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

(21) BR 112021007343-4 A2

(22) Data do Depósito: 18/10/2019

(43) Data da Publicação Nacional: 03/08/2021

(54) Título: COMPOSIÇÕES E MÉTODOS PARA EXPRESSÃO DE TRANSGENE A PARTIR DE UM LÓCUS DE ALBUMINA

(51) Int. Cl.: C12N 15/113; A61K 48/00; C12N 15/90.

(30) Prioridade Unionista: 18/10/2018 US 62/747,402; 29/04/2019 US 62/840,346.

(71) Depositante(es): INTELLIA THERAPEUTICS, INC.; REGENERON PHARMACEUTICALS, INC..

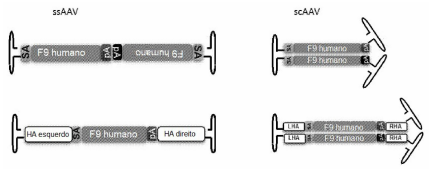
(72) Inventor(es): JONATHAN DOUGLAS FINN; HON-REN HUANG; MOITRI ROY; KEHDIH LAI; RACHEL SATTLER; CHRISTOS KYRATSOUS; CHENG WANG.

(86) Pedido PCT: PCT US2019057086 de 18/10/2019

(87) Publicação PCT: WO 2020/082042 de 23/04/2020

(85) Data da Fase Nacional: 16/04/2021

(57) Resumo: COMPOSIÇÕES E MÉTODOS PARA EXPRESSÃO DE TRANSGENE A PARTIR DE UM LÓCUS DE ALBUMINA. Trata-se de métodos para editar, por exemplo, introduzir um transgene heterólogo, no gene de albumina humano (por exemplo, no íntron 1).



“COMPOSIÇÕES E MÉTODOS PARA EXPRESSÃO DE TRANSGENE A PARTIR DE UM LÓCUS DE ALBUMINA”

[001] Este pedido reivindica o benefício de prioridade do Pedido Provisório US nº 62/747.402, depositado em 18 de outubro de 2018 e Pedido Provisório US nº 62/840.346, depositado em 29 de abril de 2019. Os relatórios descritivos de cada um dos pedidos supracitados estão incorporados no presente documento a título de referência em suas totalidades.

[002] A edição de genoma em abordagens de terapia de gene surge da ideia de que a introdução exógena do material genético ausente ou comprometido de outro modo pode corrigir uma doença genética. A terapia de gene tem sido reconhecida há muito tempo por seu grande potencial em como os praticantes abordam e tratam doenças humanas. Em vez de depender em fármacos ou cirurgia, os pacientes com fatores genéticos subjacentes podem ser tratados alvejando-se diretamente a causa subjacente. Adicionalmente, alvejando-se a causa genética, a terapia de gene pode fornecer o potencial para curar de modo eficaz os pacientes. Entretanto, as abordagens clínicas de terapia de gene ainda necessitam de melhora em diversos aspectos.

[003] São fornecidas no presente documento composições e métodos úteis para inserir e expressar um gene heterólogo (exógeno) em um lócus genômico, como um sítio de porto seguro, de uma célula hospedeira. Diversos lócus de porto seguro têm sido descritos, incluindo CCR5, HPRT, AAVS1, Rosa e albumina. Conforme descrito no presente documento, o alvejamento e a inserção de um gene exógeno no lócus de albumina (por exemplo, no íntron 1) permite que o uso de promotor endógeno da albumina acione a expressão robusta do gene exógeno. A presente divulgação é com base, em parte, na identificação de RNAs guia que alvejam especificamente sítios dentro do gene de albumina, por exemplo, íntron 1 do gene de albumina, e que fornecem a inserção e/ou a expressão eficientes de um gene exógeno. As seguintes

modalidades são fornecidas.

[004] Em um aspecto, a presente divulgação fornece um método para inserir um ácido nucleico que codifica um polipeptídeo heterólogo em um lócus de albumina de uma célula hospedeira ou população de células que compreende administrar: i) um gRNA que compreende uma sequência escolhida de: a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97; d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2-33; e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33; f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34-97; g) uma sequência que é complementar a 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2-33; h) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 98-119; i) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 98-119; e j) uma sequência selecionada a partir do grupo que consiste nas SEQ ID NOs: 120-163; ii) um agente de ligação de DNA guiado por RNA; e iii) um construto que compreende a ácido nucleico que codifica o polipeptídeo heterólogo, assim inserindo o ácido nucleico que codifica o polipeptídeo heterólogo em um lócus de albumina da célula hospedeira ou população de células.

[005] Em outro aspecto, a presente divulgação fornece um método para expressar um polipeptídeo heterólogo de um lócus de albumina de uma célula

hospedeira ou população de células que compreende administrar: i) um gRNA que compreende uma sequência escolhida de: a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97; d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2-33; e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33; f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34-97; e g) uma sequência que compreende 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2-33; ii) um agente de ligação de DNA guiado por RNA; e iii) um construto que compreende a sequência de codificação para o polipeptídeo heterólogo, assim expressando o polipeptídeo heterólogo na célula hospedeira ou população de células.

[006] Em outro aspecto, a presente divulgação fornece um método para expressar um agente terapêutico em um tipo de célula sem divisão ou população de células que compreende administrar: i) um gRNA que compreende uma sequência escolhida de: a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97; d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência

selecionada do grupo que consiste nas SEQ ID NOS: 2-33; e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33; f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34-97; e g) uma sequência que compreende 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2-33; ii) um agente de ligação de DNA guiado por RNA; e iii) um construto que compreende uma sequência de codificação para um polipeptídeo heterólogo, assim expressando o agente terapêutico no tipo de célula sem divisão ou população de células.

[007] Em algumas modalidades, o gRNA compreende uma sequência guia selecionada do grupo que consiste na SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 e SEQ ID NO: 33.

[008] Em algumas modalidades, o método é realizado in vivo. Em algumas modalidades, o método é realizado *in vitro*.

[009] Em algumas modalidades, o gRNA liga uma região a montante de um motivo adjacente de protoespaçador (PAM). Em algumas modalidades, o PAM é escolhido de NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT e NNNNRYAC.

[010] Em algumas modalidades, o gRNA é um gRNA duplo (dgRNA). Em algumas modalidades, o gRNA é um gRNA único (sgRNA). Em algumas modalidades, o sgRNA e compreende um ou mais nucleosídeos modificados. Em algumas modalidades, a Cas nuclease é uma Cas nuclease de classe 2. Em algumas modalidades, a Cas nuclease é selecionada do grupo que consiste em nuclease de

S. pyogenes, nuclease de S. aureus, nuclease de C. jejuni, nuclease de S. thermophilus, nuclease de N. meningitidis e variantes dos mesmos. Em algumas modalidades, a Cas nuclease é Cas9. Em algumas modalidades, a Cas nuclease é uma nickase.

[011] Em algumas modalidades, o construto é um construto de ácido nucleico bidirecional. Em algumas modalidades, o construto compreende: i. um primeiro segmento que compreende uma sequência de codificação para um polipeptídeo heterólogo; e ii. um segundo segmento que compreende um complemento reverso de uma sequência de codificação do polipeptídeo heterólogo. Em algumas modalidades, o construto compreende uma sequência sinal de poliadenilação. Em algumas modalidades, o construto compreende um sítio aceitante de splice. Em algumas modalidades, o construto não compreende um braço de homologia.

[012] Em algumas modalidades, o gRNA é administrado em um vetor e/ou uma nanopartícula lipídica. Em algumas modalidades, o agente de ligação de DNA guiado por RNA é administrado em um vetor e/ou uma nanopartícula lipídica. Em algumas modalidades, o construto que compreende o gene heterólogo é administrado em um vetor e/ou uma nanopartícula lipídica. Em algumas modalidades, o vetor é um vetor viral. Em algumas modalidades, o vetor viral é selecionado do grupo que consiste em um vetor viral adenoassociado (AAV), vetor de adenovírus, vetor de retrovírus e vetor de lentivírus. Em algumas modalidades, o vetor AAV é selecionado do grupo que consiste em AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV 10, AAV11, AAV 12, rh10 e híbridos dos mesmos.

[013] Em algumas modalidades, o gRNA, o agente de ligação de DNA guiado por RNA e o construto que compreende uma sequência de codificação para o polipeptídeo heterólogo, individualmente ou em qualquer combinação, são administrados simultaneamente. Em algumas modalidades, o gRNA, o agente de

ligação de DNA guiado por RNA e o construto que compreende uma sequência de codificação polipeptídeo heterólogo são administrados sequencialmente, em qualquer ordem e/ou em qualquer combinação. Em algumas modalidades, o agente de ligação de DNA guiado por RNA ou agente de ligação de DNA guiado por RNA e gRNA em combinação, é administrado antes de fornecer o construto. Em algumas modalidades, o construto que compreende uma sequência de codificação para o polipeptídeo heterólogo é administrado antes do gRNA e/ou agente de ligação de DNA guiado por RNA.

[014] Em algumas modalidades, o polipeptídeo heterólogo é um polipeptídeo secretado. Em algumas modalidades, o polipeptídeo heterólogo é um polipeptídeo intracelular.

[015] Em algumas modalidades, a célula é uma célula de fígado. Em algumas modalidades, a célula hepática é um hepatócito.

[016] Em algumas modalidades, a expressão do polipeptídeo heterólogo na célula hospedeira é aumentada em pelo menos cerca de 2%, cerca de 3%, cerca de 4%, cerca de 5%, cerca de 6%, cerca de 7%, cerca de 8%, cerca de 9%, cerca de 10%, ou mais, em relação a um nível na célula antes de administrar o gRNA, agente de ligação de DNA guiado por RNA e construto que compreende uma sequência de codificação para o polipeptídeo heterólogo.

[017] Em algumas modalidades, o gRNA que compreende a SEQ ID NO: 301.

[018] Em algumas modalidades, o gRNA media o corte específico de alvo por um agente de ligação de DNA guiado por RNA, resulta em inserção da sequência de codificação para o polipeptídeo heterólogo no íntron 1 de um gene de albumina. Em algumas modalidades, o corte resulta em uma taxa de pelo menos cerca de 10% de inserção de um ácido nucleico heterólogo na população de células. Em algumas modalidades, o corte resulta em uma taxa de entre cerca de 30 e 35%, cerca de 35 e 40%, cerca de 40 e 45%, cerca de 45 e 50%, cerca de 50 e 55%, cerca de 55 e 60%,

cerca de 60 e 65%, cerca de 65 e 70%, cerca de 70 e 75%, cerca de 75 e 80%, cerca de 80 e 85%, cerca de 85 e 90%, cerca de 90 e 95% ou cerca de 95 e 99% de inserção da sequência de codificação para o polipeptídeo heterólogo.

[019] Em algumas modalidades, a proteína de ligação de DNA guiada por RNA é uma Cas9 nuclease de *S. pyogenes*. Em algumas modalidades, a nuclease é uma clivase ou uma nickase.

[020] Em algumas modalidades, o método compreende adicionalmente administrar uma LNP que compreende o gRNA. Em algumas modalidades, o método compreende adicionalmente administrar uma LNP que compreende um mRNA que codifica o agente de ligação de DNA guiado por RNA. Em algumas modalidades, a LNP compreende o gRNA e o mRNA que codifica o agente de ligação de DNA guiado por RNA. Em algumas modalidades, o gRNA e a proteína de ligação de DNA guiada por RNA são administrados como uma RNP. Em algumas modalidades, o construto é administrado por meio de um vetor.

[021] Em um aspecto, a presente divulgação fornece a célula hospedeira produzida por qualquer um ou mais dos métodos supracitados.

[022] Em um aspecto, a presente divulgação fornece uma célula que compreende um construto de ácido nucleico bidirecional que codifica um polipeptídeo heterólogo integrado no íntron 1 de um lócus de albumina de uma célula hospedeira. Em algumas modalidades, a célula hospedeira é uma célula hepática. Em algumas modalidades, a célula hepática é um hepatócito.

BREVE DESCRIÇÃO DOS DESENHOS

[023] A Figura 1 mostra formatos de construto conforme representados em genomas de AAV. SA= aceitante de splice; pA= sequência de sinal de poliA; HA= braço de homologia; LHA= braço de homologia esquerdo; RHA= braço de homologia direito

[024] A Figura 2 mostra que vetores sem braços de homologia não são

eficazes em uma linhagem de célula hepática imortalizada (Hepal-6). Um scAAV derivado de plasmídeo P00204 que compreende braços de homologia de 200 pb resultou na expressão de hFIX nas células de divisão. Uso dos vetores de AAV derivados de P00123 (braços de homologia que carecem de scAAV) e P00147 (braços de homologia que carecem de construto bidirecional de ssAAV) não resulta na expressão detectável de hFIX.

[025] As Figuras 3A e 3B mostram resultados de teste *in vivo* de modelos de inserção com e sem braços de homologia com o uso de vetores derivados de P00123, P00147 ou P00204. A Figura 3A mostra que níveis de edição hepática, conforme medidos pela formação de indel de -60%, foram detectados em cada grupo de animais tratados com LNPs que compreendem componentes de CRISPR/Cas9. A Figura 3B mostra que animais que recebem os vetores de ssAAV sem braços de homologia (derivados de P00147) em combinação com tratamento de LNP resultam no maior nível de expressão de hFIX em soro.

[026] As Figuras 4A e 4B mostram resultados de teste *in vivo* de modelos de inserção de ssAAV com e sem braços de homologia. A Figura 4A compara a inserção alvejada com vetores derivados de plasmídeos P00350, P00356, P00362 (que tem braços de homologia assimétricos, conforme mostrado) e P00147 (construto bidirecional, conforme mostrado na Figura 4B). A Figura 4B compara inserção em um segundo sítio alvejado com vetores derivados de plasmídeos P00353, P00354 (que tem braços de homologia simétricos, conforme mostrado) e P00147.

[027] As Figuras 5A a 5D mostram resultados de inserção alvejada de construtos bidirecionais através de 20 sítios-alvo em hepatócitos de camundongo primários. A Figura 5A mostra os esquemas de cada um dos vetores testados. A Figura 5B mostra a edição, conforme medido pela formação de indel, para cada um dos grupos de tratamento através de cada combinação testada. A Figura 5C e a Figura 5D mostram que níveis significativos de edição (conforme formação de indel em um

sítio-alvo específico) não resultam necessariamente em inserção ou expressão mais eficiente dos transgenes. hSA= aceitante de splice de F9 humano; mSA= aceitante de splice de albumina de camundongo; HiBit= marcação para detecção à base de luciferase; pA= sequência de sinal de poliA; Nluc= repórter de nanoluciferase; GFP= repórter verde fluorescente.

[028] A Figura 6 mostra resultados de exame *in vivo* de inserção alvejada com construtos bidirecionais através de 10 sítios-alvo com o uso de ssAAV derivados de P00147. Conforme mostrado, os níveis significativos de formação de indel não resultam necessariamente em altos níveis de expressão de transgene.

[029] As Figuras 7A a 7D mostram resultados de exame *in vivo* de construtos bidirecionais através de 20 sítios-alvo com o uso de ssAAV derivados de P00147. A Figura 7A mostra que níveis variados de edição, conforme medido pela formação de indel, foram detectados para cada um dos grupos de tratamento através de cada combinação de LNP/vetor testada. A Figura 7B fornece dados de inserção alvejada correspondentes. Os resultados mostram correlação fraca entre formação de indel e inserção ou expressão dos construtos bidirecionais (Figura 7B e Figura 7D), e uma correlação positiva entre resultados *in vitro* e *in vivo* (Figura 7C).

[030] As Figuras 8A e 8B mostram a inserção do construto bidirecional no nível celular com o uso de método de hibridização *in situ* que usa sondas que podem detectar as junções entre o transgene de hFIX e a sequência de éxon 1 de albumina de camundongo (Figura 8A). Níveis de hFIX de circulação correlacionados com o número de células que eram positivas para o transcrito híbrido (Figura 8B).

[031] A Figura 9 mostra o efeito sobre a inserção alvejada de variação da temporização entre a entrega do ssAAV que compreende o construto de hFIX bidirecional e LNP.

[032] A Figura 10 mostra o efeito em inserção alvejada de dosagem repetida (por exemplo, 1, 2, ou 3 doses) de LNP depois da entrega do construto de hFIX

bidirecional.

[033] A Figura 11A mostra a durabilidade de expressão de hFIX *in vivo*. A Figura 11B demonstra que a expressão de íntron 1 de albumina foi prolongada.

[034] As Figuras 12A e 12B mostram que a dose de AAV ou LNP variante pode modular a quantidade de expressão de hFIX de íntron 1 do gene de albumina *in vivo*.

[035] As Figuras 13A a 13C mostram resultados de construtos bidirecionais de exame através de sítios-alvo em hepatócitos cinomolgos primários. A Figura 13A mostra níveis variados de edição, conforme medido pela formação de indel, detectados para cada uma das amostras. A Figura 13B e a Figura 13C mostram que níveis significativos de formação de indel não foram preditivos para inserção ou expressão dos construtos bidirecionais em íntron 1 do gene de albumina.

[036] As Figuras 14A a 14D mostram resultados de construtos bidirecionais de exame através de sítios-alvo em hepatócitos humanos primários. A Figura 14A mostra a edição, conforme medido pela formação de indel, detectada para cada uma das amostras. A Figura 14B, Figura 14C e 14D mostram que níveis significativos de formação de indel não foram preditivos para inserção ou expressão dos construtos bidirecionais em íntron 1 do gene de albumina.

[037] A Figura 15 mostra os resultados de estudos *in vivo* em que primatas não humanos foram dosados com LNPs juntamente com um modelo de inserção de hFIX bidirecional (derivado de P00147). Níveis de hFIX sistêmicos foram alcançados apenas em animais tratados tanto com LNPs quanto AAV, sem hFIX detectável com o uso de AAV ou LNPs sozinhos.

[038] A Figura 16A e a Figura 16B mostram níveis de expressão de Fator IX humano nas amostras plasmáticas na semana 6 após injeção.

[039] A Figura 17 mostra níveis de soro na semana 7 e % de células positivas através dos múltiplos lóbulos para cada animal.

DESCRIÇÃO DETALHADA

[040] Referência será feita agora em detalhes a determinadas modalidades da invenção, exemplos dessas são ilustrados nos desenhos anexos. Embora os presentes ensinamentos tenham sido descritos em combinação com várias modalidades, não se pretende limitar os presentes ensinamentos àquelas modalidades. Pelo contrário, os presentes ensinamentos abrangem várias alternativas, modificações e equivalentes, conforme será observado por aqueles versados na técnica.

[041] Antes de descrever os presentes ensinamentos em detalhes, deve-se entender que a divulgação não se limita às composições ou etapas de processo específicas, as quais podem variar. Deve-se notar que, conforme usado neste relatório descritivo e nas reivindicações anexas, a forma singular “um”, “uma”, “a” e “o” incluem referências plurais a menos que o contexto dite o contrário. Assim, por exemplo, a referência a “um conjugado” inclui uma pluralidade de conjugados e referência a “uma célula” inclui uma pluralidade de células e semelhantes. Conforme usado no presente documento, o termo “inclui” e suas variantes gramaticais pretendem ser não limitantes, de modo que a citação de itens em uma lista não seja a exclusão de outros itens semelhantes que podem ser substituídos ou adicionados aos itens listados.

[042] As faixas numéricas são inclusivas dos números que definem a faixa. Entende-se que os valores medidos e mensuráveis sejam aproximados, levando-se em conta dígitos significativos e o erro associado à medição. Além disso, o uso de “compreendem”, “compreende”, “que compreende”, “contêm”, “contém”, “que contém”, “incluem”, “inclui” e “que inclui” não pretende ser limitante. Deve-se entender que tanto a descrição geral anterior quanto a descrição detalhada são apenas exemplificativas e explicativas e não são restritivas dos ensinamentos.

[043] A menos que seja observado especificamente no relatório descritivo, modalidades no relatório descritivo que citam “que compreendem” vários

componentes também são contempladas como “que consistem nos” ou “que consistem essencialmente nos” componentes citados; modalidades no relatório descritivo que citam “que consistem nos” vários componentes também são contempladas como “que compreendem nos” ou “que consistem essencialmente nos” componentes citados; e modalidades no relatório descritivo que citam “que consistem essencialmente nos” vários componentes também são contempladas como “que consistem nos” ou “que compreendem nos” componentes citados (isso não se aplica intercambiavelmente ao uso desses termos nas reivindicações). O termo “ou” é usado em um sentido inclusivo, isto é, equivalente a “e/ou”, a menos que o contexto indique claramente de outra forma. O termo “cerca de”, quando usado antes de uma lista, modifica cada membro da lista. O termo “cerca de” ou “aproximadamente” significa um erro aceitável para um valor particular conforme determinado por um versado na técnica, que depende em parte de como o valor é medido ou determinado.

[044] O termo “cerca de”, quando usado antes de uma lista, modifica cada membro da lista. O termo “cerca de” ou “aproximadamente” significa um erro aceitável para um valor particular conforme determinado por um versado na técnica, que depende em parte de como o valor é medido ou determinado.

[045] Os títulos de seção usados no presente documento são para fins organizacionais apenas e não devem ser interpretados como limitação da matéria desejada de maneira nenhuma. No caso de qualquer material incorporado a título de referência contradizer qualquer termo definido neste relatório descritivo ou qualquer outro conteúdo expresso deste relatório descritivo, este relatório descritivo tem o controle. Embora os presentes ensinamentos tenham sido descritos em combinação com várias modalidades, não se pretende que os presentes ensinamentos sejam limitados a tais modalidades. Pelo contrário, os presentes ensinamentos abrangem várias alternativas, modificações e equivalentes, conforme será observado por aqueles versados na técnica.

I. DEFINIÇÕES

[046] A menos que seja afirmado de outra forma, os seguintes termos e frases, conforme usados no presente documento devem ter os seguintes significados:

[047] “Polinucleotídeo” e “ácido nucleico” são usados no presente documento para se referir a um composto multimérico que compreende nucleosídeos ou análogos de nucleosídeo que têm bases heterocíclicas nitrogenosas ou análogos de base ligados juntos ao longo de uma cadeia principal, que inclui RNA convencional, DNA, RNA e DNA misturados, e polímeros que são análogos dos mesmos. Uma “cadeia principal” de ácido nucleico pode ser constituída por uma variedade de ligações, que incluem uma ou mais por ligações de açúcar e fosfodiéster, ligações de peptídeo e ácido nucleico (“ácidos nucleicos de peptídeo” ou PNA; documento PCT nº WO 95/32305), ligações de fosforotioato, ligações de metilfosfonato ou combinações das mesmas. Porções químicas de açúcar de um ácido nucleico podem ser ribose, desoxirribose, ou compostos similares com substituições opcionais, por exemplo, substituições de 2’ metóxi ou 2’ haleto. Bases nitrogenosas podem ser bases convencionais (A, G, C, T, U), análogos das mesmas (por exemplo, uridinas modificadas, tais como 5-metoxiuridina, pseudouridina ou N1-metilpseudouridina, ou outras); inosina; derivados de purinas ou pirimidinas (por exemplo, desoxiguanosina de N^4-metila, deaza- ou aza-purinas, deaza- ou aza-pirimidinas, bases de pirimidina com grupos de substituinte na posição 5 ou 6 (por exemplo, 5-metilcitosina), bases de purina com um substituinte nas posições 2, 6 ou 8, 2-amino-6-metilaminopurina, O^6-metilguanina, 4-tio-pirimidinas, 4-amino-pirimidinas, 4-dimetil-hidrazina-pirimidinas, e O^4-alquil-pirimidinas; documento de Patente US nº 5.378.825 e PCT nº WO 93/13121). Para a discussão geral, consultar, The Biochemistry of the Nucleic Acids, 5-36, Adams et al., 11ª edição, 1992). Ácidos nucleicos podem incluir um ou mais resíduos “abásicos” quando a cadeia principal não inclui base nitrogenada para a posição (ou posições) do polímero (documento de Patente US nº 5.585.481). Um ácido nucleico

pode compreender apenas açúcares de RNA ou DNA convencionais, bases e ligações, ou pode incluir tanto componentes quanto substituições convencionais (por exemplo, nucleosídeos convencionais com substituintes de 2’ metóxi, ou polímeros que contêm tanto nucleotídeos convencionais quanto um ou mais análogos de nucleotídeo). Ácido nucleico inclui “ácido nucleico travado” (LNA), um análogo que contém um ou mais monômeros de nucleotídeo de LNA com uma unidade de furanose bicíclica travada em uma conformação de açúcar de imitação de RNA, que intensificam a afinidade de hibridização em direção às sequências de RNA e DNA complementares (Vester e Wengel, 2004, Biochemistry 43(42): 13233-41). RNA e DNA têm diferentes porções químicas de açúcar e podem se diferir pela presença de uracila ou análogos da mesma em RNA e timina ou análogos da mesma em DNA.

[048] “RNA guia”, ”gRNA” e simplesmente “guia” são usados no presente documento intercambiavelmente para se referir a um guia que compreende uma sequência guia, por exemplo, um crRNA (também conhecido como RNA de CRISPR), ou a combinação de um crRNA e um trRNA (também conhecido como tracrRNA). O crRNA e trRNA podem ser associados como uma single molécula de RNA (RNA guia único, sgRNA) ou, por exemplo, em duas moléculas de RNA separadas (RNA guia duplo, dgRNA). “RNA” ou “gRNA” se refere a cada tipo. O trRNA pode ser uma sequência de ocorrência natural ou uma sequência de trRNA com modificações ou variações em comparação com sequências de ocorrência natural. RNAs guia, tais como sgRNAs ou dgRNAs, podem incluir RNAs modificados, conforme descrito no presente documento.

[049] Conforme usado no presente documento, uma “sequência guia” se refere 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 de DNA guiado por RNA. Uma “sequência guia” também pode ser denominada uma

“sequência de alvejamento” ou uma “sequência de espaçador”. Uma sequência guia pode ser 20 pares de base em comprimento, por exemplo, no caso de Streptococcus pyogenes (isto é, Spy Cas9) e homólogos/ortólogos de 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 em 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 cerca de 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a sequência guia e a região-alvo podem ter 100% de complementaridade ou ser idênticas. Em outras modalidades, a sequência guia e a região-alvo podem conter pelo menos uma disparidade. Por exemplo, a sequência guia e a sequência-alvo podem conter 1, 2, 3 ou 4 disparidades, quando o comprimento total da sequência-alvo tem pelo menos 17, 18, 19, 20 ou mais pares de base. Em algumas modalidades, a sequência guia e a região-alvo podem conter 1-4 disparidades quando 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 disparidades quando a sequência guia compreende 20 nucleotídeos.

[050] Sequências-alvo para agentes de ligação de DNA guiados por RNA incluem ambas as fitas positivas e negativas de DNA genômico (isto é, a sequência dada e o complemento reverso da sequência), como um substrato de ácido nucleico para um agente de ligação de DNA guiado por RNA é um ácido nucleico de fita dupla. Consequentemente, quando se diz que uma sequência guia é “complementar a uma sequência-alvo”, entende-se que a sequência guia pode direcionar um RNA guia para se ligar à fita senso ou antissenso (por exemplo, complemento reverso) de uma sequência-alvo. Assim, em algumas modalidades, quando a sequência guia se liga ao

complemento reverso de uma sequência-alvo, a sequência guia é idêntica a determinados nucleotídeos da sequência-alvo (por exemplo, a sequência-alvo que não inclui o PAM) exceto pela substituição de U por T na sequência guia.

[051] Conforme usado no presente documento, um “agente de ligação de DNA guiado por RNA” significa um polipeptídeo ou complexo de polipeptídeos que tem atividade de ligação de RNA e DNA, ou uma subunidade de ligação de DNA de um tal complexo, sendo que a atividade de ligação de DNA é de sequência específica e depende da sequência do RNA. O termo agente de ligação de DNA guiado por RNA também inclui ácidos nucleicos que codificam tais polipeptídeos. Agentes de ligação de DNA guiados por RNA exemplificativos incluem Cas clivases/nickases. Agentes de ligação de DNA guiados por RNA exemplificativos podem incluir formas inativadas dos mesmos (“agentes de ligação de DNA de dCas”), por exemplo, se esses agentes forem modificados para permitir clivagem de DNA, por exemplo, por meio de fusão com um domínio de FokI clivase. “Cas nuclease”, conforme usado no presente documento, engloba Cas clivases e Cas nickases. Cas clivases e Cas nickases incluem um complexo Csm ou Cmr de um sistema de CRISPR de tipo III, a subunidade Cas 10, Csm1 ou Cmr2 do mesmo, um complexo de Cascata de um sistema de CRISPR de tipo I, a subunidade Cas3 do mesmo, e Cas nucleases de Classe 2. Conforme usado no presente documento, uma “Cas nuclease de Classe 2” é um polipeptídeo de cadeia única com atividade de ligação de DNA guiada por RNA. Cas nucleases de Classe 2 incluem Cas clivases/nickases de Classe 2 (por exemplo, variantes H840A, D10A ou N863A), que têm, adicionalmente, atividade de clivases ou nickase de DNA guiadas por RNA, e agentes de ligação de DNA de dCas de Classe 2, em que a atividade de clivase/nickase é inativada”), se esses agentes forem modificados para permitir clivagem de DNA. Cas nucleases de Classe 2 incluem, por exemplo, proteínas 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 modificações das mesmas. A proteína Cpf1, Zetsche et ah, Cell, 163: 1-13 (2015), também contém um domínio de nuclease semelhante a RuvC. Sequências de Cpf1 de Zetsche são incorporadas a título de referência em sua totalidade. Consultar, por exemplo, Zetsche, Tabelas S1 e S3. Consultar, por exemplo, Makarova et al., Nat Rev Microbiol, 13(11): 722-36 (2015); Shmakov et al., Molecular Cell, 60:385-397 (2015). Conforme usado no presente documento, a entrega de um agente de ligação de DNA guiado por RNA (por exemplo, uma Cas nuclease, uma Cas9 nuclease ou uma Cas9 nuclease de S. pyogenes) inclui entrega do polipeptídeo ou mRNA.

