggCTGGAAGAAAGCTTCCTGGTCGAA GAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAGTCGCATACCACGA AAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGACAAGGCAGACCTGA GACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGACACTTCCTGATCGAAGGAGAC CTGAACCCGGACAACAGCGACGTCGACAAGCTGTTCATCCAGCTGGTCCAGACATACAACCAGCT GTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAAGGCAATCCTGAGCGCAAGACTG AGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCCGGGAGAAAAGAAGAACGGACTGT TCGGAAACCTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAACTTCGACCTGGCA GAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGACAACCTGCTGGCAC AGATCGGAGACCAGTACGCAGACCTGTTCCTGGCAGCAAAGAACCTGAGCGACGCAATCCTGCT GAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGGCACCGCTGAGCGCAAGCATGATCAAGA GATACGACGAACACCACCAGGACCTGACACTGCTGAAGGCACTGGTCAGACAGCAGCTGCCGGA AAAGTACAAGGAAATCTTCTTCGACCAGAGCAAGAACGGATACGCAGGATACATCGACGGAGGAG CAAGCCAGGAAGAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAAGAA CTGCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTCGACAACGGAAGCAT CCCGCACCAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAAGACAGGAAGACTTCTACCCG TTCCTGAAGGACAACAGAGAAAAGATCGAAAAGATCCTGACATTCAGAATCCCGTACTACGTCGG ACCGCTGGCAAGAGGAAACAGCAGATTCGCATGGATGACAAGAAAGAGCGAAGAAACAATCACAC	**50**

	CGTGGAACTTCGAAGAAGTCGTCGACAAGGGAGCAAGCGCACAGAGCTTCATCGAAAGAATGAC AAACTTCGACAAGAACCTGCCGAACGAAAAGGTCCTGCCGAAGCACAGCCTGCTGTACGAATACT TCACAGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGAATGAGAAAGCCGGCATTC CTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGACAAACAGAAAGGTCACAGT CAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCTTCGACAGCGTCGAAATCAGCGGAG TCGAAGACAGATTCAACGCAAGCCTGGGAACATACCACGACCTGCTGAAGATCATCAAGGACAAG GACTTCCTGGACAACGAAGAAAACGAAGACATCCTGGAAGACATCGTCCTGACACTGACACTGTT CGAAGACAGAGAAATGATCGAAGAAAGACTGAAGACATACGCACACCTGTTCGACGACAAGGTCA TGAAGCAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTGAGCAGAAAGCTGATCAACGG AATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGACGGATTCGCAAACA GAAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCAAGGAAGACATCCAGAAGGCACAG GTCAGCGGACAGGGAGACAGCCTGCACGAACACATCGCAAACCTGGCAGGAAGCCCGGCAATCA AGAAGGGAATCCTGCAGACAGTCAAGGTCGTCGACGAACTGGTCAAGGTCATGGGAAGACACAA GCCGGAAAACATCGTCATCGAAATGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACA GCAGAGAAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGCCAGATCCTGAAGGA ACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTACCTGCAaAACGGAA GAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGCGACTACGACGTCGACCACATC GTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTCCTGACAAGAAGCGACAAGA ACAGAGGAAAGAGCGACAACGTCCCGAGCGAAGAAGTCGTCAAGAAGATGAAGAACTACTGGAG ACAGCTGCTGAACGCAAAGCTGATCACACAGAGAAAGTTCGACAACCTGACAAAGGCAGAGAGAG GAGGACTGAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTCGAAACAAGACAGAT CACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAAAGTACGACGAAAACGACAAGC TGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTCAGCGACTTCAGAAAGGACTTC CAGTTCTACAAGGTCAGAGAAATCAACAACTACCACCACGCACACGACGCATACCTGAACGCAGT CGTCGGAACAGCACTGATCAAGAAGTACCCGAAGCTGGAAAGCGAATTCGTCTACGGAGACTACA AGGTCTACGACGTCAGAAAGATGATCGCAAAGAGCGAACAGGAAATCGGAAAGGCAACAGCAAA GTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGACAGAAATCACACTGGCAAACGGAGAAAT CAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAAATCGTCTGGGACAAGGGAAGA GACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCAACATCGTCAAGAAGACAGAAGT	

	CCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCGAAGAGAAACAGCGACAAGCTGATCGCA AGAAAGAAGGACTGGGACCCGAAGAAGTACGGAGGATTCGACAGCCCGACAGTCGCATACAGCG TCCTGGTCGTCGCAAAGGTCGAAAAGGGAAAGAGCAAGAAGCTGAAGAGCGTCAAGGAACTGCT GGGAATCACAATCATGGAAAGAAGCAGCTTCGAAAAGAACCCGATCGACTTCCTGGAAGCAAAGG GATACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCTGTTCGAACTGGAA AACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGAAACGAACTGGCACTGC CGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTACGAAAAGCTGAAGGGAAGCCCGGA AGACAACGAACAGAAGCAGCTGTTCGTCGAACAGCACAAGCACTACCTGGACGAAATCATCGAAC AGATCAGCGAATTCAGCAAGAGAGTCATCCTGGCAGACGCAAACCTGGACAAGGTCCTGAGCGC ATACAACAAGCACAGAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCCACCTGTTCACAC TGACAAACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGACAGAAAGAGATAC ACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGAGCATCACAGGACTGTACGAAAC AAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGT CTAG	
Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 50, sequência de Kozak e 3' UTR de ALB	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCCGCCACCATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGAT GGGCAGTCATCACAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGGAAACACAGA CAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGCTGTTCGACAGCGGAGAAACAGCAGAA GCAACAAGACTGAAGAGAACAGCAAGAAGAAGATACACAAGAAGAAAGAACAGAATCTGCTACCT GCAGGAAATCTTCAGCAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACcggCTGGAAGAAAG CTTCCTGGTCGAAGAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAG TCGCATACCACGAAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGAC AAGGCAGACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGACACTTCCT GATCGAAGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTTCATCCAGCTGGTCCAG ACATACAACCAGCTGTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAAGGCAATCCT GAGCGCAAGACTGAGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCCGGGAGAAAAG AAGAACGGACTGTTCGGAAACCTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAA CTTCGACCTGGCAGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGAC	**51**

	AACCTGCTGGCACAGATCGGAGACCAGTACGCAGACCTGTTCCTGGCAGCAAAGAACCTGAGCG ACGCAATCCTGCTGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGGCACCGCTGAGCGCA AGCATGATCAAGAGATACGACGAACACCACCAGGACCTGACACTGCTGAAGGCACTGGTCAGACA GCAGCTGCCGGAAAAGTACAAGGAAATCTTCTTCGACCAGAGCAAGAACGGATACGCAGGATACA TCGACGGAGGAGCAAGCCAGGAAGAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGAC GGAACAGAAGAACTGCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTCGA CAACGGAAGCATCCCGCACCAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAAGACAGGAA GACTTCTACCCGTTCCTGAAGGACAACAGAGAAAAGATCGAAAAGATCCTGACATTCAGAATCCC GTACTACGTCGGACCGCTGGCAAGAGGAAACAGCAGATTCGCATGGATGACAAGAAAGAGCGAA GAAACAATCACACCGTGGAACTTCGAAGAAGTCGTCGACAAGGGAGCAAGCGCACAGAGCTTCAT CGAAAGAATGACAAACTTCGACAAGAACCTGCCGAACGAAAAGGTCCTGCCGAAGCACAGCCTGC TGTACGAATACTTCACAGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGAATGAGAA AGCCGGCATTCCTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGACAAACAG AAAGGTCACAGTCAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCTTCGACAGCGTCG AAATCAGCGGAGTCGAAGACAGATTCAACGCAAGCCTGGGAACATACCACGACCTGCTGAAGATC ATCAAGGACAAGGACTTCCTGGACAACGAAGAAAACGAAGACATCCTGGAAGACATCGTCCTGAC ACTGACACTGTTCGAAGACAGAGAAATGATCGAAGAAAGACTGAAGACATACGCACACCTGTTCG ACGACAAGGTCATGAAGCAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTGAGCAGAAA GCTGATCAACGGAATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGAC GGATTCGCAAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCAAGGAAGACAT CCAGAAGGCACAGGTCAGCGGACAGGGAGACAGCCTGCACGAACACATCGCAAACCTGGCAGG AAGCCCGGCAATCAAGAAGGGAATCCTGCAGACAGTCAAGGTCGTCGACGAACTGGTCAAGGTC ATGGGAAGACACAAGCCGGAAAACATCGTCATCGAAATGGCAAGAGAAAACCAGACAACACAGAA GGGACAGAAGAACAGCAGAGAAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGC CAGATCCTGAAGGAACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTA CCTGCAaAACGGAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGCGACTACG ACGTCGACCACATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTCCTGAC AAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGTCCCGAGCGAAGAAGTCGTCAAGAAGATG AAGAACTACTGGAGACAGCTGCTGAACGCAAAGCTGATCACACAGAGAAAGTTCGACAACCTGAC	

	AAAGGCAGAGAGAGGAGGACTGAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTC GAAACAAGACAGATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAAAGTACGA CGAAAACGACAAGCTGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTCAGCGACT TCAGAAAGGACTTCCAGTTCTACAAGGTCAGAGAAATCAACAACTACCACCACGCACACGACGCA TACCTGAACGCAGTCGTCGGAACAGCACTGATCAAGAAGTACCCGAAGCTGGAAAGCGAATTCGT CTACGGAGACTACAAGGTCTACGACGTCAGAAAGATGATCGCAAAGAGCGAACAGGAAATCGGAA AGGCAACAGCAAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGACAGAAATCACACTGG CAAACGGAGAAATCAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAAATCGTCTGG GACAAGGGAAGAGACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCAACATCGTCA AGAAGACAGAAGTCCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCGAAGAGAAACAGCGA CAAGCTGATCGCAAGAAAGAAGGACTGGGACCCGAAGAAGTACGGAGGATTCGACAGCCCGACA GTCGCATACAGCGTCCTGGTCGTCGCAAAGGTCGAAAAGGGAAAGAGCAAGAAGCTGAAGAGCG TCAAGGAACTGCTGGGAATCACAATCATGGAAAGAAGCAGCTTCGAAAAGAACCCGATCGACTTC CTGGAAGCAAAGGGATACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCT GTTCGAACTGGAAAACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGAAAC GAACTGGCACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTACGAAAAGCTGAA GGGAAGCCCGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAACAGCACAAGCACTACCTGGAC GAAATCATCGAACAGATCAGCGAATTCAGCAAGAGAGTCATCCTGGCAGACGCAAACCTGGACAA GGTCCTGAGCGCATACAACAAGCACAGAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCC ACCTGTTCACACTGACAAACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGAC AGAAAGAGATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGAGCATCACAGG ACTGTACGAAACAAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAGCCCGAAGAAG AAGAGAAAGGTCTAGCTAGCCATCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAG AAAATGAAGATCAATAGCTTATTCATCTCTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTA AAAAACATAAATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGA ACCTCGAG	
Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCCGCCACCATGGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACAGCGTGGGCT	**53**