[052] Conforme usado no presente documento, “ribonucleoproteína” (RNP) ou “complexo de RNP” se refere a um RNA guia juntamente com um agente de ligação de DNA guiado por RNA, tal como uma Cas nuclease, por exemplo, uma Cas clivase, Cas nickase, Cas9 clivase ou Cas9 nickase. Em algumas modalidades, o RNA guia guia o agente de ligação de DNA guiado por RNA, tal como Cas9 para uma sequência-alvo, e o RNA guia hibridiza com a sequência-alvo e o agente se liga à mesma; e a ligação pode ser seguida por clivagem ou corte.

[053] Conforme usado no presente documento, uma primeira sequência é considerada por “compreender uma sequência com pelo menos X% de identidade com” uma segunda sequência se um alinhamento da primeira sequência com a segunda sequência mostra que X% ou mais das posições da segunda sequência em sua totalidade são correspondidas pela primeira sequência. Por exemplo, a sequência AAGA compreende uma sequência com 100% de identidade com a sequência AAG uma vez que um alinhamento geraria 100% de identidade, em que existem correspondência com 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 nucleosídeo, tais como uridinas modificadas não contribuem

para as diferenças em identidade ou complementaridade entre polinucleotídeos desde que os nucleotídeos relevantes (tal 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, em que ambas têm guanosina ou guanosina modificada como um complemento). Assim, por exemplo, a sequência 5'-AXG, em que X é qualquer uridina modificada, tal como pseudouridina, pseudouridina de N1-metila ou 5-metoxiuridina, é considerada 100% idêntica a AUG em que ambas são perfeitamente complementares à mesma sequência (5'-CAU). Algoritmos de alinhamento exemplificativos são os algoritmos Smith-Waterman e Needleman-Wunsch, que são bem conhecidos na técnica. Uma pessoa versada na técnica entenderá que a escolha de configurações de algoritmo e parâmetro é apropriada para um dado par de sequências a ser alinhado; para sequências de comprimento geralmente similar e identidade esperada >50% para aminoácidos ou >75% para nucleotídeos, o algoritmo Needleman-Wunsch com configurações padrão da interface de algoritmo Needleman-Wunsch fornecido pelo EBI no servidor da web www.ebi.ac.uk é geralmente apropriado.

[054] Conforme usado no presente documento, uma primeira sequência é considerada como "X% complementar a" uma segunda sequência se X% das bases da primeira sequência fizerem par com a segunda sequência. Por exemplo, uma primeira sequência 5'AAGA3' é 100% complementar a uma segunda sequência 3'TTCT5', e a segunda sequência é 100% complementar à primeira sequência. Em algumas modalidades, uma primeira sequência 5'AAGA3' é 100% complementar a uma segunda sequência 3'TTCTGTGA5', enquanto a segunda sequência é 50% complementar à primeira sequência.

[055] "mRNA" é usado no presente documento para se referir a um polinucleotídeo 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 aminoacilados). mRNA compreenderá predominantemente RNA ou RNA modificado e pode compreender uma cadeia principal de fosfato-açúcar que inclui resíduos de ribose ou análogos dos mesmos, por exemplo, resíduos de ribose de 2'-metóxi. Em algumas modalidades, os açúcares de uma cadeia principal de fosfato-açúcar de mRNA consistem essencialmente em resíduos de ribose, resíduos de ribose de 2'-metóxi ou uma combinação dos mesmos. Bases de um mRNA podem ser bases modificadas, como pseudouridina, N-1-metil-pseudouridina, ou outras bases de ocorrência natural ou de ocorrência não natural.

[056] Sequências guia exemplificativas úteis nas composições e métodos descritos no presente documento são mostradas na Tabela 1 e ao longo do pedido.

[057] Conforme usado no presente documento, "indels" se referem às mutações de inserção/deleção que consistem em um número de nucleotídeos que são inseridos ou deletados no sítio de quebra de fita dupla (DSBs) em um ácido nucleico-alvo.

[058] Conforme usado no presente documento, "polipeptídeo" se refere a uma proteína de tipo selvagem ou variante (por exemplo, mutante, fragmento, fusão ou combinações dos mesmos). Um polipeptídeo variante pode ter pelo menos ou cerca de 5%, 10%, 15%, 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ou 100% de atividade funcional do polipeptídeo de tipo selvagem. Em algumas modalidades, a variante é pelo menos 70%, 75%, 80%, 85%, 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99% idêntica à sequência do polipeptídeo de tipo selvagem. Em algumas modalidades, um polipeptídeo variante pode ser uma variante hiperativa. Em determinados casos, a variante tem entre cerca de 80% e cerca de 120%, 140%, 160%, 180%, 200%, 300%, 400%, 500% ou mais de uma atividade funcional do polipeptídeo de tipo selvagem.

[059] Conforme usado no presente documento, uma "sequência-alvo" se refere a uma sequência de ácido nucleico em um gene-alvo que tem

complementaridade com a sequência guia do gRNA. A interação da sequência-alvo e da sequência guia direciona um agente de ligação de DNA guiado por RNA para se ligar, e potencialmente cortar ou clivar (dependendo da atividade do agente), dentro da sequência-alvo.

[060] Conforme usado no presente documento, um “gene heterólogo” se refere a um gene que foi introduzido como uma fonte exógena para um sítio dentro de um genoma de célula hospedeira (por exemplo, um sítio de íntron 1 de albumina). Isto é, o gene introduzido é heterólogo com relação a seu sítio de inserção. Um polipeptídeo expresso a partir de tal gene heterólogo é denominado um “polipeptídeo heterólogo”. O gene heterólogo pode ser de ocorrência natural ou modificado geneticamente, e pode ser de tipo selvagem ou uma variante. O gene heterólogo pode incluir sequências de nucleotídeos diferentes da sequência que codifica o polipeptídeo heterólogo (por exemplo, um sítio de entrada ribossômico interno). O gene heterólogo pode ser um gene que ocorre naturalmente no genoma hospedeiro, como um tipo selvagem ou uma variante (por exemplo, mutante). Por exemplo, embora a célula hospedeira contenha o gene de interesse (como um tipo selvagem ou como uma variante), o mesmo gene ou variante do mesmo pode ser introduzido como uma fonte exógena para, por exemplo, expressão em um lócus que é altamente expresso. O gene heterólogo também pode ser um gene que não é de ocorrência natural no genoma hospedeiro, ou que expressa um polipeptídeo heterólogo que não é de ocorrência natural no genoma hospedeiro. “Gene heterólogo”, “gene exógeno” e “transgene” são usados intercambiavelmente. Em algumas modalidades, o gene heterólogo ou transgene inclui uma sequência de ácidos nucleicos exógena, por exemplo, uma sequência de ácidos nucleicos não é endógena à célula receptora. Em algumas modalidades, o gene heterólogo ou transgene inclui uma sequência de ácidos nucleicos exógena, por exemplo, uma sequência de ácidos nucleicos que não é de ocorrência natural na célula receptora. Por exemplo, um gene heterólogo pode

ser heterólogo com relação a seu sítio de inserção e com relação a sua célula receptora.

[061] Um gene heterólogo pode ser inserido em um lócus de porto seguro no genoma sem efeitos prejudiciais significativos na célula hospedeira, por exemplo, hepatócito, por exemplo, sem causar apoptose, necrose e/ou senescência, ou sem causar mais do que 5%, 10%, 15%, 20%, 25%, 30% ou 40% de apoptose, necrose e/ou senescência em comparação com uma célula de controle. Consultar, por exemplo, Hsin et al., ”Hepatocyte death in liver inflammation, fibrosis, and tumorigenesis”, 2017. Em algumas modalidades, um lócus de porto seguro permite a superexpressão de um gene exógeno sem efeitos negativos significativos na célula hospedeira, por exemplo, hepatócito, por exemplo, sem causar apoptose, necrose e/ou senescência, ou sem causar mais do que 5%, 10%, 15%, 20%, 25%, 30% ou 40% de apoptose, necrose e/ou senescência em comparação com uma célula de controle. Em algumas modalidades, um lócus de porto seguro desejável pode ser um no qual a expressão da sequência de gene inserida não é perturbada pela expressão de leitura atenta de genes vizinhos. Em algumas modalidades, um lócus de porto seguro permite a expressão de um gene exógeno sem efeitos negativos significativos na célula hospedeira ou população de células, tais como hepatócitos ou células hepáticas, por exemplo, sem causar apoptose, necrose e/ou senescência, ou sem causar mais do que 5%, 10%, 15%, 20%, 25%, 30% ou 40% de apoptose, necrose e/ou senescência em comparação com uma célula de controle ou população de células.

[062] Em algumas modalidades, o gene heterólogo pode ser inserido em um lócus de porto seguro e usar a sequência de sinal endógena de lócus de porto seguro, por exemplo, a sequência de sinal de albumina codificada por éxon 1. Por exemplo, uma sequência de codificação pode ser inserida em íntron de albumina humana 1 de modo que esteja a jusante de e se funda à sequência de sinal de éxon de albumina

humana 1.

[063] Em algumas modalidades, o gene pode compreender sua própria sequência de sinal, pode ser inserido no lócus de porto seguro, e pode usar adicionalmente a sequência de sinal endógena de lócus de porto seguro. Por exemplo, uma sequência de codificação que compreende sua sequência de sinal nativa pode ser inserida em íntron de albumina humana 1 de modo que esteja a jusante de e se funda à sequência de sinal de albumina humana codificada por éxon 1.

[064] Em algumas modalidades, o gene pode compreender sua própria sequência de sinal e um sítio de entrada ribossômica interna (IRES), pode ser inserido no lócus de porto seguro, e pode usar adicionalmente a sequência de sinal endógena de lócus de porto seguro. Por exemplo, uma sequência de codificação que compreende sua sequência de sinal nativa e uma sequência de IRES pode ser inserida no íntron de albumina humana 1 de modo que esteja a jusante de e se funda à sequência de sinal de albumina humana codificada por éxon 1.

[065] Em algumas modalidades, o gene pode compreender sua própria sequência de sinal e IRES, pode ser inserido no lócus de porto seguro, e não usa a sequência de sinal endógena de lócus de porto seguro. Por exemplo, uma sequência de codificação que compreende sua sequência de sinal nativa e uma sequência de IRES pode ser inserida no íntron de albumina humana 1 de modo que não se funda à sequência de sinal de albumina humana codificada por éxon 1. Nessas modalidades, uma proteína é traduzida do sítio de IRES e não é quimérica (por exemplo, peptídeo de sinal de albumina fundido à proteína heteróloga), que pode ser vantajosamente não ou pouco imunogênica. Em algumas modalidades, a proteína não é secretada e/ou transportada de modo extracelular.

[066] Em algumas modalidades, o gene pode ser inserido no lócus de porto seguro e pode compreender um IRES e não usa qualquer sequência de sinal. Por exemplo, uma sequência de codificação que compreende uma sequência de IRES e

nenhuma sequência de sinal nativa pode ser inserida no íntron de albumina humana 1 de modo que não se funda à sequência de sinal de albumina humana codificada por éxon 1. Em algumas modalidades, as proteínas são traduzidas do sítio IRES sem qualquer sequência de sinal. Em algumas modalidades, a proteína não é secretada e/ou transportada de modo extracelular.

[067] Conforme usado no presente documento, um “construto de ácido nucleico bidirecional” (intercambiavelmente denominado no presente documento um ”construto bidirecional”) compreende pelo menos dois segmentos de ácido nucleico, em que um segmento (o primeiro segmento) compreende uma sequência de codificação que codifica um agente de interesse (a sequência de codificação pode ser denominada no presente documento “transgene” ou um primeiro transgene), enquanto o outro segmento (o segundo segmento) compreende uma sequência, em que o complemento da sequência codifica um agente de interesse, ou um segundo transgene.

[068] Em uma modalidade, um construto bidirecional compreende pelo menos dois segmentos de ácido nucleico em cis, em que um segmento (o primeiro segmento) compreende uma sequência de codificação (por vezes intercambiavelmente denominado no presente documento “transgene”), embora o outro segmento (o segundo segmento) compreenda uma sequência em que o complemento da sequência codifica um transgene. O primeiro transgene e o segundo transgene podem ser iguais ou diferentes. Os construtos bidirecionais podem compreender pelo menos dois segmentos de ácido nucleico em cis, em que um segmento (o primeiro segmento) compreende uma sequência de codificação que codifica um gene heterólogo em uma orientação, embora o outro segmento (o segundo segmento) compreenda uma sequência em que seu complemento codifica o gene heterólogo na outra orientação. Isto é, o primeiro segmento é um complemento do segundo segmento (não necessariamente um complemento perfeito); o complemento do segundo segmento é

o complemento reverso do primeiro segmento (não necessariamente um complemento reverso perfeito, embora ambos codifiquem a mesma proteína heteróloga). Um construto bidirecional pode compreender uma primeira sequência de codificação que codifica um gene heterólogo ligado a um aceitante de splice e uma segunda sequência de codificação, em que o complemento codifica um gene heterólogo na outra orientação, também ligado a um aceitante de splice.

[069] O agente pode ser agente terapêutico, como um polipeptídeo, RNA funcional, mRNA ou semelhantes. O transgene pode codificar um agente, como um polipeptídeo, RNA funcional ou mRNA. Em algumas modalidades, o construto de ácido nucleico bidirecional compreende pelo menos dois segmentos de ácido nucleico, em que um segmento (o primeiro segmento) compreende uma sequência de codificação que codifica um polipeptídeo de interesse, enquanto o outro segmento (o segundo segmento) compreende uma sequência em que o complemento da sequência codifica um polipeptídeo de interesse, ou um segundo transgene. Isto é, os pelo menos dois segmentos podem codificar polipeptídeos idênticos ou diferentes ou agentes idênticos ou diferentes. Quando os dois segmentos codificam um polipeptídeo idêntico, a sequência de codificação do primeiro segmento não precisa ser idêntica ao complemento da sequência do segundo segmento. Em algumas modalidades, a sequência do segundo segmento é um complemento reverso da sequência de codificação do primeiro segmento. Um construto bidirecional pode ser de fita simples ou fita dupla. O construto bidirecional divulgado no presente documento engloba um construto que tem capacidade para expressar qualquer polipeptídeo de interesse. Os construtos bidirecionais são úteis para inserção genômica de sequências de transgene, em particular, a inserção alvejada do transgene.

[070] Em algumas modalidades, um construto de ácido nucleico bidirecional compreende um primeiro segmento que compreende uma sequência de codificação que codifica um primeiro polipeptídeo (um primeiro transgene), e um segundo

segmento que compreende uma sequência em que o complemento da sequência codifica um segundo polipeptídeo (um segundo transgene). Em algumas modalidades, o primeiro e o segundo polipeptídeos são pelo menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou 100% idênticos. Em algumas modalidades, o primeiro e o segundo polipeptídeos compreendem uma sequência de aminoácidos que é pelo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou 100% idêntica, por exemplo, através de 50, 100, 200, 500, 1000 ou mais resíduos de aminoácido.

[071] Conforme usado no presente documento, um “complemento reverso” se refere a uma sequência que é uma sequência de complemento de uma sequência de referência, em que a sequência de complemento é gravada na orientação reversa. Por exemplo, para uma sequência hipotética 5’ CTGGACCGA 3’ (SEQ ID NO: 500), a sequência complementar “perfeita” é 3’ GACCTGGCT 5’ (SEQ ID NO: 501) e o complemento reverso “perfeito” é escrito 5’ TCGGTCCAG 3’ (SEQ ID NO: 502). Uma sequência de complemento reversa não precisa ser “perfeita” e pode ainda codificar o mesmo polipeptídeo ou um polipeptídeo similar como a sequência de referência. Devido à redundância de uso de códon, um complemento reverso pode divergir de uma sequência de referência que codifica o mesmo polipeptídeo. Conforme usado no presente documento, “complemento reverso” também inclui sequências que são, por exemplo, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou 100% idênticas à sequência de complemento reverso de uma sequência de referência.

II. COMPOSIÇÕES

A. COMPOSIÇÕES QUE COMPREENDEM RNA GUIA (GRNAS)

[072] São fornecidas no presente documento composições e métodos úteis para inserir e expressar um gene heterólogo (exógeno) em um lócus genômico, como um sítio de porto seguro, de uma célula hospedeira. Em particular, conforme

exemplificado no presente documento, o alvejamento e a inserção de um gene exógeno no lócus de albumina (por exemplo, no íntron 1) permite que o uso de promotor endógeno da albumina acione a expressão robusta do gene exógeno. A presente divulgação é com base, em parte, na identificação de RNAs guia que alvejam especificamente sítios dentro do íntron 1 do gene de albumina, e que fornecem a inserção e a expressão eficientes de um gene exógeno. Conforme mostrado nos Exemplos e descrito adicionalmente no presente documento, a habilidade de gRNAs identificados para mediar altos níveis de edição, conforme medido através da atividade de formação de indel, inesperadamente não se correlaciona necessariamente com o uso dos mesmos gRNAs para mediar a inserção eficiente de transgenes, conforme medido através, por exemplo, da expressão do transgene. Isto é, determinados gRNAs que têm capacidade para alcançar um nível significativo de formação de indel não têm necessariamente capacidade para mediar a inserção eficiente, e por outro lado, alguns gRNAs mostrados para alcançar baixos níveis de formação de indel podem mediar a inserção e a expressão eficiente de um transgene. Especificamente, os dados dos Exemplos indicam que gRNAs que mediam de modo eficaz a formação de indel (também chamados % de edição) não tiveram atividade de edição de indel que se correlacionaram com a atividade de edição de inserção.

[073] Em algumas modalidades, são fornecidos no presente documento composições e métodos úteis para inserir e expressar um gene exógeno em íntron 1 do gene de albumina em uma célula hospedeira. Em algumas modalidades, são divulgados no presente documento composições e métodos úteis para introduzir ou inserir um ácido nucleico heterólogo dentro de um lócus de albumina de uma célula hospedeira, por exemplo, com o uso de um RNA guia divulgado no presente documento com um agente de ligação de DNA guiado por RNA, e um construto que compreende um ácido nucleico heterólogo (“transgene”). Em algumas modalidades, são divulgadas no presente documento composições e métodos úteis para expressar

um polipeptídeo heterólogo em um lócus de albumina de uma célula hospedeira, por exemplo, com o uso de um RNA guia divulgado no presente documento com um agente de ligação de DNA guiado por RNA, e um construto que compreende um ácido nucleico heterólogo (“transgene”). Em algumas modalidades, são divulgadas no presente documento composições e métodos úteis para induzir uma quebra (por exemplo, quebra de fita dupla (DSB) ou quebra de fita simples (corte)) dentro do gene de albumina de uma célula hospedeira, por exemplo, com o uso de um RNA guia divulgado no presente documento com um agente de ligação de DNA guiado por RNA (por exemplo, um sistema de CRISPR/Cas). As composições e métodos podem ser usados in vitro ou in vivo para, por exemplo, propósitos terapêuticos.

[074] Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que se liga, ou tem capacidade para se ligar dentro de um íntron de um lócus de albumina. Em algumas modalidades, os RNAs guia divulgados no presente documento se ligam dentro de uma região de íntron 1 do gene de albumina humano (SEQ ID NO: 1). Será reconhecido que nem toda base da sequência guia deve se ligar dentro das regiões citadas. Por exemplo, em algumas modalidades, 15, 16, 17, 18, 19, 20 ou mais bases da sequência de RNA guia se ligam com as regiões citadas. Por exemplo, em algumas modalidades, 15, 16, 17, 18, 19, 20 ou mais bases contíguas da sequência de RNA guia se ligam com as regiões citadas.

[075] Em algumas modalidades, os RNAs guia divulgados no presente documento mediam um corte de alvo específico por um agente de ligação de DNA guiado por RNA (por exemplo, Cas nuclease) em um sítio dentro do íntron 1 de albumina humana (SEQ ID NO: 1). Será reconhecido que, em algumas modalidades, os RNAs guia compreendem sequências guia que se ligam, ou têm capacidade para se ligarem a uma região na SEQ ID NO: 1.

[076] Em algumas modalidades, os RNAs guia divulgados no presente

documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs:2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 164-196. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs:98-119. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada da Tabela 1. O gRNA pode compreender uma ou mais das sequências guia mostradas na Tabela 1. O gRNA pode compreender uma ou mais das sequências mostradas nas Tabelas 1, 7 e 9. O gRNA pode compreender uma ou mais das sequências mostradas nas Tabelas 2, 8 e 10. O RNA guia pode compreender uma ou mais das SEQ ID NOs: 2-33. O gRNA pode compreender uma ou mais das SEQ ID NOs: 164-196. O gRNA pode compreender uma ou mais das SEQ ID NOs: 98-119.

[077] Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs:2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 164-196. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 98-119. Em algumas modalidades, os RNAs

guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada da Tabela 1.

[078] Em algumas modalidades, o RNA guia compreende uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 e SEQ ID NO: 33.

[079] Em algumas modalidades, o RNA guia de albumina (gRNA) compreende uma sequência guia escolhida de: a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97; d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2-33; e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33; f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34-97; e g) uma sequência que é complementar a 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2-33. Em algumas modalidades, o RNA guia de albumina compreende uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, o RNA guia compreende uma

sequência selecionada do grupo que consiste nas SEQ ID NOs: 4, 13, 17, 19, 27, 28, 30, e 31.

[080] Em algumas modalidades, os RNAs guia divulgados no presente documento se ligam a uma região a montante de um motivo adjacente de protoespaçador (PAM). Conforme seria entendido por aqueles versados na técnica, a sequência de PAM ocorre na fita oposta à fita que contém a sequência-alvo. Isto é, a sequência de PAM está na fita de complemento da fita-alvo (a fita que contém a sequência-alvo à qual o RNA guia se liga). Em algumas modalidades, o PAM é selecionado do grupo que consiste em NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT e NNNNRYAC. Em algumas modalidades, o PAM é NGG.

[081] Em algumas modalidades, as sequências de RNA guia são fornecidas no presente documento complementares a uma sequência adjacente a uma sequência de PAM.

[082] Em algumas modalidades, a sequência de RNA guia compreende uma sequência que é complementar a uma sequência dentro de uma região genômica selecionada da Tabela 1 de acordo com coordenadas no genoma de referência humano hg38. Em algumas modalidades, a sequência de RNA guia compreende uma sequência que é complementar a uma sequência que compreende 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 ou 25 nucleotídeos consecutivos de dentro de uma região genômica selecionada da Tabela 1. Em algumas modalidades, a sequência de RNA guia compreende uma sequência que é complementar a uma sequência que compreende 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 ou 25 nucleotídeos consecutivos que abrangem uma região genômica selecionada da Tabela 1.

[083] Os RNAs guia divulgados no presente documento mediam um corte de alvo específico que resulta em uma quebra de fita dupla (DSB). Os RNAs guia divulgados no presente documento mediam um corte de alvo específico que resulta

em uma quebra de fita simples (SSB ou corte).

[084] Em algumas modalidades, os RNAs guia divulgados no presente documento mediam corte de alvo específico por um agente de ligação de DNA guiado por RNA (por exemplo, uma Cas nuclease, conforme divulgado no presente documento), que resulta na inserção de um ácido nucleico heterólogo dentro do íntron 1 de um gene de albumina. Em algumas modalidades, o RNA guia e/ou corte no sítio de corte resulta em uma taxa entre 30 e 35%, 35 e 40%, 40 e 45%, 45 e 50%, 50 e 55%, 55 e 60%, 60 e 65%, 65 e 70%, 70 e 75%, 75 e 80%, 80 e 85%, 85 e 90%, 90 e 95% ou 95 e 99% de inserção de um gene heterólogo. Em algumas modalidades, o RNA guia e/ou corte resulta em pelo menos 20%, pelo menos 30%, pelo menos 40%, pelo menos 50%, pelo menos 60%, pelo menos 70%, pelo menos 80%, pelo menos 90% de inserção de um ácido nucleico heterólogo. Taxas de inserção podem ser medidas *in vitro* ou *in vivo*. Por exemplo, em algumas modalidades, a taxa de inserção pode ser determinada detectando e medindo-se o ácido nucleico inserido dentro de uma população de células, e calculando-se uma porcentagem da população que contém o ácido nucleico inserido. Os métodos para medir taxas de inserção são conhecidos e estão disponíveis na técnica. Em algumas modalidades, o RNA guia permite entre 5 e 10%, 10 e 15%, 15 e 20%, 20 e 25%, 25 e 30%, 30 e 35%, 35 e 40%, 40 e 45%, 45 e 50%, 50 e 55%, 55 e 60%, 60 e 65%, 65 e 70%, 70 e 75%, 75 e 80%, 80 e 85%, 85 e 90%, 90 e 95%, 95 e 99% ou mais de expressão aumentada de um gene heterólogo. A expressão aumentada de um gene heterólogo pode ser medida *in vitro* ou *in vivo*. Por exemplo, em algumas modalidades, a expressão aumentada pode ser determinada detectando e medindo-se o nível de polipeptídeo heterólogo e comparando-se o nível com o nível de polipeptídeo heterólogoantes, por exemplo, de tratar as células ou da administração a um indivíduo. Em algumas modalidades, a expressão aumentada pode ser determinada detectando e medindo-se o nível de polipeptídeo heterólogo e comparando-se o nível com um nível de polipeptídeo

conhecido, por exemplo, um nível normal do polipeptídeo em um indivíduo saudável.

[085] Cada uma das sequências guia mostradas na Tabela 1 podem compreender adicionalmente nucleotídeos adicionais para formar um crRNA, por exemplo, sendo que a seguinte sequência de nucleotídeos exemplificativa segue a sequência guia em sua extremidade 3': GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 300) na orientação 5' para 3'. Coordenadas genômicas são de acordo com o genoma de referência humano hg38. No caso de um sgRNA, as sequências guia acima podem compreender adicionalmente nucleotídeos adicionais para formar um sgRNA, por exemplo, em que a seguinte sequência de nucleotídeos exemplificativa segue a extremidade 3' da sequência guia:

GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 301) ou

GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 302) na orientação 5' para 3'.

TABELA 1: SEQUÊNCIAS DE RNA GUIA HUMANO E COORDENADAS CROMOSSÔMICAS

ID Guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU	chr4:73405113-73405133	2
G009851	AUGCAUUUGUUUCAAAAUAU	chr4:73405000-73405020	3
G009852	UGCAUUUGUUUCAAAAUAUU	chr4:73404999-73405019	4
G009857	AUUUAUGAGAUCAACAGCAC	chr4:73404761-73404781	5
G009858	GAUCAACAGCACAGGUUUUG	chr4:73404753-73404773	6
G009859	UUAAAUAAAGCAUAGUGCAA	chr4:73404727-73404747	7
G009860	UAAAGCAUAGUGCAAUGGAU	chr4:73404722-73404742	8
G009861	UAGUGCAAUGGAUAGGUCUU	chr4:73 404715-73 40473 5	9
G009866	UACUAAAACUUUAUUUUACU	chr4:73404452-73404472	10
G009867	AAAGUUGAACAAUAGAAAAA	chr4:73404418-73404438	11

ID Guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G009868	AAUGCAUAAUCUAAGUCAAA	chr4:73405013-73405033	12
G009874	UAAUAAAAUUCAAACAUCCU	chr4:73404561-73404581	13
G012747	GCAUCUUUAAAGAAUUAUUU	chr4:73404478-73404498	14
G012748	UUUGGCAUUUAUUUCUAAAA	chr4:73404496-73404516	15
G012749	UGUAUUUGUGAAGUCUUACA	chr4:73404529-73404549	16
G012750	UCCUAGGUAAAAAAAAAAAA	chr4:73404577-73404597	17
G012751	UAAUUUUCUUUUGCGCACUA	chr4:73404620-73404640	18
G012752	UGACUGAAACUUCACAGAAU	chr4:73404664-73404684	19
G012753	GACUGAAACUUCACAGAAUA	chr4:73404665-73404685	20
G012754	UUCAUUUUAGUCUGUCUUCU	chr4:73404803-73404823	21
G012755	AUUAUCUAAGUUUGAAUAUA	chr4:73404859-73404879	22
G012756	AAUUUUUAAAAUAGUAUUCU	chr4:73404897-73404917	23
G012757	UGAAUUAUUCUUCUGUUUAA	chr4:73404924-73404944	24
G012758	AUCAUCCUGAGUUUUUCUGU	chr4:73404965-73404985	25
G012759	UUACUAAAACUUUAUUUUAC	chr4:73404453-73404473	26
G012760	ACCUUUUUUUUUUUUUACCU	chr4:73404581-73404601	27
G012761	AGUGCAAUGGAUAGGUCUUU	chr4:73404714-73404734	28
G012762	UGAUUCCUACAGAAAAACUC	chr4:73404973-73404993	29
G012763	UGGGCAAGGGAAGAAAAAAA	chr4:73405094-73405114	30
G012764	CCUCACUCUUGUCUGGGCAA	chr4:73405107-73405127	31
G012765	ACCUCACUCUUGUCUGGGCA	chr4:73405108-73405128	32
G012766	UGAGCAACCUCACUCUUGUC	chr4:73405114-73405134	33

[086] O RNA guia pode compreender adicionalmente um trRNA. Em cada modalidade de composição e método descrita no presente documento, o crRNA e trRNA pode ser associado como um RNA único (sgRNA) ou pode estar em RNAs separados (dgRNA). No contexto de sgRNAs, os componentes de crRNA e trRNA podem ser ligados covalentemente, por exemplo, por meio de uma ligação de fosfodiéster ou outra ligação covalente. Em algumas modalidades, o sgRNA

compreende uma ou mais ligações entre nucleotídeos que não são uma ligação de fosfodiéster.

[087] Em cada uma das modalidades de composição, uso e método descritas no presente documento, o RNA guia pode compreender duas moléculas de RNA como um “RNA guia duplo” ou “dgRNA”. O dgRNA compreende uma primeira molécula de RNA que compreende um crRNA que compreende, por exemplo, uma sequência guia mostrada na Tabela 1, e uma segunda molécula de RNA que compreende um trRNA. A primeira e a segunda moléculas de RNA podem não ser ligadas covalentemente, porém podem formar um duplex de RNA por meio do pareamento de base entre porções do crRNA e do trRNA.

[088] Em cada uma das modalidades de composição, uso e método descritas no presente documento, o RNA guia pode compreender uma molécula de RNA única como um “RNA guia único” ou “sgRNA”. O sgRNA pode compreender um crRNA (ou uma porção do mesmo) que compreende uma sequência guia mostrada na Tabela 1 covalentemente ligada a um trRNA. O sgRNA pode compreender 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência guia mostrada na Tabela 1. Em algumas modalidades, o crRNA e o trRNA são ligados covalentemente por meio de um ligante. Em algumas modalidades, o sgRNA forma uma estrutura de alça em formato de grampo por meio do pareamento de base entre porções do crRNA e do trRNA. Em algumas modalidades, o crRNA e o trRNA são ligados covalentemente por meio de uma ou mais ligações que não são uma ligação de fosfodiéster.

[089] Em algumas modalidades, o trRNA pode compreender toda ou uma porção de uma sequência de trRNA derivada de um sistema de CRISPR/Cas de ocorrência natural. Em algumas modalidades, o trRNA compreende um trRNA de tipo selvagem truncado ou modificado. O comprimento do trRNA depende do sistema de CRISPR/Cas usado. Em algumas modalidades, o trRNA compreende ou consiste em 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90,

100 ou mais do que 100 nucleotídeos. Em algumas modalidades, o trRNA pode compreender determinadas estruturas secundárias, tais como, por exemplo, uma ou mais estruturas de grampo ou alça em formato de grampo, ou uma ou mais estruturas de protuberância.

[090] Em algumas modalidades, a sequência alvo ou região dentro do íntron 1 de um lócus de albumina humana (por exemplo, sequências de nucleotídeo correspondentes a uma região dentro da SEQ ID NO: 1) pode ser complementar à sequência guia do RNA guia. Em algumas modalidades, o grau de complementaridade ou identidade entre uma sequência guia de um RNA guia e sua sequência-alvo correspondente pode ser de pelo menos 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% ou 100%. Em algumas modalidades, a sequência-alvo e a sequência guia do gRNA podem ser 100% complementares ou idênticas. Em outras modalidades, a sequência-alvo e a sequência guia do gRNA podem conter pelo menos uma disparidade. Por exemplo, a sequência-alvo e a sequência guia do gRNA podem conter 1, 2, 3, 4 ou 5 disparidades, em que o comprimento total da sequência guia é de cerca de 20, ou 20. Em algumas modalidades, a sequência-alvo e a sequência guia do gRNA podem conter 1-4 disparidades, em que a sequência guia é de cerca de 20, ou 20 nucleotídeos.

[091] Conforme descrito e exemplificado no presente documento, os RNAs guia de albumina podem ser usados para inserir e expressar um gene heterólogo (por exemplo, um transgene) no íntron 1 de um gene de albumina. Assim, em algumas modalidades, a presente divulgação inclui composições que compreendem um ou mais RNAs guia (gRNA) que compreendem sequências guia que direcionam um agente de ligação de DNA guiado por RNA (por exemplo, Cas9) a uma sequência de DNA-alvo em um gene de albumina.

[092] Em algumas modalidades, uma composição ou formulação divulgada no presente documento compreende um mRNA que compreende um quadro de leitura

aberto (ORF) que codifica um agente de ligação de DNA guiado por RNA, tal como uma Cas nuclease, conforme descrito no presente documento. Conforme descrito abaixo, o mRNA que compreende uma Cas nuclease pode compreender uma Cas9 nuclease, como uma Cas9 nuclease de S. *pyogenes* que tem clivase, nickase e/ou atividade de ligação de DNA sítio-específica. Em algumas modalidades, o ORF que codifica uma nuclease de DNA guiada por RNA é um “ORF de agente de ligação de DNA guiado por RNA modificado” ou simplesmente um “ORF modificado” que é usado como abreviação para indicar que o ORF é modificado.

[093] ORFs de Cas9, que incluem ORFs de Cas9 modificados, são fornecidos no presente documento e são conhecidos na técnica. Como um exemplo, o ORF de Cas9 pode ser de códon otimizado, de modo que a sequência de codificação inclua um ou mais códons alternativos para um ou mais aminoácidos. Um “códon alternativo”, conforme usado no presente documento, se refere às variações na utilização de códon para um dado aminoácido, e pode ou não ser um códon preferencial ou otimizado (códon otimizado) para um dado sistema de expressão. Uso de códon preferencial, ou códons que são bem tolerados em um dado sistema de expressão, é conhecido na técnica. As sequências codificantes de Cas9, mRNAs de Cas9 e sequências de proteínas de Cas9 dos documentos nº WO2013/176772, WO2014/065596, WO2016/106121 e WO2019/067910 são incorporadas aqui a título de referência. Em particular, os ORFs e sequências de aminoácido de Cas9 da tabela no parágrafo [0449] do documento WO2019/067910, e os mRNAs e ORFs de Cas9 dos parágrafos [0214] a [0234] do documento nº WO2019/067910 são incorporados aqui a título de referência.

[094] Em algumas modalidades, um mRNA que compreende um ORF que codifica um agente de ligação de DNA guiado por RNA, tal como uma Cas nuclease, é fornecido, usado ou administrado.

B. GRNAS E MRNAS MODIFICADOS

[095] Em algumas modalidades, o gRNA é quimicamente modificado. Um gRNA que compreende um ou mais nucleosídeos ou nucleotídeos modificados é denominado um gRNA “modificado” ou gRNA “quimicamente modificado”, para descrever a presença de um ou mais componentes de ocorrência não natural e/ou de ocorrência natural ou configurações que são usadas em vez dos resíduos A, G, C e U canônicos ou adicionalmente aos mesmos. Em algumas modalidades, um gRNA modificado é sintetizado com um nucleosídeo ou nucleotídeo não canônico, aqui o mesmo é denominado “modificado”. Nucleosídeos e nucleotídeos modificados podem incluir um ou mais dentre: (i) alteração, por exemplo, substituição, de um ou ambos dentre os oxigênios de fosfato não ligantes e/ou de um ou mais dentre os oxigênios de fosfato ligantes na ligação de cadeia principal de fosfodiéster (uma modificação de cadeia principal exemplificativa); (ii) alteração, por exemplo, substituição, de um constituinte do açúcar de ribose, por exemplo, da 2’-hidroxila no açúcar de ribose (uma modificação de açúcar exemplificativa); (iii) substituição em grande quantidade da porção química de fosfato por ligantes “defosfo” (uma modificação de cadeia principal exemplificativa); (iv) modificação ou substituição de uma nucleobase de ocorrência natural que inclui por uma nucleobase não canônica (uma modificação de base exemplificativa); (v) substituição ou modificação da cadeia principal de ribose-fosfato (uma modificação de cadeia principal exemplificativa); (vi) modificação da extremidade 3’ ou extremidade 5’ do oligonucleotídeo, por exemplo, remoção, modificação ou substituição de um grupo fosfato terminal ou conjugação de uma porção química, cap ou ligante (tal como modificações de cap 3’ ou 5’ podem compreender uma modificação de açúcar e/ou cadeia principal); e (vii) modificação ou substituição do açúcar (uma modificação de açúcar exemplificativa).

[096] Modificações químicas, tais como aquelas listadas acima, podem ser combinadas para fornecer gRNAs e/ou mRNAs modificados que compreendem 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, toda base de um gRNA é modificada, por exemplo, todas as bases têm um grupo fosfato modificado, tal como um grupo fosforotioato. Em determinadas modalidades, todos, ou substancialmente todos os grupos fosfato de uma molécula de gRNA são substituídos por grupos fosforotioato. Em algumas modalidades, gRNAs modificados compreendem pelo menos um resíduo modificado na extremidade 5' do RNA ou próximo à mesma. Em algumas modalidades, gRNAs modificados compreendem pelo menos um resíduo modificado na extremidade 3' do RNA ou próximo à mesma. Determinados gRNAs compreendem pelo menos um resíduo modificado na extremidade 5' e extremidade 3' do RNA ou próximo às mesmas.

[097] 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.

[098] Ácidos nucleicos não modificados podem ser propensos à degradação, por exemplo, por nucleases intracelulares ou aquelas encontradas em soro. Por exemplo, nucleases podem hidrolisar ligações de fosfodiéster de ácido nucleico. Consequentemente, em um aspecto, os gRNAs descritos no presente documento podem conter um ou mais nucleosídeos ou nucleotídeos modificados, por exemplo, para introduzir estabilidade em direção às nucleases intracelulares ou à base de soro. Em algumas modalidades, as moléculas de gRNA modificadas descritas no presente documento podem exibir uma resposta imunológica inata reduzida quando introduzida

em uma população de células, tanto *in vivo* quanto *ex vivo*. O termo “resposta imunológica inata” inclui uma resposta celular aos ácidos nucleicos exógenos, que incluem ácidos nucleicos de fita simples, que envolve a indução de expressão e liberação de citocina, particularmente os interferons, e morte celular.

[099] Em algumas modalidades de uma modificação de cadeia principal, o grupo fosfato de um resíduo modificado pode ser modificado substituindo-se um ou mais dentre os oxigênios por um substituinte diferente. Adicionalmente, o resíduo modificado, por exemplo, resíduo modificado presente em um ácido nucleico modificado, pode incluir a substituição em grande quantidade de uma porção química de fosfato não modificada por um grupo fosfato modificado, conforme descrito no presente documento. Em algumas modalidades, a modificação de cadeia principal da cadeia principal de fosfato pode incluir alterações que resultam em um ligante não carregado ou um ligante carregado com distribuição de carga assimétrica.

[0100] Exemplos de grupos fosfato modificados incluem, fosforotioato, fosforosselenatos, fosfatos de borano, ésteres de fosfato de borano, fosfonatos de hidrogênio, fosforoamidatos, fosfonatos de alquila ou arila e fosfotriésteres. O átomo de fósforo em um grupo fosfato não modificado é aquiral. No entanto, a substituição de um dos oxigênios não interligados por um dos átomos ou grupos de átomos acima pode gerar o átomo de fósforo quiral. O átomo de fósforo estereogênico pode ter a configuração “R” (no presente documento Rp) ou a configuração “S” (no presente documento Sp). A cadeia principal também pode ser modificada pela substituição de um oxigênio de interligação, (isto é, o oxigênio que liga o fosfato ao nucleosídeo), por nitrogênio (fosforoamidatos interligados), enxofre (fosforotioatos interligados) e carbono (metilenofosfonatos interligados). A substituição pode ocorrer no oxigênio ligante ou em ambos oxigênios ligantes.

[0101] O grupo fosfato pode ser substituído por conectores que não contêm fósforo em determinadas modificações de cadeia principal. Em algumas modalidades,

o grupo fosfato carregado pode ser substituído por uma porção química neutra. Exemplos de porções químicas que podem substituir o grupo fosfato podem incluir, sem limitação, por exemplo, fosfonato de metila, hidroxilamino, siloxano, carbonato, carboximetila, carbamato, amida, tioéter, ligante de óxido de etileno, sulfonato, sulfonamida, tioformacetal, formacetal, oxima, metilenoimino, metilenometilimino, metileno-hidrazo, metilenodimetil-hidrazo e metileno-oximetilimino.

[0102] Arcabouços que podem imitar ácidos nucleicos também podem ser construídos, em que o ligante de fosfato e açúcar de ribose são substituídos por substitutos de nucleosídeo ou nucleotídeo resistentes à nuclease. Tais modificações podem compreender modificações de cadeia principal e açúcar. Em algumas modalidades, as nucleobases podem ser pressas por uma cadeia principal substituta. Exemplos podem incluir, sem limitação, o morfolino, ciclobutila, pirrolidina e substitutos de nucleosídeo de ácido nucleico peptídico (PNA).

[0103] Os nucleosídeos modificados e nucleotídeos modificados podem incluir uma ou mais modificações no grupo açúcar, isto é, na modificação de açúcar. Por exemplo, o grupo 2’ hidroxila (OH) pode ser modificado, por exemplo substituído por diversos substituintes “oxi” ou “desóxi” diferentes. Em algumas modalidades, as modificações no grupo 2’ hidroxila podem intensificar a estabilidade do ácido nucleico desde que a hidroxila não possa mais ser desprotonada para formar um íon 2’-alcóxido.

[0104] Exemplos de modificações de grupo 2’ hidroxila podem incluir alcóxi ou arilóxi (OR, em que “R” pode ser, por exemplo, alquila, cicloalquila, arila, aralquila, heteroarila ou um açúcar); polietilenoglicóis (PEG), $O(CH_2CH_2O)_nCH_2CH_2OR$, em que R pode ser, por exemplo, H ou opcionalmente alquila substituída, e n pode ser um número inteiro de 0 a 20 (por exemplo, de 0 a 4, de 0 a 8, de 0 a 10, de 0 a 16, de 1 a 4, de 1 a 8, de 1 a 10, de 1 a 16, de 1 a 20, de 2 a 4, de 2 a 8, de 2 a 10, de 2 a 16, de 2 a 20, de 4 a 8, de 4 a 10, de 4 a 16 e de 4 a 20). Em algumas modalidades, a

modificação de grupo 2’ hidroxila pode ser 2’-O-Me. Em algumas modalidades, a modificação de grupo 2’ hidroxila pode ser uma modificação de 2’-fluoro, que substitui o grupo 2’ hidroxila por um fluoreto. Em algumas modalidades, a modificação de grupo 2’ hidroxila pode ser um 2’-H, que substitui o grupo 2’ hidroxila por um hidrogênio. Em algumas modalidades, a modificação de grupo 2’ hidroxila pode incluir ácidos nucleicos “travados” (LNA) nos quais a 2’ hidroxila pode ser conectada, por exemplo, por

[0105] uma ponte de C_{1-6} alquileno ou C_{1-6} heteroalquileno, ao 4’-carbono do mesmo açúcar de ribose, em que pontes exemplificativas podem incluir pontes de metileno, propileno, éter ou amino; O-amino (em que amino pode ser, por exemplo, NH_2; alquilamino, dialquilamino, heterociclila, arilamino, diarilamino, heteroarilamino ou di-heteroarilamino, etilenodiamina ou poliamino) e aminoalcóxi, $O(CH_2)_n$-amino, (em que amino pode ser, por exemplo, NH_2; alquilamino, dialquilamino, heterociclila, arilamino, diarilamino, heteroarilamino ou di-heteroarilamino, etilenodiamina ou poliamino). Em algumas modalidades, a modificação de grupo 2’ hidroxila pode incluir ácidos nucleicos “destravados” (UNA) em que o anel de ribose carece da ligação C2’-C3’. Em algumas modalidades, a modificação de grupo 2’ hidroxila pode incluir o grupo metoxietila (MOE), ($OCH_2CH_2OCH_3$, por exemplo, um derivado de PEG).

[0106] Modificações de “desóxi” 2’ podem incluir hidrogênio (isto é, açúcares de desoxirribose, por exemplo, nas porções suspensas de dsRNA parciais); halo (por exemplo, bromo, cloro, fluoro, ou iodo); amino (em que amino pode ser, por exemplo, NH_2; alquilamino, dialquilamino, heterociclila, arilamino, diarilamino, heteroarilamino, di-heteroarilamino, ou aminoácido); $NH(CH_2CH_2NH)_nCH2CH_2$-amino (em que amino pode ser, por exemplo, conforme descrito no presente documento), -NHC(O)R (em que R pode ser, por exemplo, alquila, cicloalquila, arila, aralquila, heteroarila ou açúcar), ciano; mercapto; alquil-tio-alquila; tioalcóxi; e alquila, cicloalquila, arila, alquenila e alquinila, que podem ser opcionalmente substituídos por, por exemplo, um

amino, conforme descrito no presente documento.

[0107] A modificação de açúcar pode compreender um grupo açúcar que também pode conter um ou mais carbonos que têm a configuração estereoquímica oposta àquela do carbono correspondente em ribose. Assim, um ácido nucleico modificado pode incluir nucleotídeos que contêm, por exemplo, arabinose, como o açúcar. Os ácidos nucleicos modificados também podem incluir açúcares abásicos. Esses açúcares abásicos também podem ser modificados adicionalmente em um ou mais dentre os átomos de açúcar constituintes. Os ácidos nucleicos modificados também podem incluir um ou mais açúcares que estão na forma L, por exemplo L-nucleosídeos.

[0108] Os nucleosídeos modificados e nucleotídeos modificados descritos no presente documento, que podem ser incorporados em um ácido nucleico modificado, podem incluir uma base modificada, também denominada uma nucleobase. Exemplos de nucleobases incluem, porém, sem limitação, adenina (A), guanina (G), citosina (C) e uracila (U). Essas nucleobases podem ser modificadas ou completamente substituídas para fornecer resíduos modificados que podem ser incorporados nos ácidos nucleicos modificados. A nucleobase do nucleotídeo pode ser independentemente selecionada dentre uma purina, uma pirimidina, um análogo de purina ou análogo de pirimidina. Em algumas modalidades, a nucleobase pode incluir, por exemplo, derivados de ocorrência natural e sintéticos de uma base.

[0109] Em modalidades que empregam um RNA guia duplo, cada um dentre o crRNA e o tracrRNA pode conter modificações. Tais modificações podem ser em uma ou ambas extremidades do crRNA e/ou tracrRNA. Em modalidades que compreendem um sgRNA, um ou mais resíduos em uma ou ambas extremidades do sgRNA podem ser quimicamente modificados, e/ou nucleosídeos internos podem ser modificados, e/ou o sgRNA inteiro pode ser quimicamente modificado. Determinadas modalidades compreendem uma modificação de extremidade 5’. Determinadas

modalidades compreendem uma modificação de extremidade 3’.

[0110] Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem um dos padrões de modificação divulgados no documento nº WO2018/107028 A1, depositado em 8 de dezembro de 2017, intitulado “Chemically Modified Guide RNAs”, cujo conteúdo está incorporado aqui a título de referência em sua totalidade. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma das estruturas/padrões de modificação divulgados no documento nº US20170114334, cujo conteúdo está incorporado aqui a título de referência em sua totalidade. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma das estruturas/padrões de modificação divulgados no documento nº WO2017/136794, cujo conteúdo está incorporado aqui a título de referência em sua totalidade.

[0111] Em algumas modalidades, o sgRNA da presente divulgação compreende os padrões de modificação mostrados abaixo na Tabela 2. “Sequência Completa” na Tabela 2 se refere a uma sequência de sgRNA para cada um dos guias listado na Tabela 1. “Sequência Completa Modificada” mostra um padrão de modificação para cada sgRNA.

TABELA 2: SGRNA E PADRÕES DE MODIFICAÇÃO PARA SGRNA DE SEQUÊNCIAS GUIA DE ALBUMINA HUMANA

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	34	mG*mA*mG*CAACCUCACUCUUGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	66

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009851	AUGCAUUUGUUUCAAAAUAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	35	mA*mU*mG*CAUUUGUUUCAAAAUAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	67
G009852	UGCAUUUGUUUCAAAAUAUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	36	mU*mG*mC*AUUUGUUUCAAAAUAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	68
G009857	AUUUAUGAGAUCAACAGCACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	37	mA*mU*mU*UAUGAGAUCAACAGCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	69
G009858	GAUCAACAGCACAGGUUUUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	38	mG*mA*mU*CAACAGCACAGGUUUUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	70

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009859	UUAAAUAAAGCAUAGUGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	39	mU*mU*mA*AAUAAAGCAUAGUGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	71
G009860	UAAAGCAUAGUGCAAUGGAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	40	mU*mA*mA*AGCAUAGUGCAAUGGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	72
G009861	UAGUGCAAUGGAUAGGUCUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	41	mU*mA*mG*UGCAAUGGAUAGGUCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	73
G009866	UACUAAAACUUUAUUUUACUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	42	mU*mA*mC*UAAAACUUUAUUUUACUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	74

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009867	AAAGUUGAACAAUAGAAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	43	mA*mA*mA*GUUGAACAAUAGAAAAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	75
G009868	AAUGCAUAAUCUAAGUCAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	44	mA*mA*mU*GCAUAAUCUAAGUCAAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmA	76
G009874	UAAUAAAAUUCAAACAUCCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	45	mU*mA*mA*UAAAAUUCAAACAUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	77
G012747	GCAUCUUUAAAGAAUUAUUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	46	mG*mC*mA*UCUUUAAAGAAUUAUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	78
G012748	UUUGGCAUUUAUUUCUAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	47	mU*mU*mU*GGCAUUUAUUUCUAAAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	79

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012749	UGUAUUUGUGAAGUCUUACAGUU UUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCG GUGCUUUU	48	mU*mG*mU*AUUUGUGAAGUC UUACAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	80
G012750	UCCUAGGUAAAAAAAAAAAAGUU UUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCG GUGCUUUU	49	mU*mC*mC*UAGGUAAAAAAA AAAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	81
G012751	UAAUUUUCUUUUGCGCACUAGUU UUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCG GUGCUUUU	50	mU*mA*mA*UUUUCUUUUGCG CACUAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	82
G012752	UGACUGAAACUUCACAGAAUGUU UUAGAGCUAGAAAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCG GUGCUUUU	51	mU*mG*mA*CUGAAACUUCAC AGAAUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	83

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012753	GACUGAAACUUCACAGAAUAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	52	mG*mA*mC*UGAAACUUCACAGAAUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	84
G012754	UUCAUUUUAGUCUGUCUUCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	53	mU*mU*mC*AUUUUAGUCUGUCUUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	85
G012755	AUUAUCUAAGUUUGAAUAUAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	54	mA*mU*mU*AUCUAAGUUUGAAUAUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	86
G012756	AAUUUUUAAAAUAGUAUUCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	55	mA*mA*mU*UUUUAAAAUAGUAUUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	87

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012757	UGAAUUAUUCUUCUGUUUAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	56	mU*mG*mA*AUUAUUCUUCUGUUUAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	88
G012758	AUCAUCCUGAGUUUUUCUGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	57	mA*mU*mC*AUCCUGAGUUUUUCUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	89
G012759	UUACUAAAACUUUAUUUUACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	58	mU*mU*mA*CUAAAACUUUAUUUUACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	90
G012760	ACCUUUUUUUUUUUUUACCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	59	mA*mC*mC*UUUUUUUUUUUUUACCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	91

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012761	AGUGCAAUGGAUAGGUCUUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	60	mA*mG*mU*GCAAUGGAUAGGUCUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	92
G012762	UGAUUCCUACAGAAAAACUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	61	mU*mG*mA*UUCCUACAGAAAAACUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	93
G012763	UGGGCAAGGGAAGAAAAAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	62	mU*mG*mG*GCAAGGGAAGAAAAAAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	94
G012764	CCUCACUCUUGUCUGGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	63	mC*mC*mU*CACUCUUGUCUGGGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	95

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012765	ACCUCACUCUUGUCUGGGCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	64	mA*mC*mC*UCACUCUUGUCUGGGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	96
G012766	UGAGCAACCUCACUCUUGUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	65	mU*mG*mA*GCAACCUCACUCUUGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	97

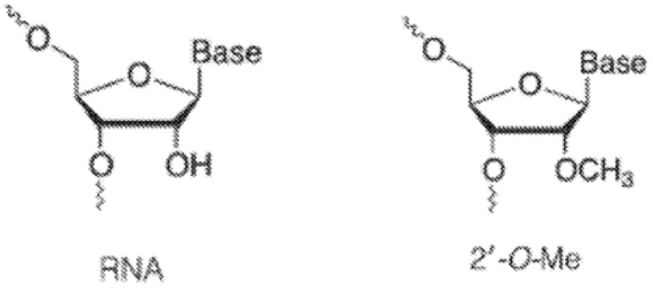
[0112] Em algumas modalidades, o sgRNA modificado compreende a seguinte sequência: mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU (SEQ ID NO: 350), em que “N” pode ser qualquer nucleotídeo natural ou não natural, e em que a totalidade de Ns compreende uma sequência guia de íntron 1 de albumina, conforme descrito na Tabela 1. Por exemplo, é engloba no presente documento a SEQ ID NO: 350, que omite os N da SEQ ID NO: 350, mas inclui a porção conservada modificada de um gRNA.

[0113] Qualquer uma das modificações descritas abaixo pode estar presente nos gRNAs e mRNAs descritos no presente documento.

[0114] Os termos “mA”, “mC”, “mU” ou “mG” podem ser usados para denotar

um nucleotídeo que foi modificado por 2’-O-Me.

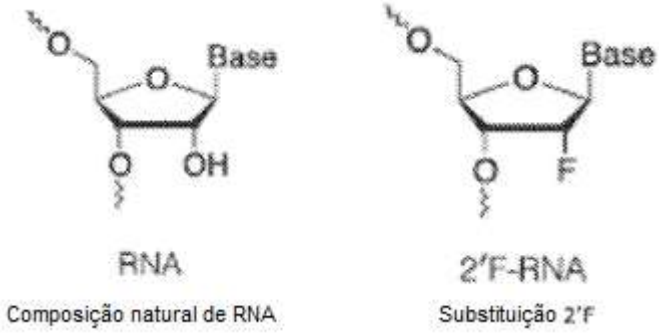
[0115] Modificação de 2’-O-metila pode ser representada como a seguir:



[0116] Outra modificação química que foi mostrada para influenciar anéis de açúcar de nucleotídeo é substituição de halogênio. Por exemplo, a substituição de 2’-fluoro (2’-F) nos anéis de açúcar de nucleotídeo podem aumentar afinidade de ligação de oligonucleotídeo e estabilidade de nuclease.

[0117] Neste pedido, os termos “fA”, “fC”, “fU” ou “fG” podem ser usados para denotar um nucleotídeo que foi substituído por 2’-F.

[0118] Substituição de 2’-F pode ser representada como a seguir:

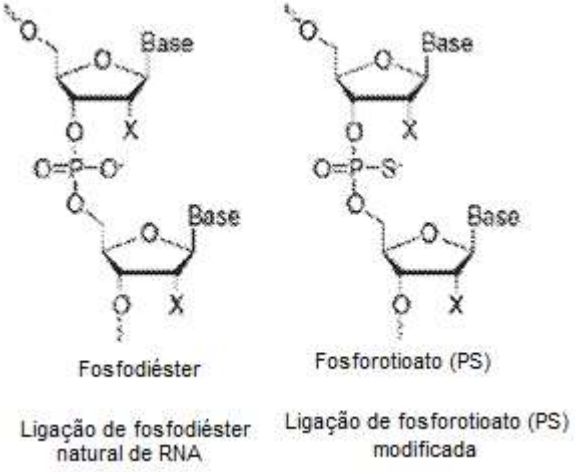


[0119] Interligação ou ligação de fosforotioato (PS) se refere a uma ligação em que um enxofre é substituído por um oxigênio de fosfato de não interligação em uma ligação fosfodiéster, por exemplo, nas ligações entre bases de nucleotídeos. Quando fosforotioatos são usados para gerar oligonucleotídeos, os oligonucleotídeos modificados também podem ser denominados S-oligos.