ID NO: 52, sequência de Kozak e 3' UTR de ALB	GGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGA CAGACACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGCGAGACCGCCGAG GCCACCAGACTGAAGAGAACCGCCAGAAGAAGATACACCAGAAGAAAGAACAGAATCTGCTACCT GCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAGGAG AGCTTCCTGGTGGAGGAGGACAAGAAGCACGAGAGACACCCCATCTTCGGCAACATCGTGGACG AGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAGCTGGTGGACAGCAC CGACAAGGCCGACCTGAGACTGATCTACCTGGCCCTGGCCCACATGATCAAGTTCAGAGGCCAC TTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGG TGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC CATCCTGAGCGCCAGACTGAGCAAGAGCAGAAGACTGGAGAACCTGATCGCCCAGCTGCCCGGC GAGAAGAAGAACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCA AGAGCAACTTCGACCTGGCCGAGGACGCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGA CCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTCCTGGCCGCCAAGAAC CTGAGCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCC TGAGCGCCAGCATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCT GGTGAGACAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTAC GCCGGCTACATCGACGGCGGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGG AGAAGATGGACGGCACCGAGGAGCTGCTGGTGAAGCTGAACAGAGAGGACCTGCTGAGAAAGCA GAGAACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGCGAGCTGCACGCCATCCTG AGAAGACAGGAGGACTTCTACCCCTTCCTGAAGGACAACAGAGAGAAGATCGAGAAGATCCTGAC CTTCAGAATCCCCTACTACGTGGGCCCCCTGGCCAGAGGCAACAGCAGATTCGCCTGGATGACC AGAAAGAGCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGC GCCCAGAGCTTCATCGAGAGAATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGC CCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACAACGAGCTGACCAAGGTGAAGTACGTG ACCGAGGGCATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCCATCGTGGACCTG CTGTTCAAGACCAACAGAAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGA GTGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACAGATTCAACGCCAGCCTGGGCACCTAC CACGACCTGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCT GGAGGACATCGTGCTGACCCTGACCCTGTTCGAGGACAGAGAGATGATCGAGGAGAGACTGAAG	

	ACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTGAAGAGAAGAAGATACACCGGCT GGGGCAGACTGAGCAGAAAGCTGATCAACGGCATCAGAGACAAGCAGAGCGGCAAGACCATCCT GGACTTCCTGAAGAGCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGC CTGACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAG CACATCGCCAACCTGGCCGGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGG TGGACGAGCTGGTGAAGGTGATGGGCAGACACAAGCCCGAGAACATCGTGATCGAGATGGCCAG AGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCAGAGAGAGAATGAAGAGAATCGAGGAG GGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCAGAGACATGTACGTGGACCAGGAGCTGG ACATCAACAGACTGAGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGA CAGCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACAGAGGCAAGAGCGACAACGTGCCC AGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGAGACAGCTGCTGAACGCCAAGCTGATCA CCCAGAGAAAGTTCGACAACCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAGCTGGACAAGG CCGGCTTCATCAAGAGACAGCTGGTGGAGACCAGACAGATCACCAAGCACGTGGCCCAGATCCT GGACAGCAGAATGAACACCAAGTACGACGAGAACGACAAGCTGATCAGAGAGGTGAAGGTGATC ACCCTGAAGAGCAAGCTGGTGAGCGACTTCAGAAAGGACTTCCAGTTCTACAAGGTGAGAGAGAT CAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTGGTGGGCACCGCCCTGATCAAG AAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGAGAAAGA TGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGCAACATC ATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCAGAAAGAGACCCCTGATCGA GACCAACGGCGAGACCGGCGAGATCGTGTGGGACAAGGGCAGAGACTTCGCCACCGTGAGAAA GGTGCTGAGCATGCCCCAGGTGAACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGC AAGGAGAGCATCCTGCCCAAGAGAAACAGCGACAAGCTGATCGCCAGAAAGAAGGACTGGGACC CCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTGGTGGTGGCCAAGGT GGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG AGAAGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGA AGGACCTGATCATCAAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCAGAAAGAGAAT GCTGGCCAGCGCCGGCGAGCTGCAGAAGGGCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAA CTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGCCCCGAGGACAACGAGCAGAAG	

	CAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCA GCAAGAGAGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACAG AGACAAGCCCATCAGAGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGC GCCCCCGCCGCCTTCAAGTACTTCGACACCACCATCGACAGAAAGAGATACACCAGCACCAAGGA GGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACCAGAATCGACCTG AGCCAGCTGGGCGGCGACGGCGGCGGCAGCCCCAAGAAGAAGAGAAAGGTGTGACTAGCCATC ACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCA TCTCTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATT TTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
Transcrição Cas9 com 5' UTR de HSD, ORF correspondente à SEQ ID NO: 54, sequência de Kozak e 3' UTR de ALB	GGGTCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC GGATCCGCCACCATGGACAAAAAATACAGCATAGGGCTAGACATAGGGACGAACAGCGTAGGGT GGGCGGTAATAACGGACGAATACAAAGTACCGAGCAAAAAATTCAAAGTACTAGGGAACACGGAC CGACACAGCATAAAAAAAAACCTAATAGGGGCGCTACTATTCGACAGCGGGGAAACGGCGGAAG CGACGCGACTAAAACGAACGGCGCGACGACGATACACGCGACGAAAAAACCGAATATGCTACCT ACAAGAAATATTCAGCAACGAAATGGCGAAAGTAGACGACAGCTTCTTCCACCGACTAGAAGAAA GCTTCCTAGTAGAAGAAGACAAAAAACACGAACGACACCCGATATTCGGGAACATAGTAGACGAA GTAGCGTACCACGAAAAATACCCGACGATATACCACCTACGAAAAAAACTAGTAGACAGCACGGA CAAAGCGGACCTACGACTAATATACCTAGCGCTAGCGCACATGATAAAATTCCGAGGGCACTTCC TAATAGAAGGGGACCTAAACCCGGACAACAGCGACGTAGACAAACTATTCATACAACTAGTACAAA CGTACAACCAACTATTCGAAGAAAACCCGATAAACGCGAGCGGGGTAGACGCGAAAGCGATACTA AGCGCGCGACTAAGCAAAAGCCGACGACTAGAAAACCTAATAGCGCAACTACCGGGGGAAAAAA AAAACGGGCTATTCGGGAACCTAATAGCGCTAAGCCTAGGGCTAACGCCGAACTTCAAAAGCAAC TTCGACCTAGCGGAAGACGCGAAACTACAACTAAGCAAAGACACGTACGACGACGACCTAGACAA CCTACTAGCGCAAATAGGGGACCAATACGCGGACCTATTCCTAGCGGCGAAAAACCTAAGCGACG CGATACTACTAAGCGACATACTACGAGTAAACACGGAAATAACGAAAGCGCCGCTAAGCGCGAGC ATGATAAAACGATACGACGAACACCACCAAGACCTAACGCTACTAAAAGCGCTAGTACGACAACAA CTACCGGAAAAATACAAAGAAATATTCTTCGACCAAAGCAAAAACGGGTACGCGGGGTACATAGA CGGGGGGGCGAGCCAAGAAGAATTCTACAAATTCATAAAACCGATACTAGAAAAAATGGACGGGA	**55**

	CGGAAGAACTACTAGTAAAACTAAACCGAGAAGACCTACTACGAAAACAACGAACGTTCGACAAC GGGAGCATACCGCACCAAATACACCTAGGGGAACTACACGCGATACTACGACGACAAGAAGACTT CTACCCGTTCCTAAAAGACAACCGAGAAAAAATAGAAAAAATACTAACGTTCCGAATACCGTACTA CGTAGGGCCGCTAGCGCGAGGGAACAGCCGATTCGCGTGGATGACGCGAAAAAGCGAAGAAAC GATAACGCCGTGGAACTTCGAAGAAGTAGTAGACAAAGGGGCGAGCGCGCAAAGCTTCATAGAA CGAATGACGAACTTCGACAAAAACCTACCGAACGAAAAAGTACTACCGAAACACAGCCTACTATAC GAATACTTCACGGTATACAACGAACTAACGAAAGTAAAATACGTAACGGAAGGGATGCGAAAACC GGCGTTCCTAAGCGGGGAACAAAAAAAAGCGATAGTAGACCTACTATTCAAAACGAACCGAAAAG TAACGGTAAAACAACTAAAAGAAGACTACTTCAAAAAAATAGAATGCTTCGACAGCGTAGAAATAA GCGGGGTAGAAGACCGATTCAACGCGAGCCTAGGGACGTACCACGACCTACTAAAAATAATAAAA GACAAAGACTTCCTAGACAACGAAGAAAACGAAGACATACTAGAAGACATAGTACTAACGCTAACG CTATTCGAAGACCGAGAAATGATAGAAGAACGACTAAAAACGTACGCGCACCTATTCGACGACAA AGTAATGAAACAACTAAAACGACGACGATACACGGGGTGGGGGCGACTAAGCCGAAAACTAATAA ACGGGATACGAGACAAACAAAGCGGGAAAACGATACTAGACTTCCTAAAAAGCGACGGGTTCGCG AACCGAAACTTCATGCAACTAATACACGACGACAGCCTAACGTTCAAAGAAGACATACAAAAAGCG CAAGTAAGCGGGCAAGGGGACAGCCTACACGAACACATAGCGAACCTAGCGGGGAGCCCGGCG ATAAAAAAAGGGATACTACAAACGGTAAAAGTAGTAGACGAACTAGTAAAAGTAATGGGGCGACAC AAACCGGAAAACATAGTAATAGAAATGGCGCGAGAAAACCAAACGACGCAAAAAGGGCAAAAAAA CAGCCGAGAACGAATGAAACGAATAGAAGAAGGGATAAAAGAACTAGGGAGCCAAATACTAAAAG AACACCCGGTAGAAAACACGCAACTACAAAACGAAAAACTATACCTATACTACCTACAAAACGGGC GAGACATGTACGTAGACCAAGAACTAGACATAAACCGACTAAGCGACTACGACGTAGACCACATA GTACCGCAAAGCTTCCTAAAAGACGACAGCATAGACAACAAAGTACTAACGCGAAGCGACAAAAA CCGAGGGAAAAGCGACAACGTACCGAGCGAAGAAGTAGTAAAAAAAATGAAAAACTACTGGCGAC AACTACTAAACGCGAAACTAATAACGCAACGAAAATTCGACAACCTAACGAAAGCGGAACGAGGG GGGCTAAGCGAACTAGACAAAGCGGGGTTCATAAAACGACAACTAGTAGAAACGCGACAAATAAC GAAACACGTAGCGCAAATACTAGACAGCCGAATGAACACGAAATACGACGAAAACGACAAACTAA TACGAGAAGTAAAAGTAATAACGCTAAAAAGCAAACTAGTAAGCGACTTCCGAAAAGACTTCCAAT TCTACAAAGTACGAGAAATAAACAACTACCACCACGCGCACGACGCGTACCTAAACGCGGTAGTA GGGACGGCGCTAATAAAAAAATACCCGAAACTAGAAAGCGAATTCGTATACGGGGACTACAAAGT	