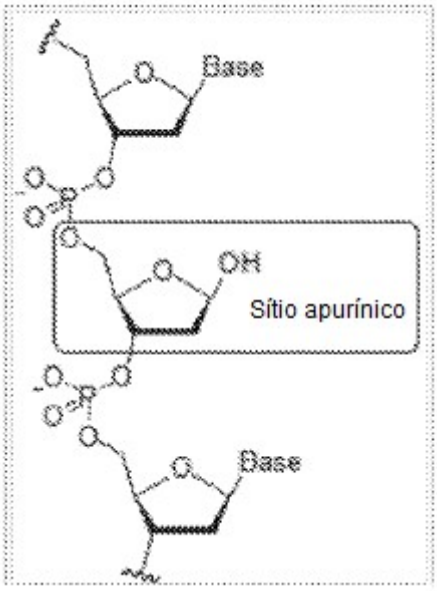
[0120] Um “*” pode ser usado para representar uma modificação de PS. Neste pedido, os termos A*, C*, U* ou G* podem ser usados para denotar um nucleotídeo que é ligado ao próximo (por exemplo, 3’) nucleotídeo por uma ligação de PS.

[0121] Neste pedido, os termos “mA*”, “mC*”, “mU*” ou “mG*” podem ser usados para denotar um nucleotídeo que foi substituído por 2’-O-Me e que é ligado ao próximo (por exemplo, 3’) nucleotídeo por uma ligação de PS.

[0122] O diagrama abaixo mostra a substituição de S- em um oxigênio de fosfato de não interligação, que gera uma ligação de PS no lugar de uma ligação de fosfodiéster:

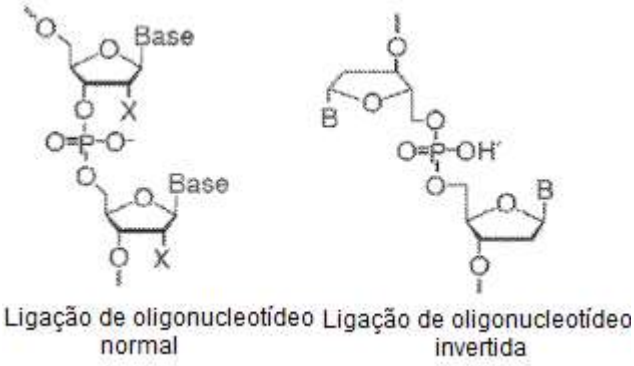


[0123] Nucleotídeos abásicos se referem àquelas que carecem de bases de nitrogenosas. A Figura abaixo representa um oligonucleotídeo com um sítio abásico (também conhecido como apurínico) que carece de uma base:



[0124] Bases invertidas se referem àquelas com ligações que são invertidas

a partir da ligação de 5’ para 3’ normal (isto é, uma ligação de 5’ para 5’ ou uma ligação de 3’ para 3’). Por exemplo:



Ligação de oligonucleotídeo normal Ligação de oligonucleotídeo invertida

[0125] Um nucleotídeo abásico pode ser fixado a uma ligação invertida. Por exemplo, um nucleotídeo abásico pode ser fixado ao nucleotídeo de terminal 5’ por meio de uma ligação de 5’ para 5’, ou um nucleotídeo abásico pode ser fixado ao nucleotídeo de terminal 3’ por meio de uma ligação de 3’ para 3’. Um nucleotídeo abásico invertido no nucleotídeo de terminal 5’ ou 3’ também pode ser denominado um cap de extremidade abásica invertida.

[0126] Em algumas modalidades, um ou mais dentre os primeiros três, quatro ou cinco nucleotídeos no terminal 5’, e um ou mais dentre os últimos três, quatro ou cinco nucleotídeos no terminal 3’ são modificados. Em algumas modalidades, a modificação é um 2’-O-Me, 2’-F, nucleotídeo abásico invertido, ligação de PS ou outra modificação de nucleotídeo bem conhecida na técnica para aumentar a estabilidade e/ou o desempenho.

[0127] Em algumas modalidades, os primeiros quatro nucleotídeos no terminal 5’ e os últimos quatro nucleotídeos no terminal 3’ são ligados às ligações de fosforotioato (PS).

[0128] Em algumas modalidades, os primeiros três nucleotídeos no terminal 5’, e os últimos três nucleotídeos no terminal 3’ compreendem um nucleotídeo modificado 2’-O-metila (2’-O-Me). Em algumas modalidades, os primeiros três nucleotídeos no terminal 5’, e os últimos três nucleotídeos no terminal 3’ compreendem um nucleotídeo modificado 2’-flúor (2’-F). Em algumas modalidades,

os primeiros três nucleotídeos no terminal 5’, e os últimos três nucleotídeos no terminal 3’ compreendem um nucleotídeo abásico invertido.

[0129] Em algumas modalidades, o RNA guia compreende um sgRNA modificado. Em algumas modalidades, o sgRNA compreende o padrão de modificação mostrado na SEQ ID NO: 350, em que N é qualquer nucleotídeo natural ou não natural, e em que a totalidade dos Ns compreende uma sequência guia que direciona uma nuclease para uma sequência-alvo no íntron 1 de albumina humana, por exemplo, conforme mostrado na Tabela 1.

[0130] Em algumas modalidades, o RNA guia compreende um sgRNA mostrado em qualquer uma das SEQ ID NOs: 34-97 ou 120-163). Em algumas modalidades, o RNA guia compreende um sgRNA mostrado em qualquer uma das SEQ ID NOs: 197-229. Em algumas modalidades, o RNA guia compreende um sgRNA que compreende qualquer uma das sequências guia das SEQ ID NOs: 2-33 ou 98-119 e os nucleotídeos da SEQ ID No: 301, em que os nucleotídeos da SEQ ID NO: 301 estão na extremidade 3’ da sequência guia, e em que o sgRNA pode ser modificado conforme mostrado na Tabela 2 ou SEQ ID NO: 350. Em algumas modalidades, o RNA guia compreende um sgRNA que compreende qualquer uma das sequências guia das SEQ ID NOs: 2-33 ou 197-229 e os nucleotídeos da SEQ ID NO: 301, em que os nucleotídeos da SEQ ID NO: 301 estão na extremidade 3’ da sequência guia, e em que o sgRNA pode ser modificado conforme mostrado na Tabela 2 ou SEQ ID NO: 350.

[0131] Conforme observado acima, em algumas modalidades, uma composição ou formulação divulgada no presente documento compreende um mRNA que compreende um quadro de leitura aberto (ORF) que codifica um agente de ligação de DNA guiado por RNA, tal como uma Cas nuclease, conforme descrito no presente documento. Em algumas modalidades, um mRNA que compreende um ORF que codifica um agente de ligação de DNA guiado por RNA, tal como uma Cas nuclease,

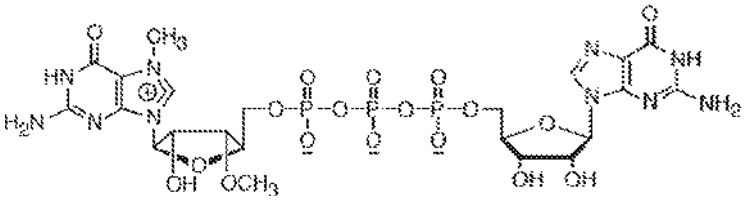
é fornecido, usado ou administrado. Em algumas modalidades, o ORF que codifica uma nuclease de DNA guiada por RNA é um “ORF de agente de ligação de DNA guiado por RNA modificado” ou simplesmente um “ORF modificado” que é usado como abreviação para indicar que o ORF é modificado.

[0132] Em algumas modalidades, o ORF modificado pode compreender uma uridina modificada pelo menos em uma, em uma pluralidade de ou em todas as posições de uridina. Em algumas modalidades, a uridina modificada é uma uridina modificada na posição 5, por exemplo, com um halogênio, metila ou etila. Em algumas modalidades, a uridina modificada é uma pseudouridina modificada na posição 1, por exemplo, com um halogênio, metila ou etila. A uridina modificada pode ser, por exemplo, pseudouridina, N1-metil-pseudouridina, 5-metoxiuridina, 5-iodouridina ou uma combinação das mesmas. 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-iodouridina. Em algumas modalidades, a uridina modificada é uma combinação de 5-iodouridina e 5-metoxiuridina.

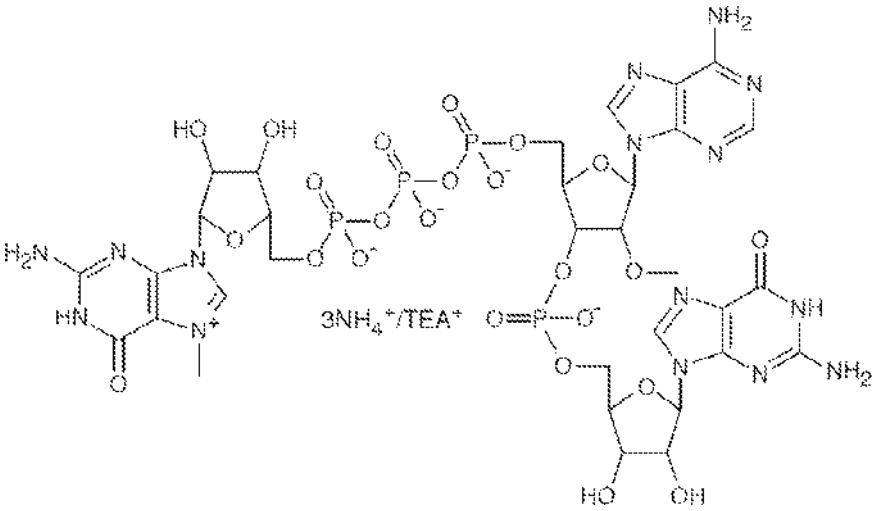
[0133] Em algumas modalidades, um mRNA divulgado no presente documento compreende um cap de 5’, tal como um Cap0, Cap1 ou Cap2. Um cap de 5’ é geralmente um ribonucleotídeo de 7-metilguanina (que pode ser modificado adicionalmente, conforme abordado abaixo, por exemplo, com relação ao ARCA)

ligado através de um 5’-trifosfato à posição 5’ do primeiro nucleotídeo da cadeia de 5’ para 3’ do mRNA, isto é, nucleotídeo próximo ao primeiro cap. Em Cap0, as riboses do primeiro e do segundo nucleotídeos próximos ao cap do mRNA compreendem ambas uma 2’-hidroxila. Em Cap1, as riboses do primeiro e do segundo nucleotídeos transcritos do mRNA compreendem um 2’-metóxi e uma 2’-hidroxila, respectivamente. Em Cap2, as riboses do primeiro e do segundo nucleotídeos próximos ao cap do mRNA compreendem ambas uma 2’-metóxi. Consultar, por exemplo, Katibah et al. (2014) *Proc Natl Acad Set EUA* 111(33): 12025-30; Abbas et al. *(2017) Proc Natl Acad Sci EUA* 114(11):E2106-E2115. A maioria dos mRNAs eucarióticos superiores endógenos, que incluem mRNAs de mamíferos, tais como mRNAs humanos, compreendem Cap1 ou Cap2. Cap0 e outras estruturas de cap que se diferem de Cap1 e Cap2 podem ser imunogênicos em mamíferos, tais como seres humanos, devido ao reconhecimento como “não próprio” pelos componentes do sistema imunológico inato, tal como IFIT-1 e IFIT-5, que pode resultar em níveis de citocina elevados que incluem interferon tipo I. Componentes do sistema imunológico inato, tais como IFIT-1 e IFIT-5 também podem competir com eIF4E pela ligação de um mRNA com um cap diferente de Cap1 ou Cap2, que inibe potencialmente a tradução do mRNA.

[0134] Um cap pode ser incluído cotranscricionalmente. Por exemplo, ARCA (análogo de cap antirreverso; Thermo Fisher Scientific nº de Cat. AM8045) é um análogo de cap que compreende um 3’-metoxi-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. ARCA resulta em um cap Cap0 no qual a posição 2’ do primeiro nucleotídeo próximo ao cap é hidroxila. Consultar, *por exemplo,* Stepinski et al., (2001) “Synthesis and properties of mRNAs containing the novel ‘antireverse’ cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'deoxy)GpppG”, *RNA* 7: 1486-1495. A estrutura de ARCA é mostrada abaixo.



[0135] CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies nº de Cat. N-7113) ou CleanCap™ GG (m7G(5')ppp(5')(2'OMeG)pG; TriLink Biotechnologies nº de Cat. N-7133) pode ser usado para fornecer uma estrutura de Cap1 cotranscricionalmente. Versões 3'-O-metiladas de CleanCap™ AG e CleanCap™ GG também estão disponíveis a partir da TriLink Biotechnologies como nº de Cat. N-7413 e N-7433, respectivamente. A estrutura de CleanCap™ AG é mostrada abaixo.



[0136] Alternativamente, um cap pode ser adicionado a um RNA pós-transcricionalmente. Por exemplo, a enzima de cap de Vaccinia está comercialmente disponível (New England Biolabs nº de Cat. M2080S) e tem atividades de trifosfatase e guanililtransferase de RNA fornecidas por sua subunidade de D1, e guanina metiltransferase, fornecida por sua subunidade D12. Dessa forma, é possível adicionar uma 7-metilguanina em um RNA, de modo a gerar Cap0 na presença de S-adenosil metionina e GTP. Consultar, por exemplo, Guo, P. e Moss, B. (1990) *Proc. Natl. Acad. Sci.* E.U.A. 87, 4023-4027; Mao, X. e Shuman, S. (1994) J. Biol. *Chem.* 269, 24472-24479.

[0137] Em algumas modalidades, o mRNA compreende adicionalmente uma

cauda poliadenilada (poli-A). Em algumas modalidades, a cauda poli-A compreende pelo menos 20, 30, 40, 50, 60, 70, 80, 90 ou 100 adeninas, opcionalmente até 300 adeninas. Em algumas modalidades, a cauda poli-A compreende 95, 96, 97, 98, 99 ou 100 nucleotídeos de adenina.

C. AGENTE DE LIGAÇÃO DE DNA GUIADO POR RNA

[0138] Conforme descrito no presente documento, os RNAs guia da presente divulgação são usados em combinação com um agente de ligação de DNA guiado por RNA para inserir e expressar um gene heterólogo (exógeno) em um lócus genômico, como um sítio de porto seguro, de uma célula hospedeira. O agente de ligação de DNA guiado por RNA pode ser uma proteína ou um ácido nucleico que codifica a proteína, como um mRNA. Em algumas modalidades, os métodos da presente divulgação incluem o uso de uma composição que compreende um RNA guia que compreende uma sequência guia da Tabela 1 e um agente de ligação de DNA guiado por RNA, por exemplo, uma nuclease, como uma Cas nuclease (por exemplo, Cas9), para formar um complexo de ribonucleoproteína.

[0139] Em algumas modalidades, o agente de ligação de DNA guiado por RNA, tal como uma Cas9 nuclease, tem atividade de clivase, que também pode ser denominada atividade de endonuclease de fita dupla. Em algumas modalidades, o agente de ligação de DNA guiado por RNA, tal como uma Cas9 nuclease, tem atividade de nickase, que também pode ser denominada atividade de endonuclease de fita simples. Em algumas modalidades, o agente de ligação de DNA guiado por RNA compreende uma Cas nuclease. Exemplos de Cas nucleases incluem aqueles dos sistemas CRISPR de tipo II de *S. pyogenes, S. aureus* e outros procariotas (consultar, por exemplo, a lista no próximo parágrafo) e versões de variante ou mutante (por exemplo, geneticamente modificada, de ocorrência não natural, de ocorrência natural ou outra variante) dos mesmos. Consultar, por exemplo, US2016/0312198 A1; US 2016/0312199 A1.

[0140] Espécies exemplificativas sem limitação das quais a Cas nuclease 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, Arthrospiraplatensis, 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.*

[0141] Em algumas modalidades, a Cas nuclease é a Cas9 nuclease de *Streptococcus pyogenes*. Em algumas modalidades, a Cas nuclease é a Cas9

nuclease de *Streptococcus thermophilus*. Em algumas modalidades, a Cas nuclease é a Cas9 nuclease de *Neisseria meningitidis*. Em algumas modalidades, a Cas nuclease é a Cas9 nuclease de *Staphylococcus aureus*. Em algumas modalidades, a Cas nuclease é a Cpf1 nuclease de *Francisella novicida*. Em algumas modalidades, a Cas nuclease é a Cpf1 nuclease de *Acidaminococcus sp.* Em algumas modalidades, a Cas nuclease é a Cpf1 nuclease de *Lachnospiraceae* bacterium ND2006. Em modalidades adicionais, a Cas nuclease é a Cpf1 nuclease de *Francisella tularensis, Lachnospiraceae bacterium, Butyrivibrio proteoclasticus, Peregrinibacteria bacterium, Parcubacteria bacterium, Smithella, Acidaminococcus, Candidatus Methanoplasma termitum, Eubacterium eligens, Moraxella bovoculi, Leptospira inadai, Porphyromonas crevioricanis, Prevotella disiens* ou *Porphyromonas macacae.* Em determinadas modalidades, a Cas nuclease é uma Cpf1 nuclease de um Acidaminococcus ou Lachnospiraceae.

[0142] Em algumas modalidades, o gRNA juntamente com um agente de ligação de DNA guiado por RNA é denominado um complexo de ribonucleoproteína (RNP). Em algumas modalidades, o agente de ligação de DNA guiado por RNA é uma Cas nuclease. Em algumas modalidades, o gRNA juntamente com uma Cas nuclease é denominado uma RNP de Cas. Em algumas modalidades, a RNP compreende componentes de Tipo I, Tipo II ou Tipo III. Em algumas modalidades, a Cas nuclease é a proteína Cas9 do sistema de CRISPR/Cas de Tipo II. Em algumas modalidades, o gRNA juntamente com Cas9 é denominado uma RNP de Cas9.

[0143] Cas9 de tipo selvagem tem dois domínios de nuclease: RuvC e HNH. O domínio RuvC cliva a fita de DNA não-alvo, e o domínio HNH cliva a fita-alvo de DNA. Em algumas modalidades, a proteína de Cas9 compreende mais que um domínio RuvC e/ou mais que um domínio HNH. Em algumas modalidades, a proteína Cas9 é uma Cas9 de tipo selvagem. Em cada uma das modalidades de composição, uso e método, a Cas induz uma quebra de fita dupla no DNA-alvo.

[0144] Em algumas modalidades, Cas nucleases quiméricas são usadas, quando um domínio ou região da proteína é substituído por uma porção de uma proteína diferente. Em algumas modalidades, um domínio de Cas nuclease pode ser substituído por um domínio de uma nuclease diferente, tal como FokI. Em algumas modalidades, uma Cas nuclease pode ser uma nuclease modificada.

[0145] Em outras modalidades, a Cas nuclease pode ser de um sistema de CRISPR/Cas de Tipo I. Em algumas modalidades, a Cas nuclease pode ser um componente do complexo em Cascata de um sistema de CRISPR/Cas de Tipo I. Em algumas modalidades, a Cas nuclease pode ser uma proteína Cas3. Em algumas modalidades, a Cas nuclease pode ser de um sistema de CRISPR/Cas de Tipo III. Em algumas modalidades, a Cas nuclease pode ter uma atividade de clivagem de RNA.

[0146] Em algumas modalidades, o agente de ligação de DNA guiado por RNA tem atividade de nickase de fita simples, isto é, pode cortar uma fita de DNA para produzir uma quebra de fita simples, também conhecida como um “corte”. Em algumas modalidades, o agente de ligação de DNA guiado por RNA compreende uma Cas nickase. Uma nickase é uma enzima que cria um corte em dsDNA, isto é, corta uma fita, porém, não a outra da hélice dupla de DNA. Em algumas modalidades, uma Cas nickase é uma versão de uma Cas nuclease (por exemplo, uma Cas nuclease abordada acima) na qual um sítio ativo endonucleolítico é inativado, por exemplo, através de uma ou mais alterações (por exemplo, mutações pontuais) em um domínio catalítico. Consultar, por exemplo, a Patente US nº 8.889.356 para a discussão de Cas nickases e alterações de domínio catalítico exemplificativo. Em algumas modalidades, uma Cas nickase, tal como uma Cas9 nickase tem um domínio RuvC ou HNH inativado.

[0147] Em algumas modalidades, o agente de ligação de DNA guiado por RNA é modificado para conter apenas um domínio de nuclease funcional. Por exemplo, a proteína de agente pode ser modificada de modo que um dos domínios de

nuclease sofra mutação ou seja completa ou parcialmente deletado para reduzir sua atividade de clivagem de ácido nucleico. Em algumas modalidades, uma nickase é usada com um domínio RuvC com atividade reduzida. Em algumas modalidades, uma nickase é usada com um domínio RuvC inativo. Em algumas modalidades, uma nickase é usada com um domínio HNH com atividade reduzida. Em algumas modalidades, uma nickase é usada com um domínio HNH inativo.

[0148] Em algumas modalidades, um aminoácido conservado dentro de um domínio de nuclease de proteína Cas é substituído para reduzir ou alterar a atividade de nuclease. Em algumas modalidades, uma Cas nuclease pode compreender uma substituição de aminoácido no domínio de nuclease semelhante a RuvC ou RuvC. Substituições de aminoácido exemplificativas no domínio de nuclease semelhante a RuvC ou RuvC incluem D10A (com base na proteína Cas9 de *S. pyogenes*). Consultar, por exemplo, Zetsche et al. (2015) *Cell* Oct 22:163(3): 759-771. Em algumas modalidades, a Cas nuclease pode compreender uma substituição de aminoácido no domínio de nuclease semelhante a HNH ou HNH. Substituições de aminoácido exemplificativas no domínio de nuclease semelhante a HNH ou HNH incluem E762A, H840A, N863A, H983A e D986A (com base na proteína Cas9 de *S. pyogenes*). Consultar, por exemplo, Zetsche et al. (2015). Substituições de aminoácido exemplificativas adicionais incluem D917A, E1006A e D1255A (com base na sequência de Francisella novicida U112 Cpf1 (FnCpf1) (UniProtKB - A0Q7Q2 (CPF1_FRATN)).

[0149] Em algumas modalidades, uma nickase é fornecida em combinação com um par de RNAs guia que são complementares às fitas de sentido e antissenso da sequência-alvo, respectivamente. Nessa modalidade, os RNAs guia direcionam a nickase a uma sequência-alvo e introduzem uma DSB gerando-se um corte em fitas opostas da sequência-alvo (isto é, corte duplo). Em algumas modalidades, uma nickase é usada juntamente com dois RNAs guia separados que alvejam fitas opostas

de DNA para produzir um corte duplo no DNA-alvo. Em algumas modalidades, uma nickase é usada juntamente com dois RNAs guia separados que são selecionados de modo a estarem em proximidade estreita para produzir um corte duplo no DNA-alvo. Em algumas modalidades, o agente de ligação de DNA guiado por RNA compreende um ou mais domínios funcionais heterólogos (por exemplo, é ou compreende um polipeptídeo de fusão).

[0150] Em algumas modalidades, o domínio funcional heterólogo pode facilitar o transporte do agente de ligação de 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 de DNA guiado por RNA pode ser fundido com 1-10 NLS (ou NLSs). Em algumas modalidades, o agente de ligação de DNA guiado por RNA pode ser fundido com 1-5 NLS (ou NLSs). Em algumas modalidades, o agente de ligação de 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 de agente de ligação de DNA guiado por RNA. O mesmo também pode ser inserido na sequência de agente de ligação de DNA guiado por RNA. Em outras modalidades, o agente de ligação de DNA guiado por RNA pode ser fundido com mais de um NLS. Em algumas modalidades, o agente de ligação de DNA guiado por RNA pode ser fundido com 2, 3, 4 ou 5 NLSs. Em algumas modalidades, o agente de ligação de DNA guiado por RNA pode ser fundido com dois NLSs. Em determinadas circunstâncias, os dois NLSs podem ser iguais (por exemplo, dois NLSs de SV40) ou diferentes. Em algumas modalidades, o agente de ligação de DNA guiado por RNA é fundido com duas sequências de NLS de SV40 ligadas no terminal carbóxi. Em algumas modalidades, o agente de ligação de 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 de DNA guiado por RNA pode ser fundido com 3 NLSs. Em algumas modalidades, o agente de ligação de DNA guiado por RNA pode

não ser fundido com NLS. Em algumas modalidades, o NLS pode ser uma sequência de monopartida, tal como, por exemplo, o NLS de SV40, PKKKRKV (SEQ ID NO: 600) ou PKKKRRV (SEQ ID NO: 601). Em algumas modalidades, o NLS pode ser uma sequência bipartida, tal como o NLS de nucleoplasmina, KRPAATKKAGQAKKKK (SEQ ID NO: 602). Em uma modalidade específica, um único NLS de PKKKRKV (SEQ ID NO: 600) pode ser ligado no terminal C do agente de ligação de DNA guiado por RNA. Um ou mais ligantes são opcionalmente incluídos no sítio de fusão.

D. CONSTRUTO DOADOR/SEQUÊNCIAS

[0151] As composições e métodos descritos no presente documento incluem o uso de um construto de ácido nucleico que compreende uma sequência que codifica um gene heterólogo a ser inserido em um sítio de corte criado por um RNA guia da presente divulgação e um agente de ligação de DNA guiado por RNA. Conforme usado no presente documento, tal construto é, por vezes, denominado um “construto doador/modelo”. Os construtos podem codificar qualquer ácido nucleico expressado (isto é, ácido nucleico que pode ser expresso), por exemplo, DNA, RNA mensageiro (mRNA), um RNA funcional, RNA interferente pequeno (siRNA), microRNA (miRNA), RNA de fita simples (ssRNA), RNAs sem codificação longos ou oligonucleotídeos antissenso.

[0152] As composições e métodos descritos no presente documento incluem o uso de um construto não bidirecional ou unidirecional, por exemplo, que codifica um transgene único, que codifica dois transgenes em *cis,* etc. O construto unidirecional pode compreender uma sequência de codificação ligada a um aceitante de splice.

[0153] As composições e métodos descritos no presente documento incluem o uso de um construto bidirecional descrito no presente documento que compreende pelo menos dois segmentos de ácido nucleico em *cis,* em que um segmento (o primeiro segmento) compreende uma sequência de codificação ou transgene”, embora o outro segmento (o segundo segmento) compreenda uma sequência em que

o complemento da sequência codifica um transgene. Um construto bidirecional pode compreender uma primeira sequência de codificação que codifica um gene heterólogo ligado a um aceitante de splice e uma segunda sequência de codificação, em que o complemento codifica um gene heterólogo na outra orientação, também ligado a um aceitante de splice.

[0154] Em algumas modalidades, os construtos divulgados no presente documento compreendem um sítio de aceitante de splice em qualquer uma das ou ambas as extremidades do construto, por exemplo, 5’ de um quadro de leitura aberto no primeiro e/ou no segundo segmentos, ou 5’ de uma ou ambas as sequências de transgene. Em algumas modalidades, o sítio de aceitante de splice compreende NAG. Em modalidades adicionais, o sítio de aceitante de splice consiste em NAG. Em algumas modalidades, o aceitante de splice é um aceitante de splice de albumina, por exemplo, um aceitante de splice de albumina usado no splicing em conjunto de éxons 1 e 2 de albumina. Em algumas modalidades, o aceitante de splice é derivado do gene de albumina humana. Em algumas modalidades, o aceitante de splice é derivado do gene de albumina de camundongo. Em algumas modalidades, o aceitante de splice é um aceitante de splice de F9 (ou “FIX”), por exemplo, o aceitante de splice de F9 usado no splicing em conjunto de éxons 1 e 2 de F9. Em algumas modalidades, o aceitante de splice é derivado do gene de F9 humano. Em algumas modalidades, o aceitante de splice é derivado do gene de F9 de camundongo. Sítios de aceitante de splice adequados adicionais úteis em eucariontes, que incluem aceitantes de splice artificiais, são conhecidos e podem ser derivados da técnica. Consultar, por exemplo, Shapiro, et al., 1987, Nucleic Acids Res., 15, 7155-7174, Burset, et al, 2001, Nucleic Acids Res., 29, 255-259.

[0155] Em algumas modalidades, a sequência de cauda de poliadenilação é codificada, por exemplo, como uma faixa “poli-A”, na extremidade 3’ do primeiro e/ou do segundo segmento. Em algumas modalidades, uma sequência de cauda de

poliadenilação é fornecida cotranscricionalmente como um resultado de uma sequência de sinal de poliadenilação que é codificada na extremidade 3’ ou próxima à mesma, do primeiro e/ou do segundo segmento. Métodos para projetar uma sequência de cauda de poliadenilação e/ou sequência de sinal de poliadenilação adequada são bem conhecidos na técnica. Sequências de aceitante de splice adequadas são divulgadas e exemplificadas no presente documento, que incluem sítios de aceitante de splice de albumina de camundongo e FIX humano. Em algumas modalidades, a sequência de sinal de poliadenilação AAUAAA (SEQ ID NO: 800) é normalmente usada em sistemas de mamífero, embora variantes, tais como UAUAAA (SEQ ID NO: 801) ou AU/GUAAA (SEQ ID NO: 802) tenham sido identificadas. Consultar, por exemplo, NJ Proudfoot, Genes & Dev. 25(17): 1770-82, 2011. Em algumas modalidades, uma sequência de cauda poliA é incluída. O comprimento do construto pode variar, dependendo do tamanho do gene a ser inserido, e pode ser, por exemplo, de 200 pares de base (pb) a cerca de 5000 pb, tal como cerca de 200 pb a cerca de 2000 pb, tal como cerca de 500 pb a cerca de 1500 pb. Em algumas modalidades, o comprimento do modelo doador de DNA é de cerca de 200 pb, ou é de cerca de 500 pb, ou é de cerca de 800 pb, ou é de cerca de 1000 pares de base, ou é de cerca de 1500 pares de base. Em outras modalidades, o comprimento do modelo doador é de pelo menos 200 pb, ou é de pelo menos 500 pb, ou é de pelo menos 800 pb, ou é de pelo menos 1000 pb, ou é de pelo menos 1500 pb.

[0156] O construto pode ser DNA ou RNA, de fita simples, fita dupla ou de fita parcialmente simples e parcialmente dupla e pode ser introduzido em uma célula hospedeira em forma linear ou circular (por exemplo, minicírculo). Consultar, por exemplo, as Publicações de Patente nº U.S. 2010/0047805, 2011/0281361, 2011/0207221. Caso introduzidas em forma linear, as extremidades da sequência doadora podem ser protegidas (por exemplo, de degradação exonucleolítica) por métodos conhecidos por aqueles versados na técnica. Por exemplo, um ou mais

resíduos de didesoxinucleotídeo são adicionados ao terminal 3’ de uma molécula linear e/ou oligonucleotídeos autocomplementares são ligados a uma ou ambas extremidades. Consultar, por exemplo, Chang et al. (1987) *Proc. Natl. Acad. Sci.* E.U.A. 84:4959-4963; Nehls et al. (1996) *Science* 272:886-889). Métodos adicionais para proteger polinucleotídeos exógenos da degradação incluem, porém, sem limitação, adição de grupo (ou grupos) amino terminal e o uso de ligações internucleotídeos modificadas, tais como, por exemplo, fosforotioatos, fosforamidatos, e resíduos de ribose de O-metila ou desoxirribose. Um construto pode ser introduzido em uma célula como parte de uma molécula de vetor que tem sequências adicionais, tais como, por exemplo, origens de replicação, promotores e genes que codificam resistência antibiótica. Um construto pode omitir elementos virais. Além disso, construtos doadores podem ser introduzidos como ácido nucleico nu, como ácido nucleico complexado com um agente, tal como um lipossoma ou poloxâmero, ou podem ser entregues por vírus (por exemplo, adenovírus, AAV, herpesvírus, retrovírus, lentivírus).

[0157] Em algumas modalidades, embora não sejam necessários para a expressão, os construtos divulgados no presente documento também podem incluir sequências reguladoras transcricionais ou translacionais, por exemplo, promotores, intensificadores, isolantes, sítios de entrada de ribossomo internos, sequências que codificam peptídeos e/ou sinais de poliadenilação.

[0158] Em algumas modalidades, os construtos que compreendem uma sequência de codificação para um polipeptídeo de interesse podem incluir uma ou mais dentre as seguintes modificações: otimização de códon (por exemplo, para códons humanos) e/ou adição de um ou mais sítios de glicosilação. Consultar, por exemplo, McIntosh et al. (2013) *Blood* (17):3335-44.

[0159] Em algumas modalidades, o construto pode ser inserido de modo que sua expressão seja acionada pelo promotor endógeno no sítio de inserção (por

exemplo, o promotor de albumina endógena quando o doador é integrado no lócus de albumina da célula hospedeira). Em tais casos, o transgene pode carecer de elementos de controle (por exemplo, promotor e/ou intensificador) que acionam sua expressão (por exemplo, um construto sem promotor). Não obstante, será aparente que, em outros casos, o construto pode compreender um promotor e/ou intensificador, por exemplo, um promotor constitutivo ou um promotor induzível ou de tecido específico (por exemplo, específico para fígado ou plaqueta) que aciona expressão da proteína funcional mediante integração. O construto pode compreender uma sequência que codifica uma proteína heteróloga a jusante de e operacionalmente ligada a uma sequência de sinal que codifica um peptídeo de sinal, por exemplo, um peptídeo de sinal de albumina, um peptídeo de sinal de uma proteína secretada por hepatócito. O construto pode compreender uma sequência que codifica uma proteína heteróloga a jusante de uma sequência de sinal que codifica um peptídeo de sinal e operacionalmente ligada à mesma a partir da proteína heteróloga. Em algumas modalidades, o construto de ácido nucleico funciona na inserção independente de homologia de um ácido nucleico que codifica uma proteína transgênica. Em algumas modalidades, o construto de ácido nucleico funciona em células sem divisão, por exemplo, células nas quais nem NHEJ, nem HR, são o mecanismo primário pelo qual a quebra de DNA de fita dupla é reparada. O ácido nucleico pode ser um construto doador independente de homologia.

[0160] O construto pode ser um construto de ácido nucleico bidirecional que compreende pelo menos dois segmentos de ácido nucleico, em que um segmento (o primeiro segmento) compreende uma sequência de codificação que codifica um agente de interesse (a sequência de codificação pode ser denominada no presente documento “transgene” ou um primeiro transgene), enquanto o outro segmento (o segundo segmento) compreende uma sequência, em que o complemento da sequência codifica um agente de interesse, ou um segundo transgene. Em algumas

modalidades, uma sequência de codificação codifica um agente terapêutico, como um polipeptídeo, RNA funcional ou intensificador. Os pelo menos dois segmentos podem codificar polipeptídeos idênticos ou diferentes ou agentes idênticos ou diferentes. Em algumas modalidades, os construtos bidirecionais divulgados no presente documento compreendem pelo menos dois segmentos de ácido nucleico, em que um segmento (o primeiro segmento) compreende uma sequência de codificação que codifica um polipeptídeo de interesse, enquanto o outro segmento (o segundo segmento) compreende uma sequência em que o complemento da sequência codifica um polipeptídeo de interesse. Quando usado em combinação com um sistema de edição de gene conforme descrito no presente documento, a bidirecionalidade dos construtos de ácido nucleico permite que o construto seja inserido em qualquer direção (não é limitada à inserção em uma direção) dentro de um sítio de inserção-alvo, que permite a expressão do polipeptídeo de interesse de a) uma sequência de codificação de um segmento (por exemplo, o segmento esquerdo que codifica ”F9 Humano” do construto de ssAAV esquerdo superior da Figura 1), ou 2) um complemento do outro segmento (por exemplo, o complemento do segmento direito que codifica “F9 Humano” indicado de cabeça para baixo no construto de ssAAV esquerdo superior da Figura 1), desse modo se intensifica a eficiência de inserção e expressão, conforme exemplificado no presente documento. A clivagem alvejada por um sistema de edição de gene pode facilitar a integração de construto e/ou expressão de transgene. Vários sistemas de edição de gene conhecidos podem ser usados na prática da presente divulgação, que incluem, por exemplo, por exemplo, sistemas de clivagem de DNA sítio-específicos, incluindo um sistema de CRISPR/Cas; sistema de nuclease de dedo de zinco (ZFN); ou sistema de nuclease efetora semelhante ao ativador de transcrição (TALEN).

[0161] Em algumas modalidades, o construto de ácido nucleico bidirecional não compreende um promotor que aciona a expressão do agente ou polipeptídeo. Por exemplo, a expressão do polipeptídeo é acionada por um promotor da célula

hospedeira (por exemplo, o promotor de albumina endógena quando o transgene é integrado em um lócus de albumina da célula hospedeira). Em algumas modalidades, o construto de ácido nucleico bidirecional inclui um primeiro segmento e um segundo segmento, sendo que cada um tem um aceitante de splice a montante de um transgene. Em determinadas modalidades, o aceitante de splice é compatível com a sequência doadora de splice do sítio de porto seguro da célula hospedeira, por exemplo, o doador de splice de íntron 1 de um gene de albumina humana.

[0162] Em algumas modalidades, o construto de ácido nucleico bidirecional compreende um primeiro segmento que compreende uma sequência de codificação para um polipeptídeo e um segundo segmento que compreende um complemento reverso de uma sequência do polipeptídeo. O mesmo é verdadeiro para agentes não polipeptídicos. Assim, a sequência de codificação no primeiro segmento tem capacidade para expressar um polipeptídeo, embora o complemento do complemento reverso no segundo segmento também tem capacidade para expressar o polipeptídeo. Conforme usado no presente documento, “sequência de codificação” ao se referir ao segundo segmento que compreende uma sequência de complemento reverso se refere à fita complementar (codificação) do segundo segmento (isto é, a sequência de codificação de complemento da sequência de complemento reverso no segundo segmento).

[0163] Em algumas modalidades, a sequência de codificação que codifica o Polipeptídeo A no primeiro segmento é menos que 100% complementar ao complemento reverso de uma sequência de codificação que também codifica o Polipeptídeo A. Isto é, em algumas modalidades, o primeiro segmento compreende uma sequência de codificação (1) para o Polipeptídeo A e o segundo segmento é um complemento reverso de uma sequência de codificação (2) para o Polipeptídeo A, em que a sequência de codificação (1) não é idêntica à sequência de codificação (2). Por exemplo, a sequência de codificação (1) e/ou sequência de codificação (2) que

codifica o Polipeptídeo A pode utilizar códons diferentes. Em algumas modalidades, uma ou ambas as sequências podem ser de códon otimizado, de modo que sequência de codificação (1) e o complemento reverso de sequência de codificação (2) possua 100% ou menos que 100% de complementaridade. Em algumas modalidades, a sequência de codificação do segundo segmento codifica o polipeptídeo com o uso de um ou mais códons alternativos para um ou mais aminoácidos do mesmo polipeptídeo codificado pela sequência de codificação no primeiro segmento. Um “códon alternativo”, conforme usado no presente documento, se refere às variações na utilização de códon para um dado aminoácido, e pode ou não ser um códon preferencial ou otimizado (códon otimizado) para um dado sistema de expressão. Usos de códon preferenciais, ou códons que são bem tolerados em um dado sistema de expressão, são conhecidos na técnica.

[0164] Em algumas modalidades, o segundo segmento compreende uma sequência de complemento reverso que adota uso de códon diferente daquele da sequência de codificação do primeiro segmento de modo a reduzir formação de grampo. Tal complemento reverso forma pares de base com menos do que todos os nucleotídeos da sequência de codificação no primeiro segmento, porém, ainda codifica, opcionalmente, o mesmo polipeptídeo. Em tais casos, a sequência de codificação, por exemplo, para o Polipeptídeo A, do primeiro segmento pode ser homóloga, porém, não idêntica à sequência de codificação, por exemplo, para o Polipeptídeo A da segunda metade do construto bidirecional. Em algumas modalidades, o segundo segmento compreende uma sequência de complemento reverso que não é substancialmente complementar (por exemplo, não mais que 70% complementar) à sequência de codificação no primeiro segmento. Em algumas modalidades, o segundo segmento compreende uma sequência de complemento reverso que é altamente complementar (por exemplo, pelo menos 90% complementar) à sequência de codificação no primeiro segmento. Em algumas modalidades, o

segundo segmento compreende uma sequência de complemento reverso que tem pelo menos cerca de 30%, cerca de 35%, cerca de 40%, cerca de 45%, cerca de 50%, cerca de 55%, cerca de 60%, cerca de 65%, cerca de 70%, cerca de 75%, cerca de 80%, cerca de 85%, cerca de 90%, cerca de 95%, cerca de 97% ou cerca de 99% de complementaridade com a sequência de codificação no primeiro segmento.

[0165] Em algumas modalidades, o segundo segmento compreende uma sequência de complemento reverso que tem 100% de complementaridade com a sequência de codificação no primeiro segmento. Isto é, a sequência no segundo segmento é um complemento reverso perfeito da sequência de codificação no primeiro segmento. A título de exemplo, o primeiro segmento compreende uma sequência hipotética 5'CTGGACCGA 3' (SEQ ID NO: 500) e o segundo segmento compreende o complemento reverso da SEQ ID NO: 1 - isto é, 5' TCGGTCCAG 3' (SEQ ID NO: 502).

[0166] Em algumas modalidades, o construto de ácido nucleico bidirecional compreende um primeiro segmento que compreende uma sequência de codificação para um polipeptídeo ou agente (por exemplo, um primeiro polipeptídeo) e um segundo segmento que compreende um complemento reverso de uma sequência de codificação de um polipeptídeo ou agente (por exemplo, um segundo polipeptídeo). Em algumas modalidades, o primeiro polipeptídeo e o segundo polipeptídeo são iguais, conforme descrito acima. Em algumas modalidades, o primeiro agente terapêutico e o segundo agente terapêutico são iguais, conforme descrito acima. Em algumas modalidades, o primeiro polipeptídeo e o segundo polipeptídeo são diferentes. Em algumas modalidades, o primeiro agente terapêutico e o segundo agente terapêutico são diferentes. Por exemplo, o primeiro polipeptídeo é Polipeptídeo A e o segundo polipeptídeo é Polipeptídeo B. Como exemplo adicional, o primeiro polipeptídeo é Polipeptídeo A e o segundo polipeptídeo é uma variante (por exemplo, um fragmento (como um fragmento funcional), mutante, fusão (incluindo a adição de

no mínimo um aminoácido em um terminal de polipeptídeo), ou combinações dos mesmos) de Polipeptídeo A. Uma sequência de codificação que codifica um polipeptídeo pode compreender opcionalmente uma ou mais sequências adicionais, como sequências que codificam sequências de aminoácidos amino ou carbóxi terminais, como uma sequência de sinal, sequência de marcação (por exemplo, HiBit), ou sequência funcional heteróloga (por exemplo, sequência de localização nuclear (NLS) ou peptídeo de autoclivagem) ligada ao polipeptídeo. Uma sequência de codificação que codifica um polipeptídeo pode compreender, opcionalmente, sequências que codificam uma ou mais sequências de peptídeo de sinal de terminal amino. Cada uma dessas sequências adicionais pode ser igual ou diferente no primeiro segmento e no segundo segmento do construto.

[0167] O construto bidirecional descrito no presente documento pode ser usado para expressar qualquer polipeptídeo de acordo com os métodos divulgados no presente documento. Em algumas modalidades, o polipeptídeo é um polipeptídeo secretado. Em algumas modalidades, o polipeptídeo é um cuja função é normalmente efetuada (por exemplo, funcionalmente ativo) como um polipeptídeo secretado. Um “polipeptídeo secretado” conforme usado no presente documento se refere a uma proteína que é secretada pela célula e/ou é funcionalmente ativa como uma proteína extracelular solúvel.

[0168] Em algumas modalidades, o polipeptídeo é um polipeptídeo intracelular. Em algumas modalidades, o polipeptídeo é um cuja função é normalmente efetuada (por exemplo, funcionalmente ativo) dentro de uma célula. Um “polipeptídeo intracelular” conforme usado no presente documento se refere a uma proteína que não é secretada pela célula, incluindo polipeptídeos citosólicos solúveis. Em algumas modalidades, o polipeptídeo é um polipeptídeo de tipo selvagem.

[0169] Em algumas modalidades, o polipeptídeo é uma proteína de fígado ou variante do mesmo. Conforme usado no presente documento, uma “proteína de

fígado” é uma proteína que é, por exemplo, produzida de modo endógeno no fígado e/ou funcionalmente ativa no fígado. Em algumas modalidades, a proteína de fígado é uma proteína de circulação produzida pelo fígado ou uma variante da mesma. Em algumas modalidades, a proteína de fígado é uma proteína que é funcionalmente ativa no fígado ou uma variante da mesma. Em algumas modalidades, a proteína de fígado exibe uma expressão elevada no fígado em comparação um ou mais outros tipos de tecido. Em algumas modalidades, o polipeptídeo é uma proteína não hepática.

[0170] Em algumas modalidades, o construto de ácido nucleico bidirecional é linear. Por exemplo, o primeiro e o segundo segmentos são unidos de uma maneira linear através de uma sequência de ligante. Em algumas modalidades, a extremidade 5’ do segundo segmento que compreende uma sequência de complemento reverso é ligada à extremidade 3’ do primeiro segmento. Em algumas modalidades, a extremidade 5’ do primeiro segmento é ligada à extremidade 3’ do segundo segmento que compreende uma sequência de complemento reverso. Em algumas modalidades, a sequência de ligante tem cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 500, 1000, 1500, 2000 ou mais nucleotídeos em comprimento. Conforme seria reconhecido por aqueles versados na técnica, outros elementos estruturais adicionais a uma sequência de ligante ou no lugar da mesma, podem ser inseridos entre o primeiro e o segundo segmentos.

[0171] Os construtos divulgados no presente documento podem ser modificados para incluir qualquer recurso estrutural adequado conforme necessário para qualquer uso particular e/ou que confere uma ou mais funções desejadas. Em algumas modalidades, o construto de ácido nucleico bidirecional divulgado no presente documento não compreende um braço de homologia. Em algumas modalidades, os construtos, por exemplo, construtos de ácido nucleico bidirecional, têm capacidade para inserção em um lócus genômico por união de extremidade não

homóloga (NHEJ). Em algumas modalidades, os construtos divulgados no presente documento são construtos doadores independentes de homologia. Em algumas modalidades, devido, em parte, à função bidirecional do construto de ácido nucleico, o construto bidirecional pode ser inserido em um lócus genômico em qualquer direção (orientação), conforme descrito no presente documento, para permitir inserção e/ou expressão eficiente de um polipeptídeo de interesse.

[0172] Em algumas modalidades, a composição descrita no presente documento compreende um ou mais sítios de entrada ribossômica internos (IRES). Primeiro identificado como um recurso de RNA de Picornavírus, IRES tem uma função importante em iniciar a síntese de proteína na ausência da estrutura de cap 5'. Um IRES pode atuar como o único sítio de ligação de ribossomo, ou pode servir como um dentre múltiplos sítios de ligação de ribossomo de polinucleotídeos. Os construtos que contêm mais do que um sítio de ligação de ribossomo funcional podem codificar diversos peptídeos ou polipeptídeos que são traduzidos independentemente pelos ribossomos (“moléculas de ácido nucleico multicistrônicas"). Alternativamente, os construtos podem compreender um IRES a fim de expressar uma proteína heteróloga que não é fundida a um polipeptídeo endógeno (isto é, um peptídeo de sinal de albumina). Exemplos de sequências de IRES que podem ser utilizadas incluem, sem limitação, aquelas de picornavírus (por exemplo, FMDV), vírus de pragas (CFFV), poliovírus (PV), vírus de encefalomiocardite (ECMV), vírus de febre aftosa (FMDV), vírus de hepatite C (HCV), vírus de febre suína clássica (CSFV), vírus de leucemia murina (MLV), vírus da imunodeficiência símia (SIV) ou vírus de paralisia de grilo (CrPV).

[0173] Em algumas modalidades, o construto de ácido nucleico compreende uma sequência que codifica um peptídeo de autoclivagem, como uma sequência 2A ou uma sequência tipo 2A. O peptídeo de autoclivagem pode ser um peptídeo P2A, um peptídeo T2A ou semelhantes. Em algumas modalidades, o peptídeo de

autoclivagem está localizado a montante do polipeptídeo de interesse. Em uma modalidade, a sequência que codifica o peptídeo 2A pode ser usado para separar a região de codificação de dois ou mais polipeptídeos de interesse. Em outra modalidade, essa sequência pode ser usada para separar a sequência de codificação do construto e a sequência de codificação do lócus endógeno (isto é, sequência de sinal endógena de albumina). Como um exemplo não limitante, a sequência que codifica o peptídeo 2A pode estar entre a região A e a região B (A-2A-B). A presença do peptídeo 2A resultaria na clivagem de uma proteína longa em proteína A, proteína B e o peptídeo 2A. A proteína A e a proteína B podem ser os polipeptídeos iguais ou diferentes de interesse.

[0174] Em algumas modalidades, um ou ambos dentre o primeiro e o segundo segmentos compreendem uma sequência de cauda de poliadenilação e/ou uma sequência de sinal de poliadenilação a jusante de um quadro de leitura aberto. Em algumas modalidades, a sequência de cauda de poliadenilação é codificada, por exemplo, como uma faixa “poli-A”, na extremidade 3’ do primeiro e/ou do segundo segmento.

III. MÉTODOS DE ENTREGA

[0175] Os RNAs guia divulgados no presente documento podem ser entregues a uma célula hospedeira ou indivíduo, in vivo ou ex vivo, com o uso de vários métodos conhecidos e adequados disponíveis na técnica. Os RNAs guia podem ser entregues juntos (individualmente ou combinados) com um agente de ligação de DNA guiado por RNA, como Cas ou ácido nucleico que codifica uma Cas9 (por exemplo, Cas9 ou um ácido nucleico que codifica uma Cas9) e um construto que compreende uma sequência que codifica um gene heterólogo a ser inserido em um sítio de corte criado por um RNA guia da presente divulgação, conforme descrito no presente documento.

[0176] Métodos de entrega de gene com base viral e não viral convencionais

podem ser usados para introduzir o RNA guia divulgado no presente documento, assim como o agente de ligação de DNA guiado por RNA e construto doador em células (por exemplo, células de mamíferos) e tecidos-alvo. Conforme fornecido adicionalmente no presente documento, ácidos nucleicos de sistemas de entrega de vetor não viral, tais como vetores não virais, vetores de plasmídeo e, por exemplo, ácido nucleico nu, e ácido nucleico complexado com um veículo de entrega, tal como um lipossoma, nanopartícula lipídica (LNP) ou poloxâmero. Sistemas de entrega de vetor viral incluem vírus de DNA e RNA.

[0177] Métodos e composições para entrega não viral de ácidos nucleicos incluem eletroporação, lipofecção, microinjeção, biolística, virossomas, lipossomas, imunolipossomas, LNPs, policação ou conjugados de ácido nucleico de lipídio, ácido nucleico nu (por exemplo, DNA/RNA nu), vírions artificial vírions e absorção de DNA intensificada por agente. Sonoporação com o uso, por exemplo, do sistema Sonitron 2000 (Rich-Mar) também pode ser usada para a entrega de ácidos nucleicos.

[0178] Sistemas de entrega de ácido nucleico exemplificativos adicionais incluem aqueles fornecidos por AmaxaBiosystems (Colônia, Alemanha), Maxcyte, Inc. (Rockville, Md.), BTX Molecular Delivery Systems (Holliston, Ma.) e Copernicus Therapeutics Inc., (consultar, por exemplo, o documento de Patente nº U.S. 6.008.336). Lipofecção é descrita, por exemplo, nos documentos de Patente nº U.S. 5.049.386; 4.946.787; e 4.897.355) e reagentes de lipofecção são vendidos comercialmente (por exemplo, Transfectam™ e Lipofectin™). A preparação de complexos de ácido nucleico de lipídio, que inclui lipossomas alvejados, tais como complexos imunolipídicos, é bem conhecida na técnica, e conforme descrito no presente documento.

[0179] Vários sistemas de entrega (por exemplo, vetores, lipossomas, LNPs) que contêm os RNAs guia, agente de ligação de DNA guiado por RNA e construto doador, individualmente ou em combinação, também podem ser administrados a um

organismo para entrega em células *in vivo* ou administrados a uma célula ou cultura de célula *ex vivo*. Administração ocorre através de qualquer uma das rotas normalmente usadas para introduzir uma molécula em contato final com o sangue, fluido ou células que incluem, porém, sem limitação, injeção, infusão, aplicação tópica e eletroporação. Métodos para administrar tais ácidos nucleicos adequados estão disponíveis e são bem conhecidos por aqueles versados na técnica.

[0180] Em algumas modalidades, as composições de RNA guia descritas no presente documento, sozinhas ou codificadas em um ou mais vetores, são formuladas em ou administradas por meio de uma nanopartícula lipídica; consultar, por exemplo, o documento nº PCT/US2017/024973, cujo conteúdo é incorporado ao presente documento, a título de referência, em sua totalidade. Qualquer formulação de nanopartícula lipídica (LNP) conhecida por aqueles versados na técnica para ter capacidade de entregar nucleotídeos aos indivíduos pode ser utilizada com os RNAs guia descritos no presente documento, assim como mRNA que codifica um agente de ligação de DNA guiado por RNA, como Cas ou Cas9, ou um agente de ligação de DNA guiado por RNA, como proteína Cas ou Cas9 por si própria.

[0181] Em algumas modalidades, os RNAs guia divulgados no presente documento podem ser entregues a uma célula hospedeira (in vitro ou in vivo) entregue por meio de uma LNP. Em algumas modalidades, o gRNA/LNP também é associado a um agente de ligação de DNA guiado por RNA, como Cas9 ou um mRNA que codifica um agente de ligação de DNA guiado por RNA, como Cas9. Em algumas modalidades, o gRNA/LNP também é associado a um construto doador, conforme descrito no presente documento.

[0182] Em algumas modalidades, a presente divulgação inclui um método para entregar os gRNAs divulgados no presente documento para uma célula in vitro, em que o gRNA é entregue por meio de uma LNP. Em algumas modalidades, o gRNA é entregue por um meio não LNP, como por meio de um sistema de AAV, e um agente

de ligação de DNA guiado por RNA (por exemplo, Cas9) ou um mRNA que codifica um agente de ligação de DNA guiado por RNA (por exemplo, Cas9), e/ou um construto doador é entregue por uma LNP.

[0183] Em algumas modalidades, a presente divulgação fornece uma composição que compreende qualquer um dos gRNAs divulgados no presente documento e uma LNP. Em algumas modalidades, a composição compreende adicionalmente uma Cas9 ou um mRNA que codifica Cas9, ou outro agente de ligação de DNA guiado por RNA descrito no presente documento. Em algumas modalidades, a composição compreende adicionalmente um construto doador conforme descrito no presente documento.

[0184] Em algumas modalidades, as LNPs compreendem lipídios ionizáveis biodegradáveis. Em algumas modalidades, as LNPs compreendem octadeca-9,12-dienoato de (9Z,12Z)-3-((4,4-bis(octiloxi)butanoil)oxi)-2-((((3-(dietilamino)propoxi)carbonil)oxi)metil)propila, também denominado 3-((4,4-bis(octiloxi)butanoil)oxi)-2-((((3-(dietilamino)propoxi)carbonil)oxi)metil)propil(9Z,12Z)-octadeca-9,12-dienoato) ou outro lipídio ionizável. Consultar, por exemplo, lipídios dos documentos nº PCT/US2018/053559 (depositado em 28 de setembro de 2018), WO/2017/173054, WO2015/095340, e WO2014/136086, assim como referências fornecidas nos mesmos. Em algumas modalidades, o termo catiônico e ionizável no contexto de lipídios LNP é intercambiável, por exemplo, em que lipídios ionizáveis são catiônicos dependendo do pH.

[0185] Em algumas modalidades, qualquer um dentre os RNAs guia descritos no presente documento, agentes de ligação de DNA guiados por RNA e/ou construtos doadores (por exemplo, construtos bidirecionais) divulgados no presente documento, sozinho ou em combinação, seja nu ou como parte de um vetor, é formulado em uma nanopartícula lipídica ou administrado por meio da mesma; consultar, por exemplo, o documento nº WO/2017/173054 cujo conteúdo está incorporado aqui a título de

referência em sua totalidade.

[0186] A eletroporação também é um meio bem conhecido para entrega de cargo, e qualquer metodologia de eletroporação pode ser usada para entrega de qualquer um dos gRNAs divulgados no presente documento. Em algumas modalidades, a eletroporação pode ser usada para entregar qualquer um dos gRNAs divulgados no presente documento, opcionalmente com um agente de ligação de DNA guiado por RNA, como Cas9 ou um mRNA que codifica um agente de ligação de DNA guiado por RNA, como Cas9 entregue pelos mesmos ou por diferentes meios. Em algumas modalidades, a eletroporação pode ser usada para entregar qualquer um dos gRNAs divulgados no presente documento e um construto doador conforme divulgado no presente documento.

[0187] Em determinadas modalidades, a presente divulgação fornece vetores de DNA ou RNA que codificam qualquer um dos RNAs guia que compreendem quaisquer uma ou mais das sequências guia descritas no presente documento. Em determinadas modalidades, a invenção compreende vetores de DNA ou RNA que codificam quaisquer uma ou mais das sequências guia descritas no presente documento. Em algumas modalidades, além de sequências de RNA guia, os vetores compreendem adicionalmente ácidos nucleicos que não codificam RNAs guia. Os ácidos nucleicos que não codificam RNA guia incluem, mas sem limitação, promotores, intensificadores, sequências reguladoras, ácidos nucleicos que codificam um agente de ligação de DNA guiado por RNA, que pode ser uma nuclease, tal como Cas9, e um construto doador que compreende um gene heterólogo. Em algumas modalidades, o vetor compreende uma ou mais sequências de nucleotídeos que codificam um crRNA, um trRNA ou um crRNA e trRNA, conforme divulgado no presente documento.