	ATACGACGTACGAAAAATGATAGCGAAAAGCGAACAAGAAATAGGGAAAGCGACGGCGAAATACT TCTTCTACAGCAACATAATGAACTTCTTCAAAACGGAAATAACGCTAGCGAACGGGGAAATACGAA AACGACCGCTAATAGAAACGAACGGGGAAACGGGGGAAATAGTATGGGACAAAGGGCGAGACTT CGCGACGGTACGAAAAGTACTAAGCATGCCGCAAGTAAACATAGTAAAAAAAACGGAAGTACAAA CGGGGGGGTTCAGCAAAGAAAGCATACTACCGAAACGAAACAGCGACAAACTAATAGCGCGAAAA AAAGACTGGGACCCGAAAAAATACGGGGGGTTCGACAGCCCGACGGTAGCGTACAGCGTACTAG TAGTAGCGAAAGTAGAAAAAGGGAAAAGCAAAAAACTAAAAAGCGTAAAAGAACTACTAGGGATAA CGATAATGGAACGAAGCAGCTTCGAAAAAAACCCGATAGACTTCCTAGAAGCGAAAGGGTACAAA GAAGTAAAAAAAGACCTAATAATAAAACTACCGAAATACAGCCTATTCGAACTAGAAAACGGGCGA AAACGAATGCTAGCGAGCGCGGGGGAACTACAAAAAGGGAACGAACTAGCGCTACCGAGCAAAT ACGTAAACTTCCTATACCTAGCGAGCCACTACGAAAAACTAAAAGGGAGCCCGGAAGACAACGAA CAAAAACAACTATTCGTAGAACAACACAAACACTACCTAGACGAAATAATAGAACAAATAAGCGAAT TCAGCAAACGAGTAATACTAGCGGACGCGAACCTAGACAAAGTACTAAGCGCGTACAACAAACAC CGAGACAAACCGATACGAGAACAAGCGGAAAACATAATACACCTATTCACGCTAACGAACCTAGG GGCGCCGGCGGCGTTCAAATACTTCGACACGACGATAGACCGAAAACGATACACGAGCACGAAA GAAGTACTAGACGCGACGCTAATACACCAAAGCATAACGGGGCTATACGAAACGCGAATAGACCT AAGCCAACTAGGGGGGGACGGGGGGGGGAGCCCGAAAAAAAAACGAAAAGTATGACTAGCCATC ACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCA TCTCTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACATAAATTTCTTTAATCATT TTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAACCTCGAG	
sequência poli-A 100	AAAAAAAAA

	CAAGATCGGTCAGAAGACTGACAAGAAGAAGAGCACACAGACTGCTGAGAACAAGAAGACTGCTG AAGAGAGAAGGAGTCCTGCAGGCAGCAAACTTCGACGAAAACGGACTGATCAAGTCGCTGCCGA ACACACCGTGGCAGCTGAGAGCAGCAGCACTGGACAGAAAGCTGACACCGCTGGAATGGTCGGC AGTCCTGCTGCACCTGATCAAGCACAGAGGATACCTGTCGCAGAGAAAGAACGAAGGAGAAACA GCAGACAAGGAACTGGGAGCACTGCTGAAGGGAGTCGCAGGAAACGCACACGCACTGCAGACA GGAGACTTCAGAACACCGGCAGAACTGGCACTGAACAAGTTCGAAAAGGAATCGGGACACATCA GAAACCAGAGATCGGACTACTCGCACACATTCTCGAGAAAGGACCTGCAGGCAGAACTGATCCTG CTGTTCGAAAAGCAGAAGGAATTCGGAAACCCGCACGTCTCGGGAGGACTGAAGGAAGGAATCG AAACACTGCTGATGACACAGAGACCGGCACTGTCGGGAGACGCAGTCCAGAAGATGCTGGGACA CTGCACATTCGAACCGGCAGAACCGAAGGCAGCAAAGAACACATACACAGCAGAAAGATTCATCT GGCTGACAAAGCTGAACAACCTGAGAATCCTGGAACAGGGATCGGAAAGACCGCTGACAGACAC AGAAAGAGCAACACTGATGGACGAACCGTACAGAAAGTCGAAGCTGACATACGCACAGGCAAGAA AGCTGCTGGGACTGGAAGACACAGCATTCTTCAAGGGACTGAGATACGGAAAGGACAACGCAGA AGCATCGACACTGATGGAAATGAAGGCATACCACGCAATCTCGAGAGCACTGGAAAAGGAAGGAC TGAAGGACAAGAAGTCGCCGCTGAACCTGTCGCCGGAACTGCAGGACGAAATCGGAACAGCATT CTCGCTGTTCAAGACAGACGAAGACATCACAGGAAGACTGAAGGACAGAATCCAGCCGGAAATCC TGGAAGCACTGCTGAAGCACATCTCGTTCGACAAGTTCGTCCAGATCTCGCTGAAGGCACTGAGA AGAATCGTCCCGCTGATGGAACAGGGAAAGAGATACGACGAAGCATGCGCAGAAATCTACGGAG ACCACTACGGAAAGAAGAACACAGAAGAAAAGATCTACCTGCCGCCGATCCCGGCAGACGAAATC AGAAACCCGGTCGTCCTGAGAGCACTGTCGCAGGCAAGAAAGGTCATCAACGGAGTCGTCAGAA GATACGGATCGCCGGCAAGAATCCACATCGAAACAGCAAGAGAAGTCGGAAAGTCGTTCAAGGA CAGAAAGGAAATCGAAAAGAGACAGGAAGAAAACAGAAAGGACAGAGAAAAGGCAGCAGCAAAG TTCAGAGAATACTTCCCGAACTTCGTCGGAGAACCGAAGTCGAAGGACATCCTGAAGCTGAGACT GTACGAACAGCAGCACGGAAAGTGCCTGTACTCGGGAAAGGAAATCAACCTGGGAAGACTGAAC GAAAAGGGATACGTCGAAATCGACCACGCACTGCCGTTCTCGAGAACATGGGACGACTCGTTCAA CAACAAGGTCCTGGTCCTGGGATCGGAAAACCAGAACAAGGGAAACCAGACACCGTACGAATACT TCAACGGAAAGGACAACTCGAGAGAATGGCAGGAATTCAAGGCAAGAGTCGAAACATCGAGATTC CCGAGATCGAAGAAGCAGAGAATCCTGCTGCAGAAGTTCGACGAAGACGGATTCAAGGAAAGAAA CCTGAACGACACAAGATACGTCAACAGATTCCTGTGCCAGTTCGTCGCAGACAGAATGAGACTGA	

	CAGGAAAGGGAAAGAAGAGAGTCTTCGCATCGAACGGACAGATCACAAACCTGCTGAGAGGATTC TGGGGACTGAGAAAGGTCAGAGCAGAAAACGACAGACACCACGCACTGGACGCAGTCGTCGTCG CATGCTCGACAGTCGCAATGCAGCAGAAGATCACAAGATTCGTCAGATACAAGGAAATGAACGCA TTCGACGGAAAGACAATCGACAAGGAAACAGGAGAAGTCCTGCACCAGAAGACACACTTCCCGCA GCCGTGGGAATTCTTCGCACAGGAAGTCATGATCAGAGTCTTCGGAAAGCCGGACGGAAAGCCG GAATTCGAAGAAGCAGACACACTGGAAAAGCTGAGAACACTGCTGGCAGAAAAGCTGTCGTCGAG ACCGGAAGCAGTCCACGAATACGTCACACCGCTGTTCGTCTCGAGAGCACCGAACAGAAAGATGT CGGGACAGGGACACATGGAAACAGTCAAGTCGGCAAAGAGACTGGACGAAGGAGTCTCGGTCCT GAGAGTCCCGCTGACACAGCTGAAGCTGAAGGACCTGGAAAAGATGGTCAACAGAGAAAGAGAA CCGAAGCTGTACGAAGCACTGAAGGCAAGACTGGAAGCACACAAGGACGACCCGGCAAAGGCAT TCGCAGAACCGTTCTACAAGTACGACAAGGCAGGAAACAGAACACAGCAGGTCAAGGCAGTCAG AGTCGAACAGGTCCAGAAGACAGGAGTCTGGGTCAGAAACCACAACGGAATCGCAGACAACGCA ACAATGGTCAGAGTAGACGTCTTCGAAAAGGGAGACAAGTACTACCTGGTCCCGATCTACTCGTG GCAGGTCGCAAAGGGAATCCTGCCGGACAGAGCAGTCGTCCAGGGAAAGGACGAAGAAGACTG GCAGCTGATCGACGACTCGTTCAACTTCAAGTTCTCGCTGCACCCGAACGACCTGGTCGAAGTCA TCACAAAGAAGGCAAGAATGTTCGGATACTTCGCATCGTGCCACAGAGGAACAGGAAACATCAAC ATCAGAATCCACGACCTGGACCACAAGATCGGAAAGAACGGAATCCTGGAAGGAATCGGAGTCAA GACAGCACTGTCGTTCCAGAAGTACCAGATCGACGAACTGGGAAAGGAAATCAGACCGTGCAGA CTGAAGAAGAGACCGCCGGTCAGATCCGGAAAGAGAACAGCAGACGGATCGGAATTCGAATCGC CGAAGAAGAAGAGAAAGGTCGAATGA	
ORF que codifica Neisseria meningitidis Cas9 (sem códons de iniciação ou parada; adequado para inclusão na sequência de codificação das proteínas de fusão)	GCAGCATTCAAGCCGAACTCGATCAACTACATCCTGGGACTGGACATCGGAATCGCATCGGTCGG ATGGGCAATGGTCGAAATCGACGAAGAAGAAAACCCGATCAGACTGATCGACCTGGGAGTCAGA GTCTTCGAAAGAGCAGAAGTCCCGAAGACAGGAGACTCGCTGGCAATGGCAAGAAGACTGGCAA GATCGGTCAGAAGACTGACAAGAAGAAGAGCACACAGACTGCTGAGAACAAGAAGACTGCTGAA GAGAGAAGGAGTCCTGCAGGCAGCAAACTTCGACGAAAACGGACTGATCAAGTCGCTGCCGAAC ACACCGTGGCAGCTGAGAGCAGCAGCACTGGACAGAAAGCTGACACCGCTGGAATGGTCGGCAG TCCTGCTGCACCTGATCAAGCACAGAGGATACCTGTCGCAGAGAAAGAACGAAGGAGAAACAGCA GACAAGGAACTGGGAGCACTGCTGAAGGGAGTCGCAGGAAACGCACACGCACTGCAGACAGGA GACTTCAGAACACCGGCAGAACTGGCACTGAACAAGTTCGAAAAGGAATCGGGACACATCAGAAA	**66**

	CCAGAGATCGGACTACTCGCACACATTCTCGAGAAAGGACCTGCAGGCAGAACTGATCCTGCTGT TCGAAAAGCAGAAGGAATTCGGAAACCCGCACGTCTCGGGAGGACTGAAGGAAGGAATCGAAAC ACTGCTGATGACACAGAGACCGGCACTGTCGGGAGACGCAGTCCAGAAGATGCTGGGACACTGC ACATTCGAACCGGCAGAACCGAAGGCAGCAAAGAACACATACACAGCAGAAAGATTCATCTGGCT GACAAAGCTGAACAACCTGAGAATCCTGGAACAGGGATCGGAAAGACCGCTGACAGACACAGAA AGAGCAACACTGATGGACGAACCGTACAGAAAGTCGAAGCTGACATACGCACAGGCAAGAAAGCT GCTGGGACTGGAAGACACAGCATTCTTCAAGGGACTGAGATACGGAAAGGACAACGCAGAAGCA TCGACACTGATGGAAATGAAGGCATACCACGCAATCTCGAGAGCACTGGAAAAGGAAGGACTGAA GGACAAGAAGTCGCCGCTGAACCTGTCGCCGGAACTGCAGGACGAAATCGGAACAGCATTCTCG CTGTTCAAGACAGACGAAGACATCACAGGAAGACTGAAGGACAGAATCCAGCCGGAAATCCTGGA AGCACTGCTGAAGCACATCTCGTTCGACAAGTTCGTCCAGATCTCGCTGAAGGCACTGAGAAGAA TCGTCCCGCTGATGGAACAGGGAAAGAGATACGACGAAGCATGCGCAGAAATCTACGGAGACCA CTACGGAAAGAAGAACACAGAAGAAAAGATCTACCTGCCGCCGATCCCGGCAGACGAAATCAGAA ACCCGGTCGTCCTGAGAGCACTGTCGCAGGCAAGAAAGGTCATCAACGGAGTCGTCAGAAGATA CGGATCGCCGGCAAGAATCCACATCGAAACAGCAAGAGAAGTCGGAAAGTCGTTCAAGGACAGA AAGGAAATCGAAAAGAGACAGGAAGAAAACAGAAAGGACAGAGAAAAGGCAGCAGCAAAGTTCA GAGAATACTTCCCGAACTTCGTCGGAGAACCGAAGTCGAAGGACATCCTGAAGCTGAGACTGTAC GAACAGCAGCACGGAAAGTGCCTGTACTCGGGAAAGGAAATCAACCTGGGAAGACTGAACGAAA AGGGATACGTCGAAATCGACCACGCACTGCCGTTCTCGAGAACATGGGACGACTCGTTCAACAAC AAGGTCCTGGTCCTGGGATCGGAAAACCAGAACAAGGGAAACCAGACACCGTACGAATACTTCAA CGGAAAGGACAACTCGAGAGAATGGCAGGAATTCAAGGCAAGAGTCGAAACATCGAGATTCCCG AGATCGAAGAAGCAGAGAATCCTGCTGCAGAAGTTCGACGAAGACGGATTCAAGGAAAGAAACCT GAACGACACAAGATACGTCAACAGATTCCTGTGCCAGTTCGTCGCAGACAGAATGAGACTGACAG GAAAGGGAAAGAAGAGAGTCTTCGCATCGAACGGACAGATCACAAACCTGCTGAGAGGATTCTGG GGACTGAGAAAGGTCAGAGCAGAAAACGACAGACACCACGCACTGGACGCAGTCGTCGTCGCAT GCTCGACAGTCGCAATGCAGCAGAAGATCACAAGATTCGTCAGATACAAGGAAATGAACGCATTC GACGGAAAGACAATCGACAAGGAAACAGGAGAAGTCCTGCACCAGAAGACACACTTCCCGCAGC CGTGGGAATTCTTCGCACAGGAAGTCATGATCAGAGTCTTCGGAAAGCCGGACGGAAAGCCGGA ATTCGAAGAAGCAGACACACTGGAAAAGCTGAGAACACTGCTGGCAGAAAAGCTGTCGTCGAGAC	

	CGGAAGCAGTCCACGAATACGTCACACCGCTGTTCGTCTCGAGAGCACCGAACAGAAAGATGTC GGGACAGGGACACATGGAAACAGTCAAGTCGGCAAAGAGACTGGACGAAGGAGTCTCGGTCCTG AGAGTCCCGCTGACACAGCTGAAGCTGAAGGACCTGGAAAAGATGGTCAACAGAGAAAGAGAAC CGAAGCTGTACGAAGCACTGAAGGCAAGACTGGAAGCACACAAGGACGACCCGGCAAAGGCATT CGCAGAACCGTTCTACAAGTACGACAAGGCAGGAAACAGAACACAGCAGGTCAAGGCAGTCAGA GTCGAACAGGTCCAGAAGACAGGAGTCTGGGTCAGAAACCACAACGGAATCGCAGACAACGCAA CAATGGTCAGAGTAGACGTCTTCGAAAAGGGAGACAAGTACTACCTGGTCCCGATCTACTCGTGG CAGGTCGCAAAGGGAATCCTGCCGGACAGAGCAGTCGTCCAGGGAAAGGACGAAGAAGACTGG CAGCTGATCGACGACTCGTTCAACTTCAAGTTCTCGCTGCACCCGAACGACCTGGTCGAAGTCAT CACAAAGAAGGCAAGAATGTTCGGATACTTCGCATCGTGCCACAGAGGAACAGGAAACATCAACA TCAGAATCCACGACCTGGACCACAAGATCGGAAAGAACGGAATCCTGGAAGGAATCGGAGTCAA GACAGCACTGTCGTTCCAGAAGTACCAGATCGACGAACTGGGAAAGGAAATCAGACCGTGCAGA CTGAAGAAGAGACCGCCGGTCAGATCCGGAAAGAGAACAGCAGACGGATCGGAATTCGAATCGC CGAAGAAGAAGAGAAAGGTCGAA	
Transcrição compreendendo SEQ ID NO: 65 (codificando Neisseria meningitidis Cas9)	GGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGATCCGCCACCATGGCAGCATTCAAGC CGAACTCGATCAACTACATCCTGGGACTGGACATCGGAATCGCATCGGTCGGATGGGCAATGGTC GAAATCGACGAAGAAGAAAACCCGATCAGACTGATCGACCTGGGAGTCAGAGTCTTCGAAAGAGC AGAAGTCCCGAAGACAGGAGACTCGCTGGCAATGGCAAGAAGACTGGCAAGATCGGTCAGAAGA CTGACAAGAAGAAGAGCACACAGACTGCTGAGAACAAGAAGACTGCTGAAGAGAGAAGGAGTCC TGCAGGCAGCAAACTTCGACGAAAACGGACTGATCAAGTCGCTGCCGAACACACCGTGGCAGCT GAGAGCAGCAGCACTGGACAGAAAGCTGACACCGCTGGAATGGTCGGCAGTCCTGCTGCACCTG ATCAAGCACAGAGGATACCTGTCGCAGAGAAAGAACGAAGGAGAAACAGCAGACAAGGAACTGG GAGCACTGCTGAAGGGAGTCGCAGGAAACGCACACGCACTGCAGACAGGAGACTTCAGAACACC GGCAGAACTGGCACTGAACAAGTTCGAAAAGGAATCGGGACACATCAGAAACCAGAGATCGGACT ACTCGCACACATTCTCGAGAAAGGACCTGCAGGCAGAACTGATCCTGCTGTTCGAAAAGCAGAAG GAATTCGGAAACCCGCACGTCTCGGGAGGACTGAAGGAAGGAATCGAAACACTGCTGATGACAC AGAGACCGGCACTGTCGGGAGACGCAGTCCAGAAGATGCTGGGACACTGCACATTCGAACCGGC AGAACCGAAGGCAGCAAAGAACACATACACAGCAGAAAGATTCATCTGGCTGACAAAGCTGAACA ACCTGAGAATCCTGGAACAGGGATCGGAAAGACCGCTGACAGACACAGAAAGAGCAACACTGAT	**67**

	GGACGAACCGTACAGAAAGTCGAAGCTGACATACGCACAGGCAAGAAAGCTGCTGGGACTGGAA GACACAGCATTCTTCAAGGGACTGAGATACGGAAAGGACAACGCAGAAGCATCGACACTGATGGA AATGAAGGCATACCACGCAATCTCGAGAGCACTGGAAAAGGAAGGACTGAAGGACAAGAAGTCG CCGCTGAACCTGTCGCCGGAACTGCAGGACGAAATCGGAACAGCATTCTCGCTGTTCAAGACAGA CGAAGACATCACAGGAAGACTGAAGGACAGAATCCAGCCGGAAATCCTGGAAGCACTGCTGAAG CACATCTCGTTCGACAAGTTCGTCCAGATCTCGCTGAAGGCACTGAGAAGAATCGTCCCGCTGAT GGAACAGGGAAAGAGATACGACGAAGCATGCGCAGAAATCTACGGAGACCACTACGGAAAGAAG AACACAGAAGAAAAGATCTACCTGCCGCCGATCCCGGCAGACGAAATCAGAAACCCGGTCGTCCT GAGAGCACTGTCGCAGGCAAGAAAGGTCATCAACGGAGTCGTCAGAAGATACGGATCGCCGGCA AGAATCCACATCGAAACAGCAAGAGAAGTCGGAAAGTCGTTCAAGGACAGAAAGGAAATCGAAAA GAGACAGGAAGAAAACAGAAAGGACAGAGAAAAGGCAGCAGCAAAGTTCAGAGAATACTTCCCGA ACTTCGTCGGAGAACCGAAGTCGAAGGACATCCTGAAGCTGAGACTGTACGAACAGCAGCACGG AAAGTGCCTGTACTCGGGAAAGGAAATCAACCTGGGAAGACTGAACGAAAAGGGATACGTCGAAA TCGACCACGCACTGCCGTTCTCGAGAACATGGGACGACTCGTTCAACAACAAGGTCCTGGTCCTG GGATCGGAAAACCAGAACAAGGGAAACCAGACACCGTACGAATACTTCAACGGAAAGGACAACTC GAGAGAATGGCAGGAATTCAAGGCAAGAGTCGAAACATCGAGATTCCCGAGATCGAAGAAGCAG AGAATCCTGCTGCAGAAGTTCGACGAAGACGGATTCAAGGAAAGAAACCTGAACGACACAAGATA CGTCAACAGATTCCTGTGCCAGTTCGTCGCAGACAGAATGAGACTGACAGGAAAGGGAAAGAAGA GAGTCTTCGCATCGAACGGACAGATCACAAACCTGCTGAGAGGATTCTGGGGACTGAGAAAGGTC AGAGCAGAAAACGACAGACACCACGCACTGGACGCAGTCGTCGTCGCATGCTCGACAGTCGCAA TGCAGCAGAAGATCACAAGATTCGTCAGATACAAGGAAATGAACGCATTCGACGGAAAGACAATC GACAAGGAAACAGGAGAAGTCCTGCACCAGAAGACACACTTCCCGCAGCCGTGGGAATTCTTCG CACAGGAAGTCATGATCAGAGTCTTCGGAAAGCCGGACGGAAAGCCGGAATTCGAAGAAGCAGA CACACTGGAAAAGCTGAGAACACTGCTGGCAGAAAAGCTGTCGTCGAGACCGGAAGCAGTCCAC GAATACGTCACACCGCTGTTCGTCTCGAGAGCACCGAACAGAAAGATGTCGGGACAGGGACACA TGGAAACAGTCAAGTCGGCAAAGAGACTGGACGAAGGAGTCTCGGTCCTGAGAGTCCCGCTGAC ACAGCTGAAGCTGAAGGACCTGGAAAAGATGGTCAACAGAGAAAGAGAACCGAAGCTGTACGAA GCACTGAAGGCAAGACTGGAAGCACACAAGGACGACCCGGCAAAGGCATTCGCAGAACCGTTCT ACAAGTACGACAAGGCAGGAAACAGAACACAGCAGGTCAAGGCAGTCAGAGTCGAACAGGTCCA	