[0188] Em algumas modalidades, o vetor compreende uma ou mais sequências de nucleotídeos que codificam um sgRNA e um mRNA que codifica um

agente de ligação de DNA guiado por RNA, que pode ser uma proteína Cas, tal como Cas9 ou Cpf1. Em algumas modalidades, o vetor compreende uma ou mais sequências de nucleotídeos que codificam um crRNA, um trRNA e um mRNA que codifica um agente de ligação de DNA guiado por RNA, que pode ser uma proteína Cas, tal como Cas9 ou Cpf1. Em uma modalidade, a Cas9 é de *Streptococcus pyogenes* (isto é, Spy Cas9). Em algumas modalidades, a sequência de nucleotídeos que codifica o crRNA, trRNA ou crRNA e trRNA (que pode ser um sgRNA) compreende ou consiste em uma sequência guia flanqueada por toda ou uma porção de uma sequência de repetição de um sistema de CRISPR/Cas de ocorrência natural. O ácido nucleico que compreende ou consiste no crRNA, trRNA ou crRNA e trRNA pode compreender adicionalmente uma sequência de vetor, em que a sequência de vetor compreende ou consiste em ácidos nucleicos que não são encontrados na natureza juntamente com o crRNA, trRNA ou crRNA e trRNA.

[0189] Em algumas modalidades, o crRNA e o trRNA são codificados por ácidos nucleicos não contíguos dentro de um vetor. Em outras modalidades, o crRNA e o trRNA podem ser codificados por um ácido nucleico contíguo. Em algumas modalidades, o crRNA e o trRNA são codificados por fitas opostas de um ácido nucleico único. Em outras modalidades, o crRNA e o trRNA são codificados pela mesma fita de um ácido nucleico único.

[0190] Em algumas modalidades, o vetor pode ser circular. Em outras modalidades, o vetor pode ser linear. Em algumas modalidades, o vetor pode ser entregue por meio de uma nanopartícula lipídica, lipossoma, nanopartícula não lipídica ou capsídeo viral. Vetores exemplificativos sem limitação incluem plasmídeos, fagomídeos, cosmídeos, cromossomos artificiais, minicromossomos, transpósons, vetores virais e vetores de expressão.

[0191] Em algumas modalidades, o vetor pode ser um vetor viral. Em algumas modalidades, o vetor viral pode ser geneticamente modificado a partir de sua

contrapartida de tipo selvagem. Por exemplo, o vetor viral pode compreender uma inserção, deleção ou substituição de um ou mais nucleotídeos para facilitar a clonagem ou de modo que uma ou mais propriedades do vetor sejam alteradas. Tais propriedades podem incluir capacidade de acondicionamento, eficiência de transdução, imunogenicidade, integração de genoma, replicação, transcrição e tradução. Em algumas modalidades, uma porção do genoma viral pode ser deletada de modo que o vírus tenha capacidade para acondicionar sequências exógenas que têm um tamanho maior. Em algumas modalidades, o vetor viral pode ter uma eficiência de transdução intensificada. Em algumas modalidades, a resposta imunológica induzida pelo vírus em um hospedeiro pode ser reduzida. Em algumas modalidades, os genes virais (tais como, por exemplo, integrase) que promovem a integração da sequência viral em um genoma hospedeiro podem sofrer mutação de modo que os vírus se tornem não integrados. Em algumas modalidades, o vetor viral pode ser deficiente de replicação. Em algumas modalidades, o vetor viral pode compreender sequências de controle transcricional ou translacional exógenas para acionar a expressão de sequências codificantes no vetor. Em algumas modalidades, o vírus pode ser dependente de auxiliar. Por exemplo, o vírus pode precisar de um ou mais vírus auxiliares para fornecer componentes virais (tais como, por exemplo, proteínas virais) necessários para amplificar e acondicionar os vetores em partículas virais. Em um tal caso, um ou mais componentes auxiliares, que incluem um ou mais vetores que codificam os componentes virais, podem ser introduzidos em uma célula hospedeira juntamente com o sistema de vetor descrito no presente documento. Em outras modalidades, o vírus pode ser isento de auxiliar. Por exemplo, o vírus pode ter capacidade para amplificar e acondicionar os vetores sem um vírus auxiliar. Em algumas modalidades, o sistema de vetor descrito no presente documento também pode codificar os componentes virais necessários para a amplificação e acondicionamento de vírus.

[0192] Vetores virais exemplificativos sem limitação incluem vetor de adenovírus associado (AAV), vetores de lentivírus, vetores de adenovírus, vetores adenovirais dependentes auxiliares (HDAd), vetores de vírus herpes simplex (HSV-1), bacteriófago T4, vetores de baculovírus e vetores de retrovírus. Em algumas modalidades, o vetor viral pode ser um vetor de AAV. Em outras modalidades, o vetor viral pode se um vetor de lentivírus.

[0193] Em algumas modalidades, “AAV” se refere a todos os serotipos, subtipos e AAV de ocorrência natural, assim como AAV recombinante. “AAV” pode ser usado para se referir ao próprio vírus ou um derivado do mesmo. O termo “AAV” inclui AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV 12, rh10, e híbridos dos mesmos, AAV aviário, AAV bovino, AAV canino, AAV equino, AAV primata, AAV não primata e AAV ovino. As sequências genômicas de vários serotipos de AAV, assim como as sequências das repetições terminais nativas (TRs), proteínas Rep e subunidades de capsídeo são conhecidas na técnica. Tais sequências podem ser encontradas na literatura ou em bancos de dados públicos, tais como GenBank. Um “vetor de AAV”, conforme usado no presente documento, se refere a um vetor de AAV que compreende uma sequência heteróloga que não é de origem AAV (isto é, uma sequência de ácidos nucleicos heteróloga ao AAV) que compreende, tipicamente, uma sequência que codifica um polipeptídeo heterólogo de interesse. O construto pode compreender um AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6,2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV12, rh10, e híbridos dos mesmos, AAV aviário, AAV bovino, AAV canino, AAV equino, AAV primata, AAV não primata e sequência de capsídeos de AAV ovino. Em geral, a sequência de ácidos nucleicos heteróloga (o transgene) é flanqueada por pelo menos uma e geralmente por duas, sequências de repetição terminal (ITRs)

invertidas de AAV. Um vetor de AAV pode ser de fita simples (ssAAV) ou autocomplementar (scAAV).

[0194] Em algumas modalidades, o lentivírus pode ser não integrado. Em algumas modalidades, o vetor viral pode ser um vetor de adenovírus. Em algumas modalidades, o adenovírus pode ser um adenovírus com alta capacidade para clonagem ou “sem coragem”, em que todas as regiões virais codificantes separadas das repetições de terminal invertidas (ITRs) de 5’ e 3’ e o sinal de acondicionamento (T) são deletadas do vírus para aumentar sua capacidade de acondicionamento. Em ainda outras modalidades, o vetor viral pode ser um vetor de HSV-1. Em algumas modalidades, o vetor à base de HSV-1 é dependente de auxiliar, e em outras modalidades é independente de auxiliar. Por exemplo, um vetor de amplicon que retém apenas a sequência de acondicionamento exige um vírus auxiliar com componentes estruturais para acondicionamento, enquanto um vetor de HSV-1 de 30 kb deletado que remove funções virais não essenciais não exige o vírus auxiliar. Em modalidades adicionais, o vetor viral pode ser bacteriófago T4. Em algumas modalidades, o bacteriófago T4 pode ter capacidade para acondicionar quaisquer moléculas de DNA ou RNA lineares ou circulares quando a cabeça do vírus é esvaziada. Em modalidades adicionais, o vetor viral pode ser um vetor de baculovírus. Em modalidades adicionais, o vetor viral pode ser um vetor de retrovírus. Em modalidades que usam vetores de AAV ou lentivirais, que têm capacidade de clonagem menor, pode ser necessário usar mais de um vetor para entregar todos os componentes de um sistema de vetor, conforme divulgado no presente documento. Por exemplo, um vetor de AAV pode conter sequências que codificam um agente de ligação de DNA guiado por RNA, tal como uma proteína Cas (por exemplo, Cas9), enquanto um segundo vetor de AAV pode conter uma ou mais sequências guia.

[0195] Em algumas modalidades, o vetor pode ter capacidade para acionar expressão de um ou mais sequências de codificação em uma célula. Em algumas

modalidades, a célula pode ser uma célula eucariótica, tal como, por exemplo, uma levedura, planta, inseto ou célula de mamífero. Em algumas modalidades, a célula eucariótica pode ser uma célula de mamífero. Em algumas modalidades, a célula eucariótica pode ser uma célula de roedor. Em algumas modalidades, a célula eucariótica pode ser uma célula de ser humano. Promotores adequados para acionar a expressão em tipos diferentes de células são conhecidos na técnica. Em algumas modalidades, o promotor pode ser de tipo selvagem. Em outras modalidades, o promotor pode ser modificado para expressão mais eficiente ou eficaz. Em ainda outras modalidades, o promotor pode ser truncado, porém, pode reter sua função. Por exemplo, o promotor pode ter um tamanho normal ou um tamanho reduzido que é adequado para acondicionamento apropriado do vetor em um vírus.

[0196] Em algumas modalidades, o vetor pode compreender uma sequência de nucleotídeos que codifica um agente de ligação de DNA guiado por RNA, tal como uma proteína Cas (por exemplo, Cas9) descrita no presente documento. Em algumas modalidades, a nuclease codificada pelo vetor pode ser uma proteína Cas. Em algumas modalidades, o sistema de vetor pode compreender uma cópia da sequência de nucleotídeos que codifica a nuclease. Em outras modalidades, o sistema de vetor pode compreender mais de uma cópia da sequência de nucleotídeos que codifica a nuclease. Em algumas modalidades, a sequência de nucleotídeos que codifica a nuclease pode ser operacionalmente ligada a pelo menos uma sequência de controle transcricional ou translacional. Em algumas modalidades, a sequência de nucleotídeos que codifica a nuclease pode ser operacionalmente ligada a pelo menos um promotor.

[0197] Em algumas modalidades, o vetor pode compreender qualquer um ou mais dentre os construtos que compreendem um gene heterólogo descrito no presente documento. Em algumas modalidades, o gene heterólogo pode ser operacionalmente ligado a pelo menos uma sequência de controle transcricional ou

translacional. Em algumas modalidades, o gene heterólogo pode ser operacionalmente ligado a pelo menos um promotor. Em algumas modalidades, o gene heterólogo não é ligado a um promotor que aciona a expressão do gene heterólogo.

[0198] Em algumas modalidades, o promotor pode ser constitutivo, induzível ou de tecido específico. Em algumas modalidades, o promotor pode ser um promotor constitutivo. Promotores constitutivos exemplificativos sem limitação incluem promotor iniciais imediatos de citomegalovírus (CMV), promotor de vírus símio (SV40), promotor tardio principal de adenovírus (MLP), promotor de vírus Rous sarcoma (RSV), promotor de vírus tumoral mamário de camundongo (MMTV), promotor de fosfoglicerato quinase (PGK), promotor de fator alfa de alongamento (EF1a), promotores de ubiquitina, promotores de actina, promotores de tubulina, promotores de imunoglobulina, um fragmento funcional dos mesmos, ou uma combinação de qualquer um dos supracitados. Em algumas modalidades, o promotor pode ser um promotor de CMV. Em algumas modalidades, o promotor pode ser um promotor de CMV truncado. Em algumas modalidades, o promotor pode ser um promotor de EF1a. Em algumas modalidades, o promotor pode ser um promotor induzível. Promotores induzíveis exemplificativos sem limitação incluem aqueles induzíveis por choque térmico, luz, produtos químicos, peptídeos, metais, esteroides, antibióticos ou álcool. Em algumas modalidades, o promotor induzível pode ser um que tem um baixo nível de expressão basal (não induzida), tal como, por exemplo, o promotor Tet-On® (Clontech).

[0199] Em algumas modalidades, o promotor pode ser um promotor de tecido específico, por exemplo, um promotor específico para expressão no fígado.

[0200] O vetor pode compreender adicionalmente uma sequência de nucleotídeo que codifica o RNA guia descrito no presente documento. Em algumas modalidades, o vetor compreende uma cópia do RNA guia. Em outras modalidades,

o vetor compreende mais que uma cópia do RNA guia. Em modalidades com mais que um RNA guia, os RNAs guia podem ser não idênticos de modo que alvejem sequências-alvo diferentes, ou podem ser idênticos pelo fato de que alvejam a mesma sequência-alvo. Em algumas modalidades, quando os vetores compreendem mais que um RNA guia, cada RNA guia pode ter outras propriedades diferentes, tal como atividade ou estabilidade dentro de um complexo com uma nuclease de DNA guiada por RNA, tal como um complexo de RNP de Cas. Em algumas modalidades, a sequência de nucleotídeos que codifica o RNA guia pode ser operacionalmente ligada a pelo menos uma sequência de controle transcricional ou translacional, tal como um promotor, uma UTR de 3' ou uma UTR de 5'. Em uma modalidade, o promotor pode ser um promotor de tRNA, por exemplo, $tRNA^{Lys3}$, ou uma quimera de tRNA. Consultar Mefferd et al., RNA. 2015 21:1683-9; Scherer et al., *Nucleic Acids Res.* 2007 35: 2620-2628. Em algumas modalidades, o promotor pode ser reconhecido por RNA polimerase III (Pol III). Exemplos sem limitação de promotores de Pol III incluem os promotores U6 e H1. Em algumas modalidades, a sequência de nucleotídeos que codifica o RNA guia pode ser operacionalmente ligada a um promotor U6 de camundongo ou ser humano. Em outras modalidades, a sequência de nucleotídeos que codifica o RNA guia pode ser operacionalmente ligada a um promotor H1 de camundongo ou ser humano. Em modalidades com mais que um RNA guia, os promotores usados para acionar a expressão podem ser iguais ou diferentes. Em algumas modalidades, o nucleotídeo que codifica o crRNA do RNA guia e o nucleotídeo que codifica o trRNA do RNA guia podem ser fornecidos no mesmo vetor. Em algumas modalidades, o nucleotídeo que codifica o crRNA e o nucleotídeo que codifica o trRNA podem ser acionados pelo mesmo promotor. Em algumas modalidades, o crRNA e trRNA podem ser transcritos em um único transcrito. Por exemplo, o crRNA e trRNA podem ser processados a partir do único transcrito para formar um RNA guia de molécula dupla. Alternativamente, o crRNA e trRNA podem

ser transcritos em um RNA guia de molécula única (sgRNA). Em outras modalidades, o crRNA e o trRNA podem ser acionados por seus promotores correspondentes no mesmo vetor. Em ainda outras modalidades, o crRNA e o trRNA podem ser codificados por diferentes vetores.

[0201] Em algumas modalidades, a sequência de nucleotídeos que codifica o RNA guia pode estar localizada no mesmo vetor que compreende a sequência de nucleotídeos que codifica um agente de ligação de DNA guiado por RNA, tal como uma proteína Cas. Em algumas modalidades, a expressão do RNA guia e do agente de ligação de DNA guiado por RNA, tal como uma proteína Cas pode ser acionada por seus próprios promotores correspondentes. Em algumas modalidades, a expressão do RNA guia pode ser acionada pelo mesmo promotor que aciona a expressão do agente de ligação de DNA guiado por RNA, tal como uma proteína Cas. Em algumas modalidades, o RNA guia e o agente de ligação de DNA guiado por RNA, tal como um transcrito de proteína Cas, podem estar contidos em um transcrito único. Por exemplo, o RNA guia pode estar dentro de uma região não traduzida (UTR) do agente de ligação de DNA guiado por RNA, tal como um transcrito de proteína Cas. Em algumas modalidades, o RNA guia pode estar dentro da UTR de 5’ do transcrito. Em outras modalidades, o RNA guia pode estar dentro da UTR de 3’ do transcrito. Em algumas modalidades, a meia-vida intracelular do transcrito pode ser reduzida contendo-se o RNA guia dentro de sua UTR de 3’ e, desse modo, se encurta o comprimento de sua UTR de 3’. Em modalidades adicionais, o RNA guia pode estar dentro de um íntron do transcrito. Em algumas modalidades, sítios de splice adequados podem ser adicionados no íntron dentro do qual o RNA guia está localizado, de modo que o RNA guia sofra, apropriadamente, splice fora do transcrito.

[0202] Em algumas modalidades, as composições compreendem um sistema de vetor. Em algumas modalidades, o sistema de vetor pode compreender um vetor único. Em outras modalidades, o sistema de vetor pode compreender dois vetores.

Em modalidades adicionais, o sistema de vetor pode compreender três vetores. Quando RNAs guia diferentes são usados para multiplexação, ou quando múltiplas cópias do RNA guia são usadas, o sistema de vetor pode compreender mais de três vetores. Em algumas modalidades, o sistema de vetor pode compreender adicionalmente um construto doador conforme descrito no presente documento. Em algumas modalidades, o sistema de vetor pode compreender adicionalmente ácidos nucleicos que codificam uma nuclease. Em algumas modalidades, o sistema de vetor pode compreender adicionalmente ácidos nucleicos que codificam RNAs guia e/ou ácido nucleico que codifica um agente de ligação de DNA guiado por RNA, que pode ser uma proteína Cas, como Cas9. Em algumas modalidades, um ácido nucleico que codifica um RNA guia e/ou um ácido nucleico que codifica um agente de ligação de DNA guiado por RNA ou nuclease estão cada um ou ambos em um vetor separado de um vetor que compreende os construtos doadores divulgados no presente documento. Em qualquer uma das modalidades, o sistema de vetor pode incluir outras sequências que incluem, porém, sem limitação, promotores, intensificadores, sequências reguladoras, conforme descrito no presente documento. Em algumas modalidades, um promotor no sistema de vetor não aciona a expressão de um transgene do construto doador (por exemplo, construto bidirecional). Em algumas modalidades, o sistema de vetor compreende uma ou mais sequências de nucleotídeos que codificam um crRNA, um trRNA ou um crRNA e trRNA. Em algumas modalidades, o sistema de vetor compreende uma ou mais sequências de nucleotídeos que codificam um sgRNA e um mRNA que codifica um agente de ligação de DNA guiada por RNA, que pode ser uma Cas nuclease (por exemplo, Cas9). Em algumas modalidades, o sistema de vetor compreende uma ou mais sequências de nucleotídeos que codificam um crRNA, um trRNA e um mRNA que codifica um agente de ligação de DNA guiada por RNA, que pode ser uma Cas nuclease, tal como, Cas9. Em algumas modalidades, a Cas9 é de *Streptococcus pyogenes* (isto é, Spy Cas9).

Em algumas modalidades, a sequência de nucleotídeos que codifica o crRNA, trRNA ou crRNA e trRNA (que pode ser um sgRNA) compreende ou consiste em uma sequência guia flanqueada por toda ou uma porção de uma sequência de repetição de um sistema de CRISPR/Cas de ocorrência natural. O sistema de vetor pode compreender um ácido nucleico que compreende ou consiste no crRNA, trRNA ou crRNA e trRNA, em que o sistema de vetor compreende ou consiste em ácidos nucleicos que não são encontrados naturalmente juntos com o crRNA, trRNA ou crRNA e trRNA.

[0203] Em algumas modalidades, o sistema de vetor pode compreender promotores induzíveis para começar a expressão apenas após o mesmo ser entregue a uma célula-alvo. Promotores induzíveis exemplificativos sem limitação incluem aqueles induzíveis por choque térmico, luz, produtos químicos, peptídeos, metais, esteroides, antibióticos ou álcool. Em algumas modalidades, o promotor induzível pode ser um que tem um baixo nível de expressão basal (não induzida), tal como, por exemplo, o promotor Tet-On® (Clontech).

[0204] Em modalidades adicionais, o sistema de vetor pode compreender promotores de tecido específico para começar a expressão apenas após o mesmo ser entregue em um tecido específico.

[0205] O vetor ou sistema de vetor pode ser entregue por lipossomo, uma nanopartícula, um exossomo ou uma microvesícula. O vetor também pode ser entregue por uma nanopartícula lipídica (LNP). Um ou mais dentre RNA guia, agente de ligação de DNA de ligação de RNA (por exemplo mRNA), ou construto doador que compreende uma sequência que codifica uma proteína heteróloga, individualmente ou em qualquer combinação, podem ser entregues por lipossoma, uma nanopartícula, um exossoma ou uma microvesícula. Um ou mais dentre RNA guia, agente de ligação de DNA de ligação de RNA (por exemplo mRNA), ou construto doador que compreende uma sequência que codifica uma proteína heteróloga, individualmente

ou em qualquer combinação, podem ser entregues por LNP. Qualquer uma das LNPs e formulações de LNP descritas no presente documento são adequadas para a entrega dos guias, Cas nuclease (ou um mRNA que codifica uma Cas nuclease), combinações dos mesmos, e/ou um construto que compreende um gene heterólogo. Em algumas modalidades, uma composição de LNP é abrangida, a qual compreende: um componente de RNA e um componente de lipídio, em que o componente de lipídio compreende um lipídio de amina, como lipídio ionizável biodegradável; e em que o componente de RNA compreende um RNA guia e/ou um mRNA que codifica uma Cas nuclease. Em alguns casos, o componente de lipídio compreende lipídio ionizável biodegradável, colesterol, DSPC e PEG-DMG.

[0206] Será evidente que um RNA guia divulgado no presente documento, um agente de ligação de DNA guiado por RNA (por exemplo, Cas nuclease ou um ácido nucleico que codifica uma Cas nuclease), e um construto doador pode ser entregue com o uso de sistemas iguais ou diferentes. Por exemplo, o RNA guia, Cas nuclease e construto podem ser carregados pelo mesmo vetor (por exemplo, AAV). Alternativamente, a Cas nuclease (como uma proteína ou mRNA) e/ou gRNA podem ser carregados por um plasmídeo ou LNP, enquanto o construto pode ser carregado por um vetor. Adicionalmente, os sistemas de entrega diferentes podem ser administrados por rotas iguais ou diferentes.

[0207] Em algumas modalidades, o método compreende administrar um RNA guia e um agente de ligação de DNA guiado por RNA (tal como um mRNA que codifica uma Cas9 nuclease) em uma LNP. Em modalidades adicionais, o método compreende administrar um construto de ácido nucleico de AAV que codifica uma proteína transgênica, tal como um construto bidirecional. LNP de CRISPR/Cas9 que compreende RNA guia e um mRNA que codifica uma Cas9, pode ser administrada por via intravenosa. Construto doador de AAV de pode ser administrado por via intravenosa.

[0208] Os sistemas de entrega diferentes podem ser entregues *in vitro* ou *in vivo* simultaneamente ou em qualquer ordem sequencial. Em algumas modalidades, o construto doador, RNA guia e Cas nuclease podem ser entregues *in vitro* ou *in vivo* simultaneamente, por exemplo, em um vetor, dois vetores, vetores individuais, uma LNP, duas LNPs, LNPs individuais ou uma combinação dos mesmos. Em algumas modalidades, o construto doador pode ser entregue *in vivo* ou *in vitro*, como um vetor e/ou associado a uma LNP, antes de (por exemplo, cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ou mais dias) entregar o RNA guia e/ou Cas nuclease, como um vetor e/ou associado a uma LNP individualmente ou em conjunto como uma ribonucleoproteína (RNP). Em algumas modalidades, o construto doador pode ser entregue em múltiplas administrações, por exemplo, todo dia, a cada dois dias, a cada três dias, a cada quatro dias, toda semana, a cada duas semanas, a cada três semanas ou a cada quatro semanas. Em algumas modalidades, o construto doador pode ser entregue em intervalos de uma semana, por exemplo, na semana 1, semana 2 e semana 3, etc. Como um exemplo adicional, o RNA guia e a Cas nuclease, como um vetor e/ou associados a uma LNP individualmente ou em conjunto como uma ribonucleoproteína (RNP), podem ser entregues *in vivo* ou *in vitro*, antes de (por exemplo, cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, ou mais dias) entregar o construto como um vetor e/ou associado a uma LNP. Em algumas modalidades, o RNA guia de albumina pode ser entregue em múltiplas administrações, por exemplo, todo dia, a cada dois dias, a cada três dias, a cada quatro dias, toda semana, a cada duas semanas, a cada três semanas ou a cada quatro semanas. Em algumas modalidades, o RNA guia de albumina pode ser entregue em intervalos de uma semana, por exemplo, na semana 1, semana 2 e semana 3, etc. Em algumas modalidades, a Cas nuclease pode ser entregue em múltiplas administrações, por exemplo, pode ser entregue todo dia, a cada dois dias, a cada três dias, a cada quatro dias, toda semana, a cada duas semanas, a cada três semanas ou a cada quatro

semanas. Em algumas modalidades, a Cas nuclease pode ser distribuída em intervalos de uma semana, por exemplo, na semana 1, semana 2 e semana 3, etc.

IV. MÉTODOS DE USO

[0209] Os gRNAs e métodos e composições associados divulgados no presente documento são úteis para inserir de modo eficaz um gene heterólogo (exógeno) no íntron 1 de um lócus de albumina humana de uma célula hospedeira. Em algumas modalidades, a presente divulgação fornece um método para inserir um gene heterólogo com o íntron 1 de um lócus de albumina humana de uma célula hospedeira, que compreende administrar a uma célula hospedeira (in vivo ou in vitro) um RNA guia conforme descrito no presente documento (qualquer um da SEQ ID NO: 2-33), um agente de ligação de DNA guiado por RNA (por exemplo, Cas nuclease conforme descrito no presente documento), e um construto doador que compreende uma sequência que codifica um polipeptídeo heterólogo de interesse.

[0210] Os gRNAs e métodos e composições associados divulgados no presente documento são úteis para expressar um gene heterólogo (exógeno) no íntron 1 de um lócus de albumina humana de uma célula hospedeira. Em algumas modalidades, a presente divulgação fornece um método para expressar um gene heterólogo com o íntron 1 de um lócus de albumina humana de uma célula hospedeira, que compreende administrar a uma célula hospedeira (in vivo ou in vitro) um RNA guia conforme descrito no presente documento (qualquer um da SEQ ID NO: 2-33), um agente de ligação de DNA guiado por RNA (por exemplo, Cas nuclease conforme descrito no presente documento), e um construto doador que compreende uma sequência que codifica um polipeptídeo heterólogo de interesse.

[0211] Os gRNAs e métodos e composições associados divulgados no presente documento são úteis para tratar um distúrbio associado ao fígado em um indivíduo, conforme descrito no presente documento. Em algumas modalidades, a presente divulgação fornece um método para tratar um distúrbio associado ao fígado,

que compreende administrar a uma célula hospedeira (in vivo ou in vitro) um RNA guia conforme descrito no presente documento (qualquer um da SEQ ID NO: 2-33), um agente de ligação de DNA guiado por RNA (por exemplo, Cas nuclease conforme descrito no presente documento), e um construto doador que compreende uma sequência que codifica um polipeptídeo de interesse.

[0212] As composições e métodos da presente divulgação são úteis e aplicáveis para uma faixa de células hospedeiras. Em algumas modalidades, a célula hospedeira é uma célula de fígado, célula neuronal ou célula muscular. Em algumas modalidades, a célula hospedeira é qualquer célula sem divisão adequada. Conforme usado no presente documento, uma “célula sem divisão” se refere a células que são terminalmente diferenciadas e não se dividem, assim como células quiescentes que não se dividem, porém, permanecem a habilidade em reentrar na divisão e proliferação celular. Células hepáticas, por exemplo, retêm a habilidade em dividir (por exemplo, quando lesionadas ou ressecadas), porém, tipicamente não se dividem. Durante a divisão celular mitótica, a recombinação homóloga é um mecanismo pelo qual o genoma é protegido e quebras de fita dupla são reparadas. Em algumas modalidades, uma célula “sem divisão” se refere a uma célula na qual a recombinação homóloga (HR) não é o mecanismo primário pelo qual quebras de DNA de fita dupla são reparadas na célula, por exemplo, em comparação com uma célula de divisão de controle. Em algumas modalidades, uma célula “sem divisão” se refere a uma célula na qual a união de extremidade não homóloga (NHEJ) é o mecanismo primário pelo qual quebras de DNA de fita dupla são reparadas na célula, por exemplo, em comparação com uma célula de divisão de controle. Tipos de célula sem divisão foram descritos na literatura, por exemplo, através de mecanismos de reparo de quebra de DNA de fita dupla de NHEJ ativos. Consultar, por exemplo, Iyama, DNA Repair (Amst.) 2013, 12(8): 620-636. Em algumas modalidades, a célula hospedeira inclui, porém, sem limitação, uma célula hepática, uma célula muscular ou uma célula neuronal. Em

algumas modalidades, a célula hospedeira é um hepatócito, tal como um hepatócito de camundongo, cino ou ser humano. Em algumas modalidades, a célula hospedeira é um miócito, tal como um monócito de camundongo, cino ou ser humano. Em algumas modalidades, é fornecida no presente documento uma célula hospedeira, descrita acima, que compreende o construto bidirecional divulgado no presente documento. Em algumas modalidades, a célula hospedeira expressa o polipeptídeo de transgene codificado pelo construto bidirecional divulgado no presente documento. Em algumas modalidades, é fornecida no presente documento uma célula hospedeira produzida por um método divulgado no presente documento. Em determinadas modalidades, a célula hospedeira é produzida administrando-se ou entregando-se a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema de ZFN, TALEN ou CRISPR/Cas9.

[0213] Em algumas modalidades, o método compreende adicionalmente alcançar um efeito durável, por exemplo, efeito de pelo menos 1 mês, 2 meses, 6 meses, 1 ano ou 2 anos. Em algumas modalidades, o método compreende adicionalmente alcançar o efeito terapêutico de uma maneira durável e prolongada, por exemplo, efeito de pelo menos 1 mês, 2 meses, 6 meses, 1 ano ou 2 anos. Em algumas modalidades, o nível de atividade e/ou nível de Fator IX de circulação é estável por pelo menos 1 mês, 2 meses, 6 meses, 1 ano ou mais. Em algumas modalidades, uma atividade e/ou nível de estado estável de proteína de FIX é alcançado por pelo menos 7 dias, pelo menos 14 dias, ou pelo menos 28 dias. Em modalidades adicionais, o método compreende manter atividade e/ou níveis de Fator IX após uma dose única por pelo menos 1, 2, 4 ou 6 meses, ou pelo menos 1, 2, 3, 4 ou 5 anos.

[0214] Em modalidades adicionais que envolvem a inserção no lócus de albumina, os níveis de albumina de circulação do indivíduo são normais. O método

pode compreender manter os níveis de albumina de circulação do indivíduo dentro de ±5%, ±10%, ±15%, ±20% ou ±50% dos níveis de albumina de circulação normais. Em determinadas modalidades, os níveis de albumina do indivíduo estão inalterados em comparação com os níveis de albumina de indivíduos não tratados pelo menos pela semana 4, semana 8, semana 12 ou semana 20. Em determinadas modalidades, os níveis de albumina do indivíduo caem transitoriamente e, então, retornam aos níveis normais. Em particular, os métodos podem compreender a não detecção de nenhuma alteração significativa em níveis de albumina no plasma.

[0215] Em algumas modalidades, a invenção compreende um método ou uso de modificação de (por exemplo, criação de uma quebra de fita dupla em) um gene de albumina, tal como um gene de albumina humana, que compreende administrar ou entregar a uma célula hospedeira ou população de células hospedeiras qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. Em algumas modalidades, a invenção compreende um método ou uso de modificação de (por exemplo, criação de uma quebra de fita dupla em) uma região de íntron 1 de albumina, tal como um íntron 1 de albumina humana, que compreende administrar ou entregar a uma célula hospedeira ou população de células hospedeiras qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. Em algumas modalidades, a invenção compreende um método ou uso de modificação de (por exemplo, criação de uma quebra de fita dupla em) um lócus genômico humano, como um sítio de porto seguro, tal como tecido hepático ou célula hospedeira de hepatócito, que compreende administrar ou entregar a uma célula hospedeira ou população de células hospedeiras qualquer um ou mais dentre os gRNAs, construto

doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. A inserção em um lócus genômico, como um sítio de porto seguro, como um sítio de porto seguro de lócus de albumina (por exemplo, íntron 1), permite a superexpressão do gene de Fator IX sem efeitos prejudiciais significativos na célula hospedeira ou população de células, como hepatócitos ou células de fígado. Em algumas modalidades, a invenção compreende um método ou uso de modificação de (por exemplo, criação de uma quebra de fita dupla em) de íntron 1 de um lócus de albumina humana que compreende administrar ou entregar a uma célula hospedeira ou população de células hospedeiras qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. Em algumas modalidades, o RNA guia compreende uma sequência guia que contém pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos que se ligam dentro do íntron 1 de um lócus de albumina humana (SEQ ID NO: 1). Em algumas modalidades, o RNA guia compreende pelo menos 15, 16, 17, 18, 19, ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, o RNA guia compreende uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma

sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é selecionada do grupo que consiste nas SEQ ID NOs: 34-97. Em algumas modalidades, o método é realizado *in vitro*. Em algumas modalidades, o método é realizado *in vivo*. Em algumas modalidades, o construto doador é um construto bidirecional que compreende uma sequência que codifica o Fator IX. Em algumas modalidades, a célula hospedeira é uma célula hepática, dessa forma. Em modalidades adicionais, a célula hepática é um hepatócito.

[0216] Em algumas modalidades, a invenção compreende um método ou uso de introdução de um ácido nucleico de Fator IX em uma célula hospedeira ou população de células hospedeiras que compreende administrar ou entregar qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. Em algumas modalidades, o RNA guia compreende uma sequência guia que contém pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos que têm capacidade para ligar a uma região dentro do íntron 1 de lócus de albumina humana (SEQ ID NO: 1). Em algumas modalidades, o RNA guia compreende pelo menos 15, 16, 17, 18, 19, ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, o RNA guia compreende uma

sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13,19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é selecionada do grupo que consiste nas SEQ ID NOs: 34-97. Em algumas modalidades, o método é *in vitro*. Em algumas modalidades, o método é *in vivo*. Em algumas modalidades, o construto doador é um construto bidirecional que compreende uma sequência que codifica o Fator IX. Em algumas modalidades, a célula hospedeira é uma célula hepática, ou a população de células hospedeiras são células hepáticas, tal como hepatócito.

[0217] Em algumas modalidades, a invenção compreende um método ou uso de expressão de Fator IX em uma célula hospedeira ou uma população de células hospedeiras que compreende administrar ou entregar qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma

sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento. Em algumas modalidades, o RNA guia compreende uma sequência guia que contém pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos que têm capacidade para ligar a uma região dentro do íntron 1 de lócus de albumina humana (SEQ ID NO: 1). Em algumas modalidades, o RNA guia compreende pelo menos 15, 16, 17, 18, 19, ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, o RNA guia compreende uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13,19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é selecionada do grupo que consiste nas SEQ ID NOs: 34-97. Em algumas modalidades, o método é *in vitro*. Em

algumas modalidades, o método é *in vivo*. Em algumas modalidades, o construto doador é um construto bidirecional que compreende uma sequência que codifica o Fator IX. Em algumas modalidades, a célula hospedeira é uma célula hepática, ou a população de células hospedeiras são células hepáticas, tal como hepatócito.

[0218] Em algumas modalidades, a invenção compreende um método ou uso de tratamento de hemofilia (por exemplo, hemofilia A ou hemofilia B) que compreende administrar ou entregar qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende uma sequência que codifica o Fator IX), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease) descritos no presente documento em um indivíduo que necessita dos mesmos. Em algumas modalidades, o RNA guia compreende uma sequência guia que contém pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos que têm capacidade para ligar a uma região dentro do íntron 1 de lócus de albumina humana (SEQ ID NO: 1). Em algumas modalidades, o RNA guia compreende pelo menos 15, 16, 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, o RNA guia compreende uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13,19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32, 33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97. Em algumas modalidades, os RNAs guia divulgados no presente

documento compreendem uma sequência guia que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que tem pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2-33. Em algumas modalidades, os RNAs guia divulgados no presente documento compreendem uma sequência guia que é selecionada do grupo que consiste nas SEQ ID NOs: 34-97. Em algumas modalidades, o construto doador é um construto bidirecional que compreende uma sequência que codifica um polipeptídeo heterólogo. Em algumas modalidades, a célula hospedeira é uma célula hepática, ou a população de células hospedeiras são células hepáticas, tais como hepatócitos.

[0219] Conforme usado no presente documento, “hemofilia” se refere a um distúrbio causado por um gene ou polipeptídeo de Fator IX ausente ou defeituoso. Hemofilia também se refere a um distúrbio causado por um gene ou polipeptídeo de Fator VIII ausente ou defeituoso. O distúrbio inclui afecções que são herdadas e/ou adquiridas (por exemplo, causadas por uma mutação espontânea no gene), e inclui hemofilia A e hemofilia B. A hemofilia A é causada por deficiência de Fator VIII. A hemofilia B é causada por deficiência de Fator IX. Em algumas modalidades, o gene de Fator IX defeituoso ou polipeptídeo resulta em nível de Fator IX reduzido no plasma e/ou uma atividade de coagulação reduzida de Fator IX. Conforme usado no presente documento, hemofilia inclui hemofilia leve, moderada e grave. Por exemplo, indivíduos com menos que cerca de 1% de fator ativo são classificados com hemofilia grave, aqueles com cerca de 1-5% de fator ativo têm hemofilia moderada, e aqueles com hemofilia leve têm entre cerca de 5 e 40% de níveis normais de fator de coagulação ativo.

[0220] Em algumas modalidades, o construto doador compreende uma sequência que codifica o Fator IX, em que a sequência de Fator IX é o Fator IX de tipo

selvagem. Em algumas modalidades, a sequência codifica uma variante de Fator IX. Por exemplo, a variante pode ter atividade de coagulação aumentada em relação ao Fator IX de tipo selvagem. Por exemplo, o Fator IX variante pode compreender uma ou mutações, tal como uma substituição de aminoácido na posição R338 (por exemplo, R338L), em relação ao Fator IX de tipo selvagem. Em algumas modalidades, a sequência codifica uma variante de Fator IX que é 80%, 85%, 90%, 93%, 95%, 97%, 99% idêntica ao Fator IX de tipo selvagem, que tem pelo menos 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% ou mais de atividade em comparação com o Fator IX de tipo selvagem. Em algumas modalidades, a sequência codifica um fragmento de Fator IX, em que o fragmento tem pelo menos 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% ou mais de atividade em comparação com Fator IX de tipo selvagem.

[0221] Em algumas modalidades, o construto doador compreende uma sequência que codifica uma variante de Fator IX, em que a variante de Fator IX ativa coagulação na ausência de seu cofator, Fator VIII (expressão resulta em atividade mimética de FVIII terapeuticamente relevante). Tais variantes de Fator IX podem manter, adicionalmente, a atividade de Fator IX de tipo selvagem. Por exemplo, tal variante de Fator IX pode compreender uma subestação de aminoácido na posição L6, V181, K265, 1383, E185, ou uma combinação dos mesmos em relação ao Fator IX de tipo selvagem. Por exemplo, tal variante de Fator IX pode compreender uma mutação L6F, uma mutação V181I, uma mutação K265A, uma mutação I383V, uma mutação E185D ou uma combinação das mesmas em relação ao Fator IX de tipo selvagem.

[0222] As composições e métodos da presente divulgação são úteis para inserção eficiente de um gene heterólogo de interesse e expressão segura do polipeptídeo heterólogo (por exemplo, um polipeptídeo terapêutico). Em algumas modalidades, o polipeptídeo é um polipeptídeo secretado. Em algumas modalidades, o polipeptídeo é um cuja função é normalmente efetuada (por exemplo,

funcionalmente ativo) como um polipeptídeo secretado. Um “polipeptídeo secretado” conforme usado no presente documento se refere a uma proteína que é secretada pela célula e/ou é funcionalmente ativa como uma proteína extracelular solúvel.

[0223] Em algumas modalidades, o polipeptídeo é um polipeptídeo intracelular. Em algumas modalidades, o polipeptídeo é um cuja função é normalmente efetuada (por exemplo, funcionalmente ativo) dentro de uma célula. Um “polipeptídeo intracelular” conforme usado no presente documento se refere a uma proteína que não é secretada pela célula, incluindo polipeptídeos citosólicos solúveis. Uma ou mais sequências de IRES e/ou peptídeo de autoclivagem podem flanquear um polipeptídeo intracelular, por exemplo, em ou próximo de uma extremidade do polipeptídeo, tal como uma extremidade de terminal amino do polipeptídeo.

[0224] Em algumas modalidades, o polipeptídeo é um polipeptídeo de tipo selvagem. Em algumas modalidades, o polipeptídeo é polipeptídeo (por exemplo, um mutante hiperativo de um polipeptídeo de tipo selvagem) variante (por exemplo, mutante). Em algumas modalidades, o polipeptídeo é uma proteína de fígado. Em algumas modalidades, o polipeptídeo é uma proteína não hepática. Em algumas modalidades, o polipeptídeo é Fator IX, ou uma variante do mesmo. Em algumas modalidades, o polipeptídeo de fígado é, por exemplo, um polipeptídeo para abordar um distúrbio de fígado como, sem limitação, tirosinemia, doença de Wilson, doença de Tay-Sachs, hiperbilirubinema (Crigler-Najjar), porfiria aguda intermitente, citrulinemia tipo 1, colestase intra-hepática familiar progressiva ou doença da urina do xarope de bordo.

[0225] Em algumas modalidades, a expressão do polipeptídeo pela célula hospedeira (in vitro ou in vivo) é aumentada em pelo menos 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, ou mais em relação a um nível expresso pela célula hospedeira antes de fornecer as composições divulgadas no presente documento. Em modalidades

adicionais, a expressão do polipeptídeo heterólogo pode ser aumentada para níveis pelo menos detectáveis ou níveis terapeuticamente eficazes.

[0226] Em algumas modalidades, a expressão do polipeptídeo pela célula hospedeira (in vitro ou in vivo) é aumentada para pelo menos 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, ou mais de um nível normal conhecido (por exemplo, um nível de um polipeptídeo em um indivíduo saudável).

[0227] Em algumas modalidades, a expressão do polipeptídeo pela célula hospedeira (in vitro ou in vivo) é aumentada a pelo menos cerca de 10 μg/ml, 15 μg/ml, 20 μg/ml, 25 μg/ml, 30 μg/ml, 35 μg/ml, 40 μg/ml, 45 μg/ml, 50 μg/ml, 55 μg/ml, 60 μg/ml, 65 μg/ml, 70 μg/ml, 75 μg/ml, 80 μg/ml, 85 μg/ml, 90 μg/ml, 95 μg/ml, 100 μg/ml, 120 μg/ml, 140 μg/ml, 160 μg/ml, 180 μg/ml, 200 μg/ml, 225 μg/ml, 250 μg/ml, 275 μg/ml, 300 μg/ml, 325 μg/ml, 350 μg/ml, 400 μg/ml, 450 μg/ml, 500 μg/ml, 550 μg/ml, 600 μg/ml, 650 μg/ml, 700 μg/ml, 750 μg/ml, 800 μg/ml, 850 μg/ml, 900 μg/ml, 1000 μg/ml, 1100 μg/ml, 1200 μg/ml, 1300 μg/ml, 1400 μg/ml, 1500 μg/ml, 1600 μg/ml, 1700 μg/ml, 1800 μg/ml, 1900 μg/ml, 2000 μg/ml, ou mais, conforme determinado, por exemplo, na célula, plasma, e/ou soro de um indivíduo. Métodos para detectar e medir níveis de polipeptídeo em várias amostras são bem conhecidos na técnica.

[0228] Em algumas modalidades, as composições e métodos da presente divulgação são úteis para tratar uma doença associada ao fígado. Conforme usado no presente documento, uma “doença associada ao fígado” se refere a doenças que causam danos ao tecido de fígado diretamente, doenças que resultam de danos ao tecido de fígado e/ou distúrbios de órgãos não hepáticos ou tecido que resultou de um defeito no fígado. Exemplos de doença associada ao fígado incluem, sem limitação, tirosinemia, doença de Wilson, doença de Tay-Sachs, hiperbilirubinema (Crigler-Najjar), porfiria intermitente aguda, citrulinemia tipo 1, colestase intra-hepática familiar progressiva e doença da urina do xarope de bordo.

[0229] Conforme descrito no presente documento, quaisquer um ou mais dentre o RNA guia divulgados no presente documento, agente de ligação de DNA guiado por RNA e construto doador que compreende um transgene, podem ser entregues com o uso de qualquer sistema de entrega adequado e método conhecido na técnica. As composições podem ser entregues *in vitro* ou *in vivo* simultaneamente ou em qualquer ordem sequencial. Em algumas modalidades, o construto doador, RNA guia e o agente de ligação de DNA guiado por RNA podem ser entregues *in vitro* ou *in vivo* simultaneamente, por exemplo, em um vetor, dois vetores, vetores individuais, uma LNP, duas LNPs, LNPs individuais ou uma combinação dos mesmos. Em algumas modalidades, o construto doador pode ser entregue *in vivo* ou *in vitro*, como um vetor e/ou associado a uma LNP, antes de (por exemplo, cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ou mais dias) entregar o RNA guia e/ou agente de ligação de DNA guiado por RNA, como um vetor e/ou associado a uma LNP individualmente ou em conjunto como uma ribonucleoproteína (RNP). Como exemplo adicional, o RNA guia e o agente de ligação de DNA guiado por RNA, como um vetor e/ou associados a uma LNP individualmente ou em conjunto como uma ribonucleoproteína (RNP), podem ser entregues *in vivo* ou *in vitro*, antes de (por exemplo, cerca de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, ou mais dias) entregar o construto como um vetor e/ou associado a uma LNP. Em algumas modalidades, o RNA guia e o agente de ligação de DNA guiado por RNA estão associados a uma LNP e são entregues à célula hospedeira ou antes da entrega do construto doador.

[0230] Em algumas modalidades, o construto doador compreende uma sequência que codifica o Fator IX ou variantes dos mesmos. Por exemplo, a variante possui atividade aumentada em relação ao polipeptídeo de tipo selvagem. Em algumas modalidades, a sequência codifica uma variante de polipeptídeo que é 80%, 85%, 90%, 93%, 95%, 97%, 99% idêntica a uma sequência de polipeptídeo de tipo selvagem, que tem pelo menos 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100%

ou mais de atividade em comparação com o polipeptídeo de tipo selvagem. Em algumas modalidades, a sequência codifica um fragmento de um polipeptídeo de tipo selvagem, em que o fragmento possui pelo menos 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% ou mais de atividade em comparação com o polipeptídeo de tipo selvagem.

[0231] Em algumas modalidades, uma administração única de um construto doador que compreende um gene heterólogo, RNA guia e agente de ligação de DNA guiado por RNA é suficiente para aumentar a expressão de um polipeptídeo de interesse a um nível desejável. Em outras modalidades, mais do que uma administração de uma composição que compreende um construto doador que compreende um gene heterólogo, RNA guia e agente de ligação de DNA guiado por RNA podem ser benéficas para maximizar efeitos terapêuticos.

[0232] Em algumas modalidades, os RNAs guia, agente de ligação de DNA guiado por RNA e construto doador são administrados individualmente ou em qualquer combinação de modo intravenoso. Em algumas modalidades, os RNAs guia, agente de ligação de DNA guiado por RNA e construto doador são administrados individualmente ou em qualquer combinação à circulação hepática.

[0233] Em algumas modalidades, o hospedeiro ou indivíduo é um mamífero. Em algumas modalidades, o hospedeiro indivíduo é um ser humano. Em algumas modalidades, o hospedeiro ou indivíduo é um roedor (por exemplo, camundongo).

[0234] Esta descrição e modalidades exemplificativas não devem ser tomadas como limitantes. Para os fins deste relatório descritivo e reivindicações anexas, a menos que seja indicado de outra forma, todos os números que expressam quantidades, porcentagens ou proporções, e outros valores numéricos usados no relatório descritivo e nas reivindicações, devem ser entendidos como modificados em todos os casos pelo termo “cerca de”, na medida em que os mesmos não estão ainda modificados assim. Consequentemente, a menos que seja indicado o contrário, os

parâmetros numéricos apresentados no seguinte relatório descritivo e reivindicações anexas são aproximações que podem variar dependendo das propriedades desejadas a serem obtidas. Pelo menos, e sem tentar limitar a aplicação da doutrina de equivalentes ao escopo da reivindicação, cada parâmetro numérico deve ao menos ser interpretado à luz do número de dígitos significativos apresentados e aplicando-se as técnicas comuns de arredondamento.

EXEMPLOS

[0235] Os seguintes exemplos são fornecidos para ilustrar determinadas modalidades divulgadas e não devem ser interpretadas como limitantes do escopo desta divulgação de maneira alguma.

EXEMPLO 1- MATERIAIS E MÉTODOS

CLONAGEM E PREPARAÇÃO DE PLASMÍDEO

[0236] Um construto de inserção bidirecional flanqueado por suas ITRs de AAV2 foi sintetizado e clonado em pUC57-Kan por um fornecedor comercial. O construto resultante (P00 147) foi usado como o vetor de clonagem parental para outros vetores. Os outros construtos de inserção (sem ITRs) também foram comercialmente sintetizados e clonados em pUC57. O plasmídeo purificado foi digerido com a enzima de restrição BglII (New England BioLabs, nº de cat R0144S), e os construtos de inserção foram clonados no vetor parental. Plasmídeo foi propagado em E. coli Quimicamente Competente Stbl3™ (Thermo Fisher, nº de cat C737303).

PRODUÇÃO DE AAV

[0237] A tripla transfecção em células HEK293 foi usada para acondicionar genomas com construtos de interesse para produção de AAV8 e AAV-DJ e vetores resultantes foram purificados tanto a partir de células lisadas quanto de meio de cultura através do método de ultracentrifugação de gradiente de iodixanol (Consultar, por exemplo, Lock et al., Hum Gene Ther. Outubro de 2010;21(10): 1259-71). Os plasmídeos usados na transfecção tripla que conteve o genoma com construtos de

interesse são citados nos Exemplos por um número “PXXXX”, consultar também, por exemplo, a Tabela 9. AAV isolado foi dialisado em tampão de armazenamento (PBS com 0,001% de Pluronic F68). O titulador de AAV foi determinado por qPCR com o uso de iniciadores/sonda localizados dentro da região de ITR.

TRANSCRIÇÃO *IN VITRO* (“IVT”) DE MRNA NUCLEASE

[0238] mRNA de Cas9 de *Streptococcus pyogenes* (“Spy”) com cap e poliadenilado que contém pseudo-U de N1-metila foi gerado por transcrição *in vitro* com o uso de um modelo de DNA de plasmídeo linearizado e RNA polimerase de T7. De modo geral, DNA de plasmídeo que contém um promotor T7 e uma região de 100 nt poli (A/T) foi linearizado incubando-se a 37 °C com Xbal para completar a digestão seguida por inativação térmica de Xbal a 65 °C. O plasmídeo linearizado foi purificado a partir de sais de enzima e tampão. 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 pseudo-UTP de N1-metila (Trilink); 10 mM de ARCA (Trilink); 5 U/µl de RNA polimerase de T7 (NEB); 1 U/µl de inibidor de Rnase Murino (NEB); 0,004 U/µl de pirofosfatase de *E. coli* inorgânica (NEB); e 1x tampão de reação. 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 modelo de DNA. O mRNA de Cas9 foi purificado com o uso de um kit de Limpeza de Transcrição MegaClear de acordo com o protocolo do fabricante (ThermoFisher). Alternativamente, o mRNA de Cas9 foi purificado com o uso de precipitação de LiCl, precipitação de acetato de amônia e precipitação de acetato de sódio ou com o uso de um método de precipitação de LiCl seguido por purificação adicional através de filtragem de fluxo tangencial. A concentração de transcrito foi determinada medindo-se a absorção de luz em 260 nm (Nanodrop), e a transcrição foi analisada por eletroforese capilar pelo Bioanlayzer (Agilent).

[0239] Os mRNAs de Cas9 abaixo compreendem ORF de Cas9 SEQ ID

NO:703 ou SEQ ID NO: 704 ou uma sequência da Tabela 24 do documento nº PCT/US2019/053423 (que é incorporado aqui a título de referência).

FORMULAÇÕES LIPÍDICAS PARA ENTREGA DE MRNA E GRNA DE CAS9

[0240] mRNA e gRNA de Cas9 foram entregues às células e animais com o uso de formulações lipídicas que compreendem lipídio ionizável octadeca-9,12-dienoato de (9Z,12Z)-3-((4,4-bis(octiloxi)butanoil)oxi)-2-((((3-(dietilamino)propoxi)carbonil)oxi)metil)propila, também denominado 3-((4,4-bis(octiloxi)butanoil)oxi)-2-((((3-(dietilamino)propoxi)carbonil)oxi)metil)propil(9Z,12Z)-octadeca-9,12-dienoato), colesterol, DSPC e PEG2k-DMG.

[0241] Para experimentos que utilizam formulações lipídicas pré-misturadas (denominadas no presente documento “pacotes de lipídio”), os componentes foram reconstituídos em 100% de etanol em uma razão molar de lipídio ionizável:colesterol:DSPC:PEG2k-DMG de 50:38:9:3, antes de serem misturados com cargas de RNA (por exemplo, mRNA e gRNA de Cas9) em uma razão molar de amina de lipídio para fosfato de RNA (N:P) de cerca de 6,0, conforme descrito adicionalmente no presente documento.

[0242] Para experimentos que utilizam os componentes formulados como nanopartículas lipídicas (LNPs), os componentes foram dissolvidos em 100% de etanol em várias razões molares. As cargas de RNA (por exemplo, mRNA e gRNA de Cas9) foram dissolvidas em 25 mM de citrato, 100 mM de NaCl, pH 5,0, que resulta em uma concentração de carga de RNA de aproximadamente 0,45 mg/ml.

[0243] Para os experimentos descritos no Exemplo 2, as LNPs foram formadas através da mistura microfluídica das soluções de lipídio e RNA com o uso de um instrumento de Nanossistemas de Precisão NanoAssemblr™ Benchtop de acordo com o protocolo do fabricante. Uma razão de 2:1 de solvente aquoso para orgânico foi mantida durante a mistura com o uso de taxas de fluxo diferentes. Após misturar, as LNPs foram coletadas, diluídas em água (aproximadamente 1:1 em v/v),

mantidas por 1 hora em temperatura ambiente, e diluídas adicionalmente com água (aproximadamente 1:1 em v/v) antes da troca de tampão final. A troca de tampão final em 50 mM de Tris, 45 mM de NaCl, 5% (em p/v) de sacarose, pH 7,5 (TSS) foi completada com colunas de dessalinização de PD-10 (GE). Caso necessário, as formulações foram concentradas por centrifugação com filtros centrífugos de 100 kDa da Amicon (Millipore). A mistura resultante foi, então, filtrada com o uso de um filtro estéril de 0,2 μm. A LNP final foi armazenada a -80 °C até o uso posterior. As LNPs foram formuladas em uma razão molar de lipídio ionizável:colesterol:DSPC:PEG2k-DMG de 45:44:9:2, com uma razão molar de amina de lipídio para fosfato de RNA (N:P) de cerca de 4,5 e uma razão de gRNA para mRNA de 1:1 em peso.

[0244] Para os experimentos descritos em outros exemplos, as LNPs foram preparadas com o uso de uma técnica de fluxo cruzado que utiliza mistura por jato de colisão do lipídio em etanol com dois volumes de soluções de RNA e um volume de água. O lipídio em etanol foi misturado através de uma mistura cruzada com os dois volumes de solução de RNA. Uma quarta corrente de água foi misturada com a corrente de saída do cruzamento através de uma baliza em linha (Consultar o documento nº W0O2016010840 Figura 2). As LNPs foram mantidas por 1 hora em temperatura ambiente, e diluídas adicionalmente com água (aproximadamente 1:1 em v/v). LNPs diluídas foram concentradas com o uso de filtragem de fluxo tangencial em um cartucho de folha plana (Sartorius, 100 kD de MWCO) e, então, o tampão trocado através de diafiltração em 50 mM de Tris, 45 mM de NaCl, 5% (em p/v) de sacarose, pH 7,5 (TSS). Alternativamente, a troca de tampão final em TSS foi completada com colunas de dessalinização PD-10 (GE). Caso necessário, as formulações foram concentradas por centrifugação com filtros centrífugos de 100 kDa da Amicon (Millipore). A mistura resultante foi, então, filtrada com o uso de um filtro estéril de 0,2 μm. A LNP final foi armazenada a 4 °C ou- 80 °C até o uso posterior. As LNPs foram formuladas em uma razão molar de lipídio ionizável:colesterol:DSPC:PEG2k-DMG de

50:38:9:3, com uma razão molar de amina de lipídio para fosfato de RNA (N:P) de cerca de 6,0 e uma razão de gRNA para mRNA de 1:1 em peso.

CULTURA CELULAR E ENTREGA *IN VITRO* DE MRNA DE CAS9, GRNA E CONSTRUTOS DE INSERÇÃO

CÉLULAS HEPA1-6

[0245] Células Hepa1-6 foram colocadas em placa em densidade de 10000 células/poço em placas de 96 poços. 24 horas depois, células foram tratadas com LNP e AAV. Antes do tratamento o meio foi aspirado dos poços. LNP foi diluída em 4 ng/ul em meio de DMEM+10% de FBS e diluída adicionalmente em 2 ng/ul em 10% de FBS (em DMEM) e incubada a 37 °C por 10 min (em uma concentração final de 5% de FBS). MOI-alvo de AAV foi 1e6, diluído em meio de DMEM+10% de FBS. 50 µl da LNP diluída acima em 2 ng/ul foram adicionados às células (o que entrega um total de 100 ng de carga de RNA) seguido por 50 µl de AAV. O tratamento de LNP e AAV teve minutos de diferença. O volume total de meio em células foi de 100 µl. Após 72 horas depois do tratamento e 30 dias depois do tratamento, o sobrenadante dessas células tratadas foi coletado para análise ELISA de FIX humano conforme descrito abaixo.

HEPATÓCITOS PRIMÁRIOS

[0246] Hepatócitos de camundongo primários (PMH), hepatócitos de cino primários (PCH) e hepatócitos de humano primários (PHH) foram descongelados e ressuspensos em meio de descongelamento de hepatócito com complementos (ThermoFisher) seguido por centrifugação. O sobrenadante foi descartado, e as células peletizadas ressuspensas em meio de plaqueamento de hepatócito mais pacote de complemento (ThermoFisher). Células foram contadas e colocadas em placa em placas de 96 poços revestidas com colágeno I Bio-coat em uma densidade de 33000 células/poço para PHH e 50000 células/poço para PCH e 15000 células/poço para PMH. Permitiu-se que células em placa assentassem e aderissem por 5 horas em uma incubadora de cultura de tecido a 37 °C e atmosfera de 5% de

CO_2. Após incubação, as células foram verificadas para formação de monocamada e foram lavadas três vezes com manutenção de hepatócito antes e incubadas a 37 °C.