	GAAGACAGGAGTCTGGGTCAGAAACCACAACGGAATCGCAGACAACGCAACAATGGTCAGAGTA GACGTCTTCGAAAAGGGAGACAAGTACTACCTGGTCCCGATCTACTCGTGGCAGGTCGCAAAGG GAATCCTGCCGGACAGAGCAGTCGTCCAGGGAAAGGACGAAGAAGACTGGCAGCTGATCGACGA CTCGTTCAACTTCAAGTTCTCGCTGCACCCGAACGACCTGGTCGAAGTCATCACAAAGAAGGCAA GAATGTTCGGATACTTCGCATCGTGCCACAGAGGAACAGGAAACATCAACATCAGAATCCACGAC CTGGACCACAAGATCGGAAAGAACGGAATCCTGGAAGGAATCGGAGTCAAGACAGCACTGTCGTT CCAGAAGTACCAGATCGACGAACTGGGAAAGGAAATCAGACCGTGCAGACTGAAGAAGAGACCG CCGGTCAGATCCGGAAAGAGAACAGCAGACGGATCGGAATTCGAATCGCCGAAGAAGAAGAGAA AGGTCGAATGATAGCTAGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGT GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTG CCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATT CTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGC ATGCTGGGGATGCGGTGGGCTCTATGG	
Sequência de aminoácidos de Neisseria meningitidis Cas9	MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRR LTRRRAHRLLRTRRLLKREGVLQAANFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRG YLSQRKNEGETADKELGALLKGVAGNAHALQTGDFRTPAELALNKFEKESGHIRNQRSDYSHTFSRKD LQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAER FIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEAS TLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISF DKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVI NGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLR LYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFN GKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKG KKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTID KETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSSRPEAVHEYVTP LFVSRAPNRKMSGQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEA HKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYY LVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTG NINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPPVRSGKRTADGSEFESPKKK	**68**

	RKVE	
RNA guia G390	mG*mC*mC*GAGUCUGGAGAGCUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**69**
RNA guia G502	mA*mC*mA*CAAAUACCAGUCCAGCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**70**
RNA guia G509	mA*mA*mA*GUUCUAGAUGCCGUCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**71**
RNA guia G534	mA*mC*mG*CAAAUAUCAGUCCAGCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**72**
sgRNA compreendendo padrão de modificação	mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**74**
RNA guia G562	mC*mC*mA*AUAUCAGGAGACUAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	**75**

*** = Ligação PS; m = nucleotídeo 2'-O-Me; N =** qualquer nucleotídeo natural ou não natural

Sequência GFP:

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCC
CGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAG
CCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTT
AACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGT
GTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC
ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG
CCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGC
GATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAAC
GACGGCCAGTGAATTCTAATACGACTCACTATAGGGTCCCGCAGTCGG
CGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC
GGATCCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC
CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGT
GTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGA
AGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCG
TGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG
TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCC
GCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG
CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC
AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCG
GCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGT
CCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGC
TGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAATAGGAATTATGCAGTCTAGCCATCACATTTAAAAGCATCTCA
GCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAATAGCTTATTCAT
CTCTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAAACAT
AAATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAA
TGGAAAGAACCTCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAATCTAGACTTAAGCTTGATGAGCTCTAGCTTGGCGT

AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT
TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAAC
GCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG
CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAA
AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG
TGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC
TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAG
ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC
GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTA
CCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTT
CATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA
GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACG
CTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATT
AATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA
TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA

TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGC
AGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTG
TGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACA
TAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCT
GGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGG
GCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTA
TTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT
GCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAA
ATAGGCGTATCACGAGGCCCTTTCGTCG

REIVINDICAÇÕES

1. Método de administração de um mRNA a uma célula-tronco hematopoiética e/ou célula progenitora (HSPC) ou a uma população de HSPC, **caracterizado** pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo o mRNA, um lipídio com amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG;

b. colocar a HSPC ou população de HSPC em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a HSPC ou a população de HSPC *in vitro*;

administrando, desse modo, o mRNA à HSPC ou à população de HSPC.

2. Método de administração de um mRNA a uma HSPC, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo o mRNA e um lipídio com amina;

b. colocar a célula em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a HSPC *in vitro*;

administrando, desse modo, o mRNA à HSPC.

3. Método de administração de um mRNA a uma célula-tronco ou a uma população de células-tronco, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo o mRNA;

b. colocar a população de células-tronco em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a população de células-tronco *in vitro*;

administrando, desse modo, o mRNA à população de células-tronco.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, **caracterizado** pelo fato de que o mRNA codifica uma nuclease Cas.

5. Método de introdução de um mRNA da nuclease Cas e um gRNA em uma HSPC, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo o mRNA da nuclease Cas, um gRNA, um lipídio com amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG;

b. colocar a HSPC em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a HSPC;

introduzindo, desse modo, o mRNA da nuclease Cas e o gRNA na HSPC.

6. Método de produção de uma HSPC geneticamente manipulada *in vitro*, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo um mRNA da nuclease Cas, um gRNA, um lipídio com amina, um lipídio auxiliar, um lipídio neutro e um lipídio com PEG;

b. colocar a HSPC em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a HSPC *in vitro*;

produzindo, desse modo, uma HSPC geneticamente manipulada.

7. Método de introdução de um mRNA da nuclease Cas e um gRNA em uma célula-tronco, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo o mRNA da nuclease Cas, gRNA e um lipídio com amina;

b. colocar a célula-tronco em contato com a composição de

LNP pré-incubada *in vitro*; e

c. cultivar a célula-tronco;

introduzindo, desse modo, o mRNA da nuclease Cas e o gRNA na célula-tronco.

8. Método de produção de uma célula-tronco geneticamente manipulada, como uma HSPC *in vitro*, caracterizado pelo fato de que compreende:

a. pré-incubar um fator sérico com uma composição de LNP compreendendo um mRNA da nuclease Cas, gRNA e um lipídio biodegradável;

b. colocar a célula em contato com a composição de LNP pré-incubada *in vitro*; e

c. cultivar a célula *in vitro*;

produzindo, desse modo, uma célula-tronco geneticamente modificada, como uma HSPC.

9. Método, de acordo com a reivindicação 4, **caracterizado** pelo fato de que a composição de LNP compreende ainda um gRNA.

10. Método, de acordo com qualquer uma das reivindicações 4 a 9, **caracterizado** pelo fato de que a nuclease Cas é uma nuclease Cas de Classe 2.

11. Método, de acordo com a reivindicação 10, **caracterizado** pelo fato de que a nuclease Cas de Classe 2 é uma nuclease Cas9.

12. Método, de acordo com a reivindicação 11, **caracterizado** pelo fato de que a nuclease Cas9 é uma Cas9 de *S. pyogenes*.

13. Método, de acordo com a reivindicação 10, **caracterizado** pelo fato de que a nuclease Cas de Classe 2 é uma nuclease Cpf1.

14. Método, de acordo com qualquer uma das reivindicações 5 a 13, **caracterizado** pelo fato de que o gRNA é um RNA guia

duplo (dgRNA).

15. Método, de acordo com qualquer uma das reivindicações 5 a 13, **caracterizado** pelo fato de que o gRNA é um RNA guia único (sgRNA).

16. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda uma etapa de lavagem após a etapa de contato.

17. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a etapa de contato dura entre cerca de 1 minuto e cerca de 72 horas.

18. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a etapa de contato dura entre cerca de 1 minuto e cerca de 24 horas.

19. Método, de acordo com a reivindicação 17 ou 18, **caracterizado** pelo fato de que a etapa de contato dura entre cerca de 2 horas e cerca de 24 horas.

20. Método, de acordo com qualquer uma das reivindicações 17 a 19, **caracterizado** pelo fato de que a etapa de contato dura entre cerca de 4 horas e cerca de 12 horas.

21. Método, de acordo com qualquer uma das reivindicações 17 a 20, **caracterizado** pelo fato de que a etapa de contato dura de cerca entre 6 horas e cerca de 12 horas.

22. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a sobrevida de células pós-transfecção é de pelo menos 60%.

23. Método, de acordo com a reivindicação 22, **caracterizado** pelo fato de que a sobrevida de células pós-transfecção é de pelo menos 70%.

24. Método, de acordo com a reivindicação 22, **caracterizado** pelo fato de que a sobrevida de células pós-transfecção é de pe-

lo menos 80%.

25. Método, de acordo com a reivindicação 22, **caracterizado** pelo fato de que a sobrevida de células pós-transfecção é de pelo menos 90%.

26. Método, de acordo com a reivindicação 22, **caracterizado** pelo fato de que a sobrevida de células pós-transfecção é de pelo menos 95%.

27. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda a pré-incubação do fator sérico e da composição de LNP por cerca de 30 segundos ou durante a noite.

28. Método, de acordo com a reivindicação 27, **caracterizado** pelo fato de que compreende a pré-incubação por cerca de 1 minuto a 1 hora.

29. Método, de acordo com a reivindicação 27, **caracterizado** pelo fato de que compreende a pré-incubação por cerca de 1-30 minutos.

30. Método, de acordo com a reivindicação 27, **caracterizado** pelo fato de que compreende a pré-incubação por cerca de 1-10 minutos.

31. Método, de acordo com a reivindicação 27, **caracterizado** pelo fato de que compreende a pré-incubação por cerca de 5 minutos.

32. Método, de acordo com a reivindicação 27 ou 31, **caracterizado** pelo fato de que compreende a pré-incubação por 5 minutos ± 2 minutos.

33. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a pré-incubação ocorre a cerca de 4 °C.

34. Método, de acordo com qualquer uma das reivindica-

ções precedentes, **caracterizado** pelo fato de que a pré-incubação ocorre a cerca de 25 °C.

35. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a pré-incubação ocorre a cerca de 37 °C.

36. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a etapa de pré-incubação compreende um tampão.

37. Método, de acordo com a reivindicação 36, **caracterizado** pelo fato de que o tampão compreende ou consiste em um meio de cultura de HSPC.

38. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a composição de LNP é pré-incubada com soro.

39. Método, de acordo com a reivindicação 38, **caracterizado** pelo fato de que o soro é soro de mamífero, camundongo, primata ou humano.

40. Método, de acordo com qualquer uma das reivindicações 1 a 37, **caracterizado** pelo fato de que a composição de LNP é pré-incubada com um fator sérico isolado.

41. Método, de acordo com a reivindicação 40, **caracterizado** pelo fato de que o fator sérico é uma ApoE.

42. Método, de acordo com a reivindicação 40, **caracterizado** pelo fato de que o fator sérico é escolhido dentre ApoE2, ApoE3 e ApoE4.

43. Método, de acordo com qualquer uma das reivindicações 40 a 42, **caracterizado** pelo fato de que a ApoE é uma proteína humana recombinante.

44. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a etapa de cultura

compreende a expansão da célula-tronco, HSPC ou população de HSPC em um tampão de cultura de HSPC.

45. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda a troca do meio de cultura entre as etapas de contato e cultura.

46. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a etapa de cultura compreende um expansor de células-tronco.

47. Método, de acordo com qualquer uma das reivindicações 1 a 2, 4 a 6 ou 8 a 46, **caracterizado** pelo fato de que a HSPC é uma célula-tronco hematopoiética (HSC).

48. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a célula-tronco, a HSPC ou a população de HSPC é uma célula ou amostra humana.

49. Método, de acordo com qualquer uma das reivindicações 5 a 48, **caracterizado** pelo fato de que o mRNA e o ácido nucleico de RNA guia são formulados em uma única composição de LNP.

50. Método, de acordo com qualquer uma das reivindicações 5 a 48, **caracterizado** pelo fato de que o mRNA e o gRNA são coencapsulados na composição de LNP.

51. Método, de acordo com qualquer uma das reivindicações 5 a 48, **caracterizado** pelo fato de que o mRNA e o gRNA são encapsulados separadamente nas LNPs.

52. Método, de acordo com qualquer uma das reivindicações 5 a 48, **caracterizado** pelo fato de que o mRNA é formulado em uma primeira composição de LNP e o ácido nucleico de RNA guia é formulado em uma segunda composição de LNP.R

53. Método, de acordo com a reivindicação 52, **caracterizado** pelo fato de que a primeira e a segunda composições de LNP são administradas simultaneamente.

54. Método, de acordo com a reivindicação 52, **caracterizado** pelo fato de que a primeira e a segunda composições de LNP são administradas sequencialmente.R

55. Método, de acordo com qualquer uma das reivindicações 52 a 54, **caracterizado** pelo fato de que a primeira e a segunda composições de LNP são combinadas antes da etapa de pré-incubação.

56. Método, de acordo com qualquer uma das reivindicações 52 a 54, **caracterizado** pelo fato de que a primeira e a segunda composições de LNP são pré-incubadas separadamente.

57. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda a introdução de um ácido nucleico modelo na célula.

58. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a composição de LNP compreende: um componente de RNA e um componente lipídico, em que o componente lipídico compreende um lipídio com amina, um lipídio neutro, um lipídio auxiliar e um lipídio do tipo stealth; e em que a razão N/P é cerca de 1-10.

59. Método, de acordo com a reivindicação 58, **caracterizado** pelo fato de que o componente lipídico compreende o Lipídio A ou seu análogo de acetal.

60. Método, de acordo com a reivindicação 58, **caracterizado** pelo fato de que o componente lipídico compreende:

cerca de 40-60% em mol de lipídio com amina;R

cerca de 5-15% em mol de lipídio neutro; e

cerca de 1,5-10% em mol de lipídio com PEG,

em que o restante do componente lipídico é lipídio auxiliar, e

em que a razão N/P da composição de LNP é cerca de 3-

10.

61. Método, de acordo com a reivindicação 58, **caracterizado** pelo fato de que o componente lipídico compreende:

cerca de 50-60% em mol de lipídio com amina;R

cerca de 8-10% em mol de lipídio neutro; e

cerca de 2,5-4% em mol de lipídio com PEG,

em que o restante do componente lipídico é lipídio auxiliar, e

em que a razão N/P da composição de LNP é cerca de 3-8.

62. Método, de acordo com a reivindicação 56, **caracterizado** pelo fato de que o componente lipídico compreende:

cerca de 50-60% em mol de lipídio com amina;R

cerca de 5-15% em mol de DSPC; e

cerca de 2,5-4% em mol de lipídio com PEG,

em que o restante do componente lipídico é colesterol e

em que a razão N/P da composição de LNP é cerca de 3-8.

63. Método, de acordo com a reivindicação 58, **caracterizado** pelo fato de que o componente lipídico compreende:

48-53% em mol de Lipídio A;R

cerca de 8-10% em mol de DSPC; e

1,5-10% em mol de lipídio com PEG,

em que o restante do componente lipídico é colesterol e

em que a razão N/P da composição de LNP é de 3-8 ±0,2.

64. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que o RNA é um RNA modificado.

65. Método, de acordo com a reivindicação 64, **caracterizado** pelo fato de que o RNA modificado é um mRNA modificado.

66. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que o RNA compreende

um quadro de leitura aberto que codifica um agente de ligação ao DNA guiado por RNA, em que o quadro de leitura aberto tem um teor de uridina que varia de seu teor mínimo de uridina a 150% do teor de uridina mínimo.

67. Composição, como definida em qualquer uma das reivindicações precedentes, **caracterizada** pelo fato de que o RNA compreende um quadro de leitura aberto que codifica um agente de ligação ao DNA guiado por RNA, em que o quadro de leitura aberto tem um teor de dinucleotídeo de uridina que varia de seu teor de dinucleotídeo de uridina mínimo a 150% do teor de dinucleotídeo de uridina mínimo.

68. Composição, de acordo com qualquer uma das reivindicações precedentes, **caracterizada** pelo fato de que o RNA compreende uma sequência com pelo menos 90% de identidade com qualquer uma das SEQ ID NO: 1, 4, 10, 14, 15, 17, 18, 20, 21, 23, 24, 26, 27, 29, 30, 50, 52, 54, 65 ou 66, em que o mRNA compreende um quadro de leitura aberto que codifica um agente de ligação ao DNA guiado por RNA.

69. Método, de acordo com qualquer uma das reivindicações 5 a 68, **caracterizado** pelo fato de que o gRNA é um gRNA modificado.

70. Método, de acordo com a reivindicação 69, **caracterizado** pelo fato de que o gRNA compreende uma modificação escolhida dentre um nucleotídeo modificado por 2'-O-metila (2'-O-Me), uma ligação fosforotioato (PS) entre nucleotídeos; e um nucleotídeo modificado por 2'-fluoro (2'-F).

71. Método, de acordo com a reivindicação 69 ou 70, **caracterizado** pelo fato de que o gRNA compreende uma modificação em um ou mais dos cinco primeiros nucleotídeos na extremidade 5’.

72. Método, de acordo com qualquer uma das reivindica-

ções 69 a 71, **caracterizado** pelo fato de que o gRNA compreende uma modificação em um ou mais dos cinco últimos nucleotídeos na extremidade 3'.

73. Método, de acordo com qualquer uma das reivindicações 69 a 72, **caracterizado** pelo fato de que o gRNA compreende ligações PS entre os quatro primeiros nucleotídeos.

74. Método, de acordo com qualquer uma das reivindicações 69 a 73, **caracterizado** pelo fato de que o gRNA compreende ligações PS entre os quatro últimos nucleotídeos.

75. Método, de acordo com qualquer uma das reivindicações 69 a 74, **caracterizado** pelo fato de que compreende ainda nucleotídeos modificados por 2'-O-Me nos três primeiros nucleotídeos na extremidade 5'.

76. Método, de acordo com qualquer uma das reivindicações 69 a 75, **caracterizado** pelo fato de que compreende ainda nucleotídeos modificados por 2'-O-Me nos três últimos nucleotídeos na extremidade 3'.

77. Método, de acordo com as reivindicações 1 a 2, 4 a 6 ou 8 a 76, **caracterizado** pelo fato de que a HSPC ou a população de HSPC é CD34+.

78. Método, de acordo com as reivindicações 1 a 2, 4 a 6 ou 8 a 77, **caracterizado** pelo fato de que a HSPC ou a população de HSPC é CD34+CD90+.

79. Célula-tronco ou população de células-tronco manipulada, **caracterizada** pelo fato de que é produzida pelo método como definido em qualquer uma das reivindicações precedentes.

80. HSPC ou população de HSPC manipulada, **caracterizada** pelo fato de que é produzida pelo método como definido em qualquer uma das reivindicações precedentes.

81. HSPC ou população de HSPC, de acordo com a reivin-

dicação 78, **caracterizada** pelo fato de que a HSPC manipulada reside dentro de um tecido ou órgão, por exemplo, medula óssea, sangue ou outro tecido dentro de um paciente, por exemplo, após transplante de uma HSPC manipulada.

82. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a célula-tronco, a HSPC ou a população de HSPC é autóloga em relação a um paciente ao qual será administrada a célula.

83. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a célula-tronco, a HSPC ou a população de HSPC é alogênica em relação a um paciente ao qual será administrada a referida célula.

84. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda a realização da edição de genes por CRISPR-Cas na célula-tronco, na HSPC ou na população de HSPC.

85. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que compreende ainda a detecção da edição de genes na célula-tronco, na HSPC ou na população de HSPC.

86. Método, de acordo com a reivindicação 84 ou 85, **caracterizado** pelo fato de que a edição de genes é medida como porcentagem de edição ou porcentagem de modificação do DNA.

87. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de edição é de pelo menos 40%.

88. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de edição é de pelo menos 60%.

89. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de edição é de pelo menos 70%.

90. Método, de acordo com a reivindicação 86, **caracteri-**

zado pelo fato de que a porcentagem de edição é de pelo menos 80%.

91. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de edição é de pelo menos 90%.

92. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de edição é de pelo menos 95%.

93. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 40%.

94. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 60%.

95. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 70%.

96. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 80%.

97. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 90%.

98. Método, de acordo com a reivindicação 86, **caracterizado** pelo fato de que a porcentagem de modificação do DNA é de pelo menos 95%.

99. Método, de acordo com qualquer uma das reivindicações precedentes, **caracterizado** pelo fato de que a célula-tronco, a HSPC ou a população de HSPC é de uma amostra de medula óssea.

Fig. 1A

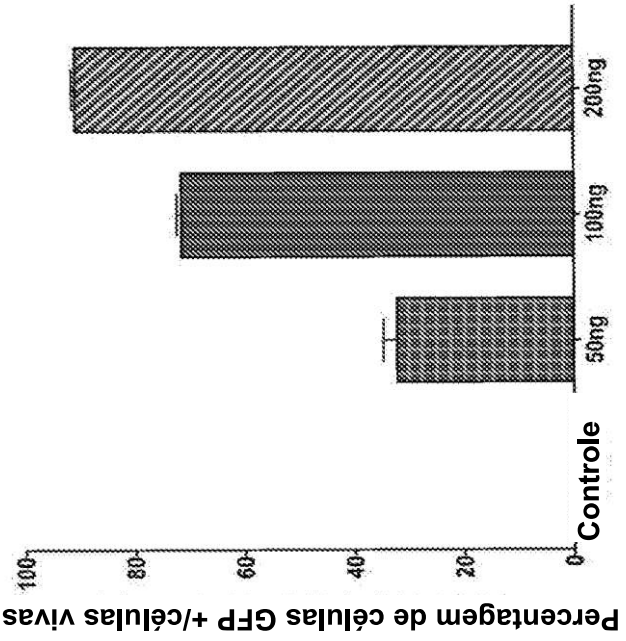

Percentagem de células GFP +/células vivas
0
20
40
60
80
100
Controle
50ng
100ng
200ng

Fig. 1B

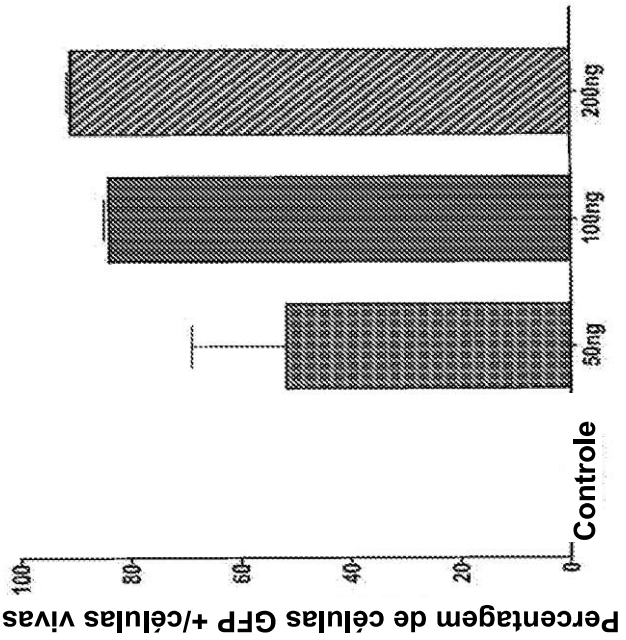

Percentagem de células GFP +/células vivas
0
20
40
60
80
100
Controle
50ng
100ng
200ng

Fig. 2A

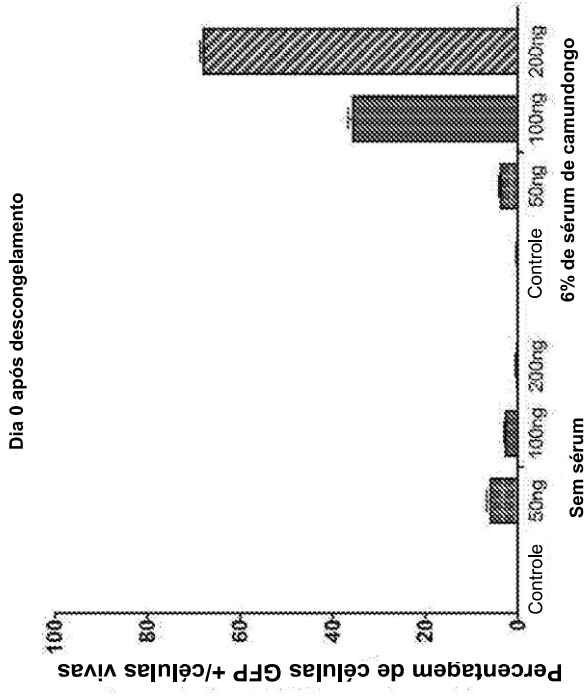

Dia 0 após descongelamento
Percentagem de células GFP +/células vivas
100
80
60
40
20
0
Controle
50ng
100ng
200ng
Controle
50ng
100ng
200ng
Sem sérum
6% de sérum de camundongo

Fig. 2B

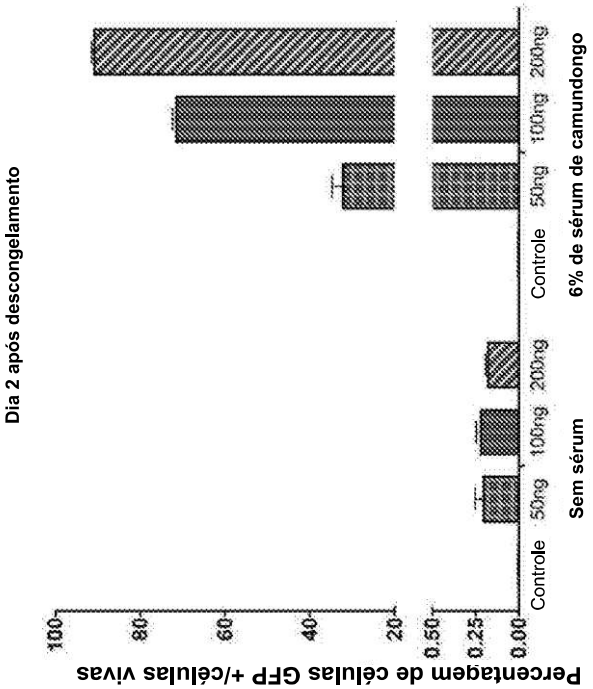

Dia 2 após descongelamento
Percentagem de células GFP +/células vivas
100
80
60
40
20
0.50
0.25
0.00
Controle
50ng
100ng
200ng
Controle
50ng
100ng
200ng
Sem sérum
6% de sérum de camundongo

Fig. 3A

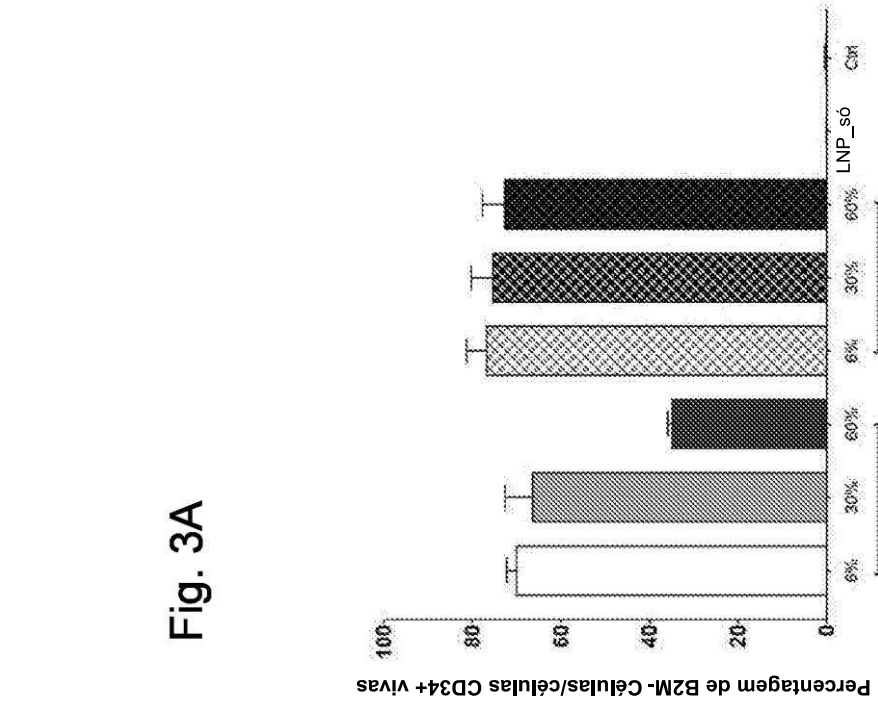
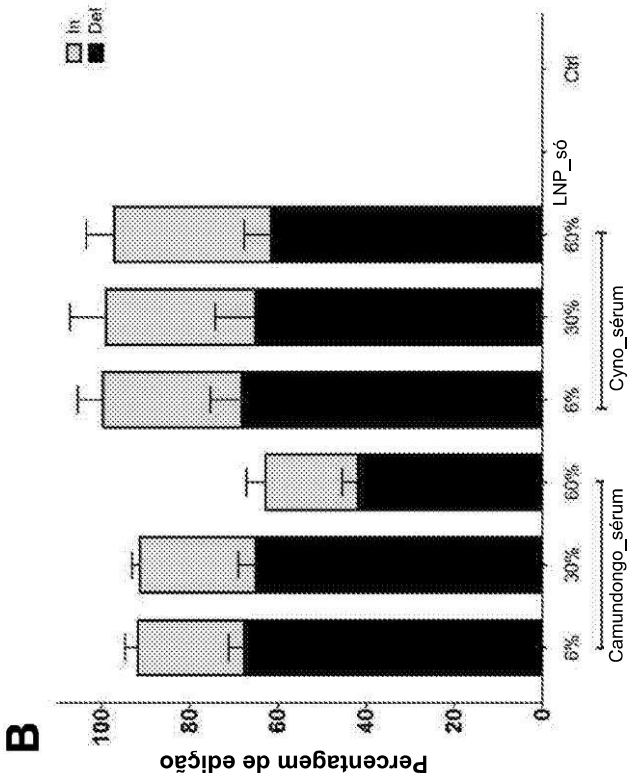

Percentagem de B2M- Células/células CD34+ vivas
100
80
60
40
20
0
6%
30%
60%
6%
30%
60%
LNP_só
Ctrl
Camundongo_sérum
Cyno_sérum

Fig. 3B


B
In
Del
Percentagem de edição
100
80
60
40
20
0
6%
30%
60%
6%
30%
60%
LNP_só
Ctrl
Camundongo_sérum
Cyno_sérum

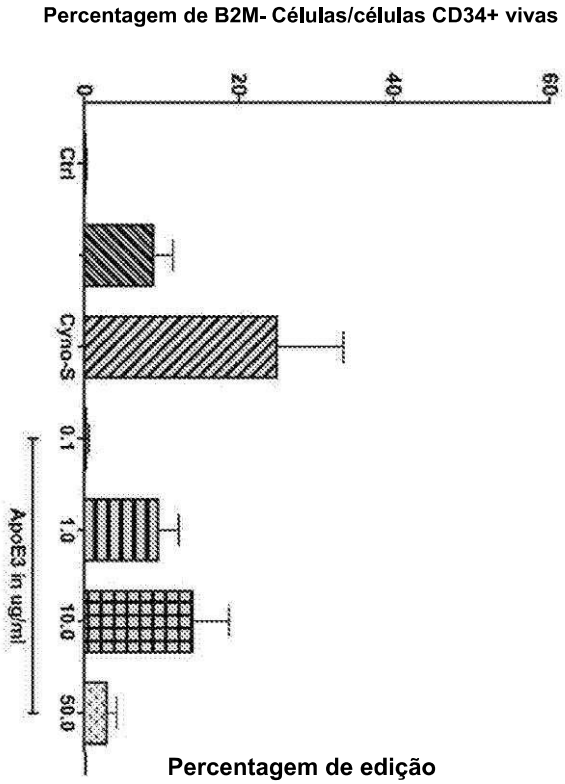

Percentagem de B2M- Células/células CD34+ vivas
0
20
40
60
Ctrl
Cyno-S
0.1
1.0
10.0
50.0
ApoE3 in ug/ml

Fig. 4A

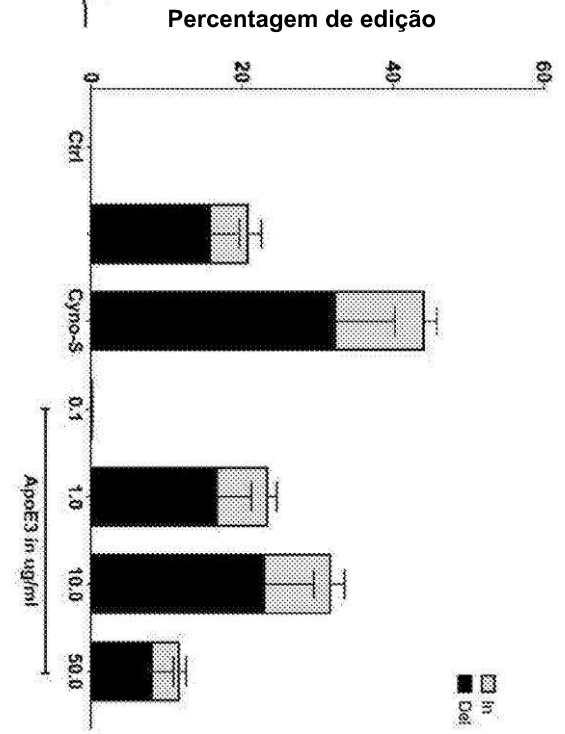

Percentagem de edição
0
20
40
60
Ctrl
Cyno-S
0.1
1.0
10.0
50.0
ApoE3 in ug/ml
In
Del

Fig. 4B

Fig. 5

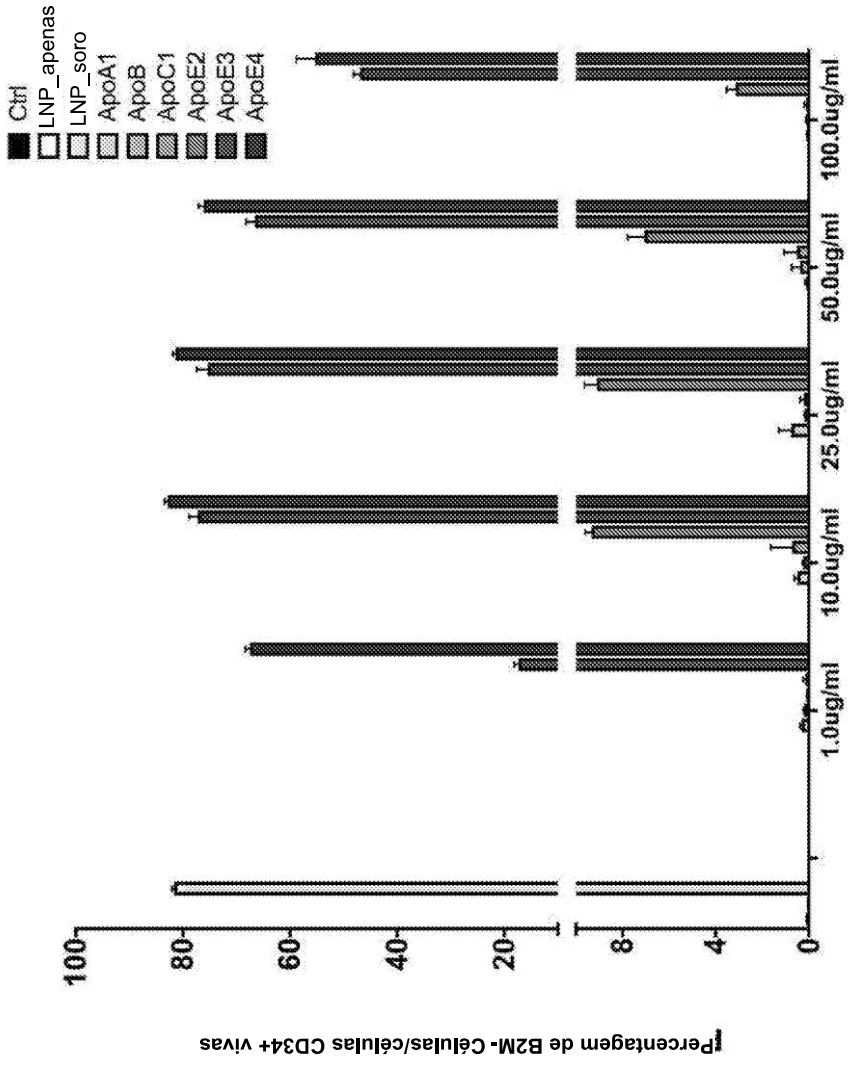

Ctrl
LNP_apenas
LNP_soro
ApoA1
ApoB
ApoC1
ApoE2
ApoE3
ApoE4
Percentagem de B2M- Células/células CD34+ vivas
100
80
60
40
20
8
4
0
1.0ug/ml
10.0ug/ml
25.0ug/ml
50.0ug/ml
100.0ug/ml

Fig. 6A

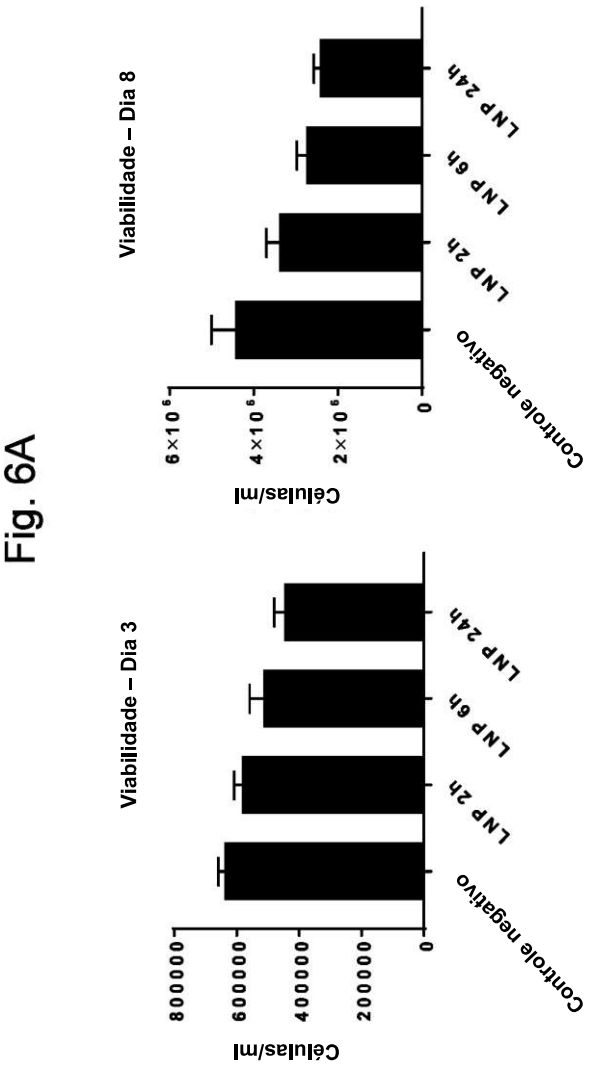
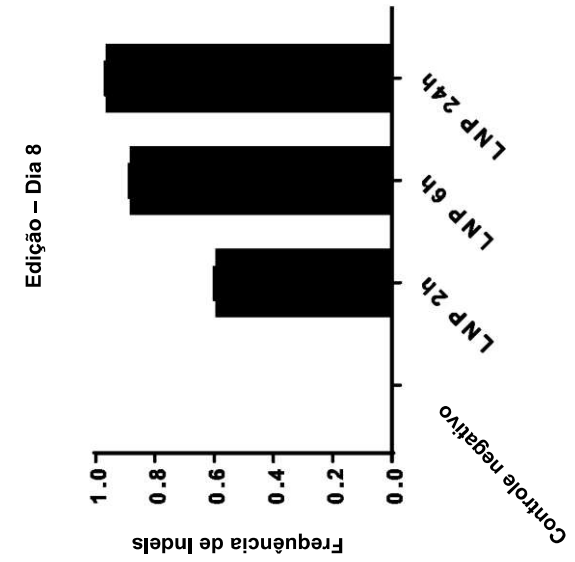

Viabilidade – Dia 8
Células/ml
6×10⁶
4×10⁶
2×10⁶
0
Controle negativo
LNP 2h
LNP 6h
LNP 24h
Viabilidade – Dia 3
Células/ml
800000
600000
400000
200000
0
Controle negativo
LNP 2h
LNP 6h
LNP 24h

Fig. 6B


Edição – Dia 8
Frequência de Indels
1.0
0.8
0.6
0.4
0.2
0.0
Controle negativo
LNP 2h
LNP 6h
LNP 24h

RESUMO

Patente de Invenção: **"MÉTODO IN VITRO DE ADMINISTRAÇÃO DE MRNA USANDO NANOPARTÍCULAS DE LIPÍDIO".**

A presente invenção refere-se a composições e métodos para a introdução de um mRNA em células-tronco, como HSPCs, e para a administração de componentes de edição de genes a essas células *in vitro*. Por exemplo, a divulgação se refere à modificação de uma sequência gênica usando um complexo de CRISPR-Cas9 em HSPCs, e métodos e sistemas de administração para atingir tal modificação gênica em HSPCs.