[0247] Para experimentos que utilizam entrega de pacote de lipídio, mRNA e gRNA de Cas9 foram, cada um, separadamente diluídos em 2 mg/ml no meio de manutenção e 2,9 µl de cada um foram adicionados aos poços (em uma placa de Eppendorf de 96 poços) que contêm 12,5 µl de 50 mM de citrato de sódio, 200 mM de cloreto de sódio a pH 5 e 6,9 µl de água. 12,5 µl de formulação de pacote de lipídio foram, então, adicionados, seguido por 12,5 µl de água e 150 µl de TSS. Cada poço foi diluído em 20 ng/µl (com relação ao teor de RNA total) com o uso de meio de manutenção de hepatócito e, então, diluído em 10 ng/µl (com relação ao teor de RNA total) com 6% de soro de camundongo fresco. O meio foi aspirado das células antes da transfecção e 40 µl das misturas de pacote de lipídio/RNA foram adicionados às células, seguido pela adição de AAV (diluído em meio de manutenção) em uma MOI de 1e5. O meio foi coletado 72 horas depois do tratamento para análise e células foram coletadas para análise adicional, conforme descrito no presente documento.

ENSAIOS DE LUCIFERASE

[0248] Para experimentos que envolvem detecção de NanoLuc em meio celular, um volume de Substrato de Ensaio de Luciferase Nano-Glo® foi combinado com 50 volumes de Tampão de Ensaio de Luciferase Nano-Glo®. O ensaio foi executado em um executador Promega Glomax em um tempo de integração de 0,5 segundo com o uso de diluição de 1:10 de amostras (50 µl de reagente + 40 µl de água + 10 µl de meio celular).

[0249] Para experimentos que envolvem a detecção da marcação HiBit em meio celular, Proteína LgBiT e Substrato Extracelular de HiBiT Nano-GloR foram diluídos em 1:100 e 1:50, respectivamente, em Tampão Extracelular de HiBiT Nano-GloR em temperatura ambiente. O ensaio foi executado em um executador Promega Glomax em um tempo de integração de 1,0 segundo com o uso de diluição de 1:10

de amostras (50 µl de reagente + 40 µl de água + 10 µl de meio celular).

ENTREGA IN VIVO DE LNP E/OU AAV

[0250] Camundongos foram dosados com AAV, LNP, tanto AAV quanto LNP ou veículo (PBS + 0,001% de Plurônico para veículo de AAV, TSS para veículo de LNP) por meio da veia caudal lateral. AAV foi administrado em um volume de 0,1 ml por animal com quantidades (genomas vetoriais/camundongo, “vg/ms”) conforme descrito no presente documento. LNPs foram diluídas em TSS e administradas em quantidades, conforme indicado no presente documento, em cerca de 5 µl/grama de peso corporal. Tipicamente, camundongos foram injetados primeiro com AAV e, então, com LNP, caso aplicável. Em vários pontos no tempo depois do tratamento, soro e/ou tecido hepático foi coletado para determinadas análises conforme descrito adicionalmente abaixo.

ANÁLISE DE ELISA DE FATOR IX HUMANO (HFIX)

[0251] Para estudos *in vitro*, os níveis de Fator IX humano total secretados em meio celular foram determinados com o uso de um Kit de ELISA de Fator IX Humano (Abeam, nº de cat ab188393) de acordo com o protocolo do fabricante. Níveis de hFIX secretado foram quantificados a partir de uma curva padrão com o uso de ajuste logístico de 4 parâmetros e expressos como ng/ml de meio.

[0252] Para estudos *in vivo*, o sangue foi coletado e o soro ou plasma foi isolado, conforme indicado. Os níveis de Fator IX humano total foram determinados com o uso de um Kit de ELISA de Fator IX Humano (Abeam, nº de cat ab188393) de acordo com o protocolo do fabricante. Níveis de hFIX em soro ou plasma foram quantificados a partir de uma curva padrão com o uso de ajuste logístico de 4 parâmetros e expressos como µg/ml de soro.

SEQUENCIAMENTO DE PRÓXIMA GERAÇÃO (“NGS”) E ANÁLISE PARA EFICIÊNCIA DE CLIVAGEM NO ALVO

[0253] O sequenciamento profundo foi utilizado para identificar a presença de

inserções e deleções introduzidas pela edição de gene, por exemplo, dentro do íntron 1 de albumina. Iniciadores de PCR foram projetados ao redor do sítio-alvo e a área genômica de interesse foi amplificada. Projeto de sequência iniciadora foi realizado como é padrão no campo.

[0254] PCR adicional foi realizada de acordo com os protocolos do fabricante (Illumina) para adicionar química para sequenciamento. Os amplicons foram sequenciados em um instrumento Illumina MiSeq. As leituras foram alinhadas ao genoma de referência após eliminação daqueles que têm baixa pontuação de qualidade. Os arquivos resultantes que contêm as leituras foram mapeados para o genoma de referência (arquivos BAM), quando as leituras que sobrepuseram a região-alvo de interesse foram selecionadas e o número de leituras de tipo selvagem contra o número de leituras que contêm uma inserção ou deleção (“indel”) foi calculado.

[0255] A porcentagem de edição (por exemplo, a “eficiência de edição” ou “percentual de edição”) é definida como o número total de leituras de sequência com inserções ou deleções (“indels”) sobre o número total de leituras de sequência, que incluem tipo selvagem.

ANÁLISE DE HIBRIDIZAÇÃO *IN SITU*

[0256] BaseScope (ACDbio, Newark, CA) é uma tecnologia de hibridização *in situ* de RNA especializada que pode fornecer a detecção específica de junções de éxon, por exemplo, em um transcrito de mRNA híbrida que contém um transgene de inserção (hFIX) e sequência de codificação do sítio de inserção (éxon 1 de albumina). BaseScope foi usado para medir a porcentagem de células hepáticas que expressam o mRNA híbrido.

[0257] Para detectar o mRNA híbrido, duas sondas contra os mRNAs híbridos que podem surgir depois da inserção de um construto bidirecional foram projetadas pela ACDbio (Newark, CA). Uma das sondas foi projetada para detectar um mRNA híbrido que resulta da inserção do construto em uma orientação, enquanto a outra

sonda foi projetada para detectar um mRNA híbrido que resulta da inserção do construto na outra orientação. Fígados de diferentes grupos de camundongos foram coletados e cortados recentemente congelados. O ensaio BaseScope, que usa uma sonda única ou sondas agrupadas foi realizado de acordo com o protocolo de fabricação. Lâminas foram digitalizados e analisados pelo software HALO. O antecedente (grupo tratado salino) desse ensaio foi de 0,58%.

EXEMPLO 2 - TESTE IN VITRO DE MODELOS DE INSERÇÃO PARA ÍNTRON 1 DE ALBUMINA COM E SEM BRAÇOS DE HOMOLOGIA

[0258] Nesse Exemplo, células Hepa1-6 foram cultivadas e tratadas com modelos de inserção que carregam AAV de várias formas (por exemplo, que têm um genoma de fita simples (“ssAAV”) ou um genoma autocomplementar (“scAAV”)), na presença ou ausência de LNP que entrega mRNA de Cas9 e G000551, por exemplo, conforme descrito no Exemplo 1 (n = 3). O AAV e a LNP foram preparados conforme descrito no Exemplo 1. Depois do tratamento, o meio foi coletado para níveis de Fator IX humano, conforme descrito no Exemplo 1.

[0259] Células Hepa1-6 são uma linhagem de célula hepática de camundongo imortalizada que continua a se dividir em cultura. Conforme mostrado na Figura 2 (ponto no tempo de 72 horas depois do tratamento), o vetor (scAAV derivado de plasmídeo P00204) que compreende 200 pb de braços de homologia resultou na expressão detectável de hFIX, por exemplo, após a inserção no íntron 1 de albumina nas células no ciclo. Uso dos vetores de AAV derivados de P00123 (braços de homologia que carecem de scAAV) e P00147 (braços de homologia que carecem de construto bidirecional de ssAAV) não resulta na expressão detectável de hFIX nesse experimento. As células foram mantidas em cultura e esses resultados foram confirmados quando ensaiados novamente em 30 dias depois do tratamento (dados não mostrados).

EXEMPLO 3 - TESTE *IN VIVO* DE MODELOS DE INSERÇÃO PARA ÍNTRON

1 DE ALBUMINA COM E SEM BRAÇOS DE HOMOLOGIA

[0260] Nesse Exemplo, camundongos foram tratados com AAV derivado dos mesmos plasmídeos (P00123, P00204 e P00147) conforme testado *in vitro* no Exemplo 2. Os materiais de dosagem foram preparados e dosados conforme descrito no Exemplo 1. Camundongos C57B1/6 foram dosados (n = 5 para cada grupo) com genomas de vetor 3e11 cada (vg/ms) seguido por LNP que compreende G000551 (“G551”) em uma dose de 4 mg/kg (com relação ao teor de carga de RNA total). Quatro semanas depois da dose, os animais sofreram eutanásia e o tecido hepático e os soros foram coletados para edição e expressão de hFIX, respectivamente.

[0261] Conforme mostrado na Figura 3A e na Tabela 12, os níveis de edição hepática de -60% foram detectados em cada grupo de animais tratados com LNP que compreende gRNA que alveja íntron 1 de albumina murina. No entanto, apesar dos níveis robustos e consistentes de edição em cada grupo de tratamento, os animais que receberam o vetor de ssAAV sem braços de homologia (vetor derivado de P00147) em combinação com tratamento de LNP resultaram no nível mais elevado de expressão de hFIX em soro (Figura 3B e Tabela 13).

TABELA 12: INDEL %

Modelo	Indel Médio (%)	Desvio Padrão de Indel (%)
scAAV Sem Corte (P00123)	66,72	4,09
ssAAV Sem Corte (POO 147)	68,10	2,27
HR de ssAAV (P00204)	70,16	3,68
LNP apenas	68,24	6,47
Veículo	0,28	0,08

TABELA 13: NÍVEIS DE FATOR IX (UG/ML)

Modelo	Fator IX Médio (ug/ml)	Desvio Padrão de Fator IX (ug/ml)
scAAV Sem Corte (P00123)	0,75	0,28
ssAAV Sem Corte (POO 147)	2,92	1,04
HR de ssAAV (P00204)	0,96	0,35
LNP apenas	0	0
Veículo	0	0

EXEMPLO 4 - TESTE *IN VIVO* DE MODELOS DE INSERÇÃO DE SSAAV PARA ÍNTRON 1 DE ALBUMINA COM E SEM BRAÇOS DE HOMOLOGIA

[0262] O experimento descrito nesse Exemplo examinou o efeito de incorporação de braços de homologia em vetores de ssAAV *in vivo*.

[0263] Os materiais de dosagem usados nesse experimento foram preparados e dosados conforme descrito no Exemplo 1. Camundongos C57B1/6 foram dosados (n = 5 para cada grupo) com genomas de vetor 3e11 vg/ms seguido por LNP que compreende G000666 (“G666”) ou G000551 (“G551”) em uma dose de 0,5 mg/kg (com relação ao teor de carga de RNA total). Quatro semanas depois da dose, os soros de animais foram coletados para expressão de hFIX.

[0264] Conforme mostrado na Figura 4A e na Tabela 14, o uso dos vetores de ssAAV com braços de homologia assimétricos (braços de 300/600 pb, braços de 300/2000 pb e braços de 300/1500 pb para vetores derivados de plasmídeos P00350, P00356 e P00362, respectivamente) para inserção no sítio de íntron 1 de albumina alvejado por G551 resultou nos níveis de hFIX de circulação que estiveram abaixo do limite inferior de detecção para o ensaio. No entanto, o uso do vetor de ssAAV (derivado de P00147) sem braços de homologia e que tem dois quadros de leitura abertos (ORF) de hFIX em uma orientação bidirecional resultou em níveis detectáveis de hFIX de circulação em cada animal.

[0265] De maneira similar, o uso dos vetores de ssAAV com braços de

homologia simétricos dos plasmídeos P00353 e P00354, respectivamente, para inserção no sítio de íntron 1 de albumina alvejado por G666 resultou em níveis inferiores, porém, detectáveis, em comparação com o uso do bidirecional sem braços de homologia (derivados de P00147) (consultar a Figura 4B e a Tabela 15).

TABELA 14: HFIX

Níveis Séricos AAV	FIX Sérico Médio (ug/ml)	Desvio Padrão de FIX Sérico (ug/ml)
P00147	5,13	1,31
P00350	-0,22	0,08
P00356	-0,23	0,04
P00362	-0,09	0,16

TABELA 15: NÍVEIS SÉRICOS DE HFIX

AAV	FIX Sérico Médio (ug/ml)	Desvio Padrão de FIX Sérico (ug/ml)
P00147	7,72	4,67
P00353	0,20	0,23
P00354	0,46	0,26

EXEMPLO 5- EXAME *IN VITRO* DE CONSTRUTOS BIDIRECIONAIS ATRAVÉS DE 20 SÍTIOS-ALVO NO ÍNTRON 1 DE ALBUMINA EM HEPATÓCITOS DE CAMUNDONGO PRIMÁRIOS

[0266] Tendo demonstrado que os construtos bidirecionais que não têm braços de homologia superaram vetores com outras configurações para inserção no íntron 1 de albumina, o experimento descrito nesse Exemplo examinou os efeitos de auxiliar os aceitantes de splice. Esses construtos bidirecionais variados foram testados através de um painel de sítios-alvo que utiliza 20 diferentes gRNAs que alvejam íntron 1 de albumina murina em hepatócitos de camundongo primários (PMH).

[0267] Os materiais de entrega de pacote de ssAAV e lipídio testados nesse

Exemplo foram preparados e entregues aos PMH, conforme descrito no Exemplo 1, com o AAV em uma MOI de 1e5. Depois do tratamento, meio de DNA e célula genômico isolado foi coletado para análise de edição e expressão de transgene, respectivamente. Cada um dos vetores compreendeu um repórter que pode ser medido através da detecção de fluorescência à base de luciferase, conforme descrito no Exemplo 1, plotada na Figura 5C como unidades de luciferase relativa (“RLU”). Por exemplo, os vetores de AAV que compreendem os ORFs de hFIX continham um peptídeo de HiBit fundido em suas extremidades 3’, e o vetor de AAV que compreende apenas genes repórter compreendia ORF de aNanoLuc (adicionalmente a GFP). Esquemas de cada um dos vetores testados são fornecidos na Figura 5A. Os gRNAs testados são mostrados na Figura 5B e na Figura 5C, com o uso de um número encurtado para aqueles listados na Tabela 5 (por exemplo, quando os zeros iniciais são omitidos, por exemplo, quando “G551” corresponde a “GOOO551” na Tabela 5).

[0268] Conforme mostrado na Figura 5B e na Tabela 16, níveis consistentes, porém, variados de edição foram detectados para cada um dos grupos de tratamento através de cada combinação testada. A expressão de transgene que usa várias combinações de modelo e RNA guia é mostrada na Figura 5C. Conforme mostrado na Figura 5D, um nível significativo de formação de indel não resulta necessariamente em expressão mais eficiente dos transgenes. Nem todos os guias que geraram indels significativas em altos níveis de proteínas com o mesmo modelo de inserção conforme medido pelas atividades de luciferase relativas. Ao usar modelos derivados de P00411 e P00418, os valores de R^2 foram 0,54 e 0,37 entre indel e atividade de luciferase, respectivamente, quando os guias com menos que 10% de edição não estão incluídos (Figura 5D). De modo interessante, apesar de diferenciarem ORFs e aceitantes de splice, os níveis relativos de expressão conforme medido em RLUs foram consistentes entre os três vetores testados, demonstrando a robustez, reprodutibilidade e modularidade do sistema de construto bidirecional, por exemplo, para uso na inserção

de transgenes de interesse no íntron 1 de albumina (consultar a Figura 5C e a Tabela 17). O aceitante de splice de albumina de camundongo e aceitante de splice de FIX humano resultaram, cada um, na expressão de transgene eficaz.

TABELA 16: % DE INDEL

ID guia	P00411		P00418		P00415	
	Indel Médio (%)	Desvio Padrão de Indel (%)	Indel Médio (%)	Desvio Padrão de Indel (%)	Indel Médio (%)	Desvio Padrão de Indel (%)
G000551	67,4	1,42	70,67	2,29	66,73	4,90
G000552	90,93	0,15	91,10	2,43	90,37	1,01
G000553	77,80	3,83	77,47	1,87	80,50	0,85
G000554	72,37	6,49	70,53	3,16	70,60	2,91
G000555	35,37	2,63	35,77	9,34	40,47	4,75
G000666	62,47	3,87	50,90	19,41	65,90	3,99
G000667	30,57	2,73	25,30	3,67	31,67	2,29
G000668	63,60	2,02	66,65	4,60	68,30	4,90
G000669	19,10	2,51	19,33	1,53	18,70	1,25
G000670	47,80	3,27	49,10	4,42	51,97	2,06
GO11722	4,20	0,72	4,27	1,20	4,20	0,26
GO 11723	5,63	1,27	6,07	0,15	5,93	0,15
GO11724	6,10	1,28	8,50	2,69	7,13	1,27
GO 11725	1,93	0,29	2,60	0,79	2,53	0,65
GO 11726	10,73	1,46	11,70	0,50	12,43	1,33
GO 11727	14,20	1,56	14,80	2,36	16,20	2,69
GO11728	10,55	1,20	13,65	0,92	15,50	1,56
GO 11729	5,00	0,10	5,63	0,25	6,00	1,01
G011730	7,83	0,97	9,13	0,59	7,33	0,59
G011731	23,70	0,66	25,27	1,21	24,87	1,01
AAV Apenas	0,15	0,07	0,05	0,07	0,10	0,00

TABELA 17: EXPRESSÃO DE LUCIFERASE

ID guia	P00411		P00418		P00415	
	Luciferase Média (RLU)	Desvio Padrão de Luciferase (RLU)	Luciferase Média (RLU)	Desvio Padrão de Luciferase (RLU)	Luciferase Média (RLU)	Desvio Padrão de Luciferase (RLU)
G000551	58000,00	4331,28	41800,00	2165,64	78633,33	20274,70
G000552	95700,00	10573,08	80866,67	27911,35	205333,33	30664,86
G000553	205333,33	52993,71	177333,33	32929,22	471666,67	134001,00
G000554	125333,33	55949,38	91933,33	19194,10	232666,67	67002,49
G000555	59933,33	11566,04	77733,33	11061,80	155666,67	15947,83
G000666	88500,00	28735,87	93266,67	30861,19	313000,00	15394,80
G000667	75333,33	22653,11	68966,67	27222,11	153000,00	30805,84
G000668	164000,00	56320,51	133400,00	65111,29	429000,00	120751,80
G000669	28933,33	11636,29	22033,33	2413,16	46466,67	6543,19
G000670	162666,67	32959,57	200000,00	33867,39	424666,67	36473,73
GO 11722	16766,67	3384,28	8583,33	4103,10	24000,00	8915,16
GO 11723	22733,33	7252,82	17133,33	4905,44	26100,00	8109,87
GO 11724	17300,00	2400,00	28033,33	9091,94	30933,33	3365,02
GO 11725	8253,33	1163,20	8890,00	1429,27	20366,67	13955,05
GO 11726	12223,33	3742,54	11610,00	2490,44	14950,00	8176,03
GO 11727	35600,00	8128,35	36300,00	12301,22	86700,00	5023,94
GO 11728	14900,00	5011,99	22466,67	7130,45	38166,67	13829,08
GO 11729	10460,00	2543,95	11223,33	2220,28	26966,67	16085,50
G011730	14833,33	2307,24	21700,00	8681,59	41233,33	25687,03
G011731	16433,33	3274,65	22566,67	2205,30	20756,67	13096,20
AAV Apenas	217,00	15,56	215,00	15,56	207,00	1,41

EXEMPLO 6- EXAME *IN VIVO* DE CONSTRUTOS BIDIRECIONAIS ATRAVÉS DE SÍTIOS-ALVO DE ÍNTRON 1 DE ALBUMINA

[0269] O ssAAV e as LNPs testados nesse Exemplo foram preparados e

entregues em camundongos C57B1/6, conforme descrito no Exemplo 1, para avaliar o desempenho dos construtos bidirecionais através de sítios-alvo *in vivo*. Quatro semanas depois da dose, os animais sofreram eutanásia e o tecido hepático e os soros foram coletados para edição e expressão de hFIX, respectivamente.

[0270] Em um experimento inicial, 10 diferentes formulações de LNP que contêm 10 diferentes gRNAs que alvejam íntron 1 de albumina foram entregues aos camundongos juntamente com ssAAV derivado de P00147. O AAV e a LNP foram entregues em 3e11 vg/ms e 4 mg/kg (com relação ao teor de carga de RNA total), respectivamente (n = 5 para cada grupo). Os gRNAs testados nesse experimento são mostrados na Figura 6. Conforme mostrado na Figura 6 e na Tabela 18 conforme observado *in vitro*, um nível significativo de formação de indel não foi preditivo para inserção ou expressão dos transgenes.

TABELA 18 - NÍVEIS DE SORO DE HFIX E % DE INDEL

Guia	Indel Médio (%)	Desvio Padrão de Indel (%)	Luciferase Média (RLU)	Desvio Padrão de Luciferase (RLU)
G000551	75,02	1,27	3,82	3,38
G000555	51,18	1,19	32,56	9,05
G000553	62,78	2,64	25,07	4,04
G000667	52,96	4,96	32,03	6,74
G000554	55,24	2,28	29,48	7,34
G000552	67,56	1,73	14,79	5,34
G000668	43,14	5,78	26,72	7,97
G000669	50,68	2,97	10,70	4,43
G000666	64,62	1,34	26,19	5,56
G000670	55,90	1,30	30,96	8,44

[0271] Em um experimento separado, um painel de 20 gRNAs que alvejam os 20 diferentes sítios-alvo testados *in vitro* no Exemplo 5 foram testados *in vivo*. Para essa finalidade, formulações de LNP que contêm os 20 gRNAs que alvejam íntron 1 de albumina foram entregues aos camundongos juntamente com ssAAV derivado de P00147. O AAV e a LNP foram entregues em 3e11 vg/ms e 1 mg/kg (com relação ao

teor de carga de RNA total), respectivamente. Os gRNAs testados nesse experimento são mostrados na Figura 7A e na Figura 7B.

TABELA 19 EDIÇÃO NO FÍGADO

Guia	Edição Hepática Média (%)	Desvio Padrão de Edição Hepática (%)
G000551	59,48	4,02
G000555	58,72	3,65
G000553	51,26	2,81
G000554	33,04	8,76
G000555	12,72	4,46
G000666	53,60	4,92
G000667	26,74	4,98
G000668	39,22	3,04
G000669	33,34	4,77
G000670	47,50	5,58
GO 11722	10,34	1,68
GO 11723	4,02	0,84
GO 11724	2,46	0,64
GO 11725	8,26	1,24
GO 11726	6,90	1,01
GO 11727	13,33	6,43
GO 11728	35,78	9,34
GO 11729	4,62	1,46
G011730	12,68	3,14
G011731	26,70	1,86

TABELA 20: NÍVEIS DE HFIX SÉRICO

Guia	Semana 1		Semana 2		Semana 4	
	FIX Médio (ug/ml)	Desvio Padrão de FIX (ug/ml)	FIX Médio (ug/ml)	Desvio Padrão de FIX (ug/ml)	FIX Médio (ug/ml)	Desvio Padrão de FIX (ug/ml)
GOOO551	10,88	2,74	10,25	2,51	9,39	3,48
GOOO555	13,34	2,09	12,00	2,75	12,43	2,57
GOOO553	17,64	4,34	20,27	6,35	15,31	2,43
G000554	12,79	4,99	14,29	6,09	12,74	4,93
GOOO555	11,94	5,79	11,99	5,76	8,61	4,02
G000666	21,63	1,32	20,65	1,55	17,23	0,62
G000667	16,77	2,86	12,35	2,85	12,57	5,60
G000668	21,35	1,51	18,20	3,18	17,72	2,25
G000669	5,76	2,10	6,72	2,93	3,39	0,78
G000670	18,18	2,17	19,16	3,05	15,49	3,61
G011722	8,07	1,74	7,74	2,41	8,07	1,74
G011723	2,11	0,28	1,65	0,28	2,11	0,28
G011724	0,92	0,43	0,60	0,30	0,92	0,43
G011725	1,75	0,77	1,14	0,67	1,75	0,77
G011726	0,59	0,30	1,01	0,64	0,59	0,30
G011727	6,71	2,80	6,90	3,68	6,71	2,80
G011728	11,77	3,12	12,29	3,43	11,77	3,12
G011729	0,94	0,35	0,89	0,29	0,94	0,35
G011730	5,93	1,77	6,33	1,73	5,93	1,77
G011731	3,56	0,87	3,78	0,50	3,56	0,87
AAV Apenas	0,00	0,00	0,00	0,00	0,00	0,00
Veículo	0,00	0,00	0,00	0,00	0,00	0,00
Soro Humano	3,63	0,32	3,61	0,35	3,28	0,03

[0272] Conforme mostrado, na Figura 7A e Tabela 19, os níveis variados de edição foram detectados para cada um dos grupos de tratamento através de cada

combinação de LNP/vetor testada. No entanto, conforme mostrado na Figura 7B e consistente com os dados *in vitro* descritos no Exemplo 5, níveis superiores de edição não resultaram necessariamente em níveis superiores de expressão dos transgenes *in vivo*, o que indica uma carência de correlação entre edição e inserção/expressão dos construtos bidirecionais. De fato, uma correlação muito pequena existe entre a quantidade de edição alcançada e a quantidade de expressão de hFIX, conforme visto na plotagem fornecida na Figura 7D e na Tabela 20. Em particular, um valor de R^2 de apenas 0,34 é calculado entre os conjuntos de dados de edição e expressão para esse experimento, quando esses gRNAs que alcançam menos que 10% de edição são removidos da análise. Interessantemente, conforme mostrado na Figura 7C, é fornecida uma plotagem de correlação que compara os níveis de expressão conforme medidos em RLU a partir do experimento *in vitro* do Exemplo 5 com os níveis de expressão de transgene *in vivo* detectados nesse experimento, sendo que um valor de R^2 de 0,70 demonstra uma correlação positiva entre o primário exame de célula e os tratamentos *in vivo*.

[0273] Para avaliar a inserção do construto bidirecional no nível celular, tecidos hepáticos de animais tratados foram examinados com o uso de um método de hibridização *in situ* (BaseScope), por exemplo, conforme descrito no Exemplo 1. Esse ensaio utilizou sondas que podem detectar as junções entre o transgene de hFIX e a sequência de éxon 1 de albumina de camundongo, como um transcrito híbrido. Conforme mostrado na Figura 8A, células positivas para o transcrito híbrido foram detectadas em animais que receberam tanto AAV quanto LNP. Especificamente, quando AAV apenas é administrado, menos que 1,0% das células foram positivas para o transcrito híbrido. Com a administração de LNPs que compreendem G011723, G000551 ou G000666, 4,9%, 19,8% ou 52,3% das células foram positivas para o transcrito híbrido. Adicionalmente, conforme mostrado na Figura 8B, níveis de hFIX de circulação correlacionados com o número de células que eram positivas para o

transcrito híbrido. Finalmente, o ensaio utilizou sondas agrupadas que podem detectar a inserção do construto de hFIX bidirecional em qualquer orientação. No entanto, quando uma sonda única que detecta apenas uma única orientação foi usada, a quantidade de células que foram positivas para o transcrito híbrido foi de cerca de metade daquelas detectadas com o uso das sondas agrupadas (em um exemplo, 4,46% contra 9,68%), o que sugere que o construto bidirecional, de fato, tem capacidade para inserção no íntron 1 de albumina em qualquer orientação, o que dá origem às transcrições híbridas expressas que se correlacionam com a quantidade de expressão de transgene no nível de proteína. Esses dados mostram que os níveis de proteína de circulação alcançados são dependentes do guia usado para inserção.

EXEMPLO 7- TEMPORIZAÇÃO DE ENTREGA DE AAV E LNP *IN VIVO*

[0274] Nesse Exemplo, a temporização entre a entrega de ssAAV que compreende o construto de hFIX bidirecional e LNP para inserção alvejada no íntron 1 de albumina foi examinada.

[0275] O ssAAV e as LNPs testados nesse Exemplo foram preparados e entregues aos camundongos, conforme descrito no Exemplo 1. A formulação de LNP conteve G000551 e o modelo bidirecional foi entregue como ssAAV derivado de P00147. O AAV e a LNP foram entregues em 3e11 vg/ms e 4 mg/kg (com relação ao teor de carga de RNA total), respectivamente (n = 5 para cada grupo). Um coorte de “Molde apenas” recebeu AAV apenas, e um coorte de “PBS” não recebeu AAV ou LNP. Um coorte recebeu AAV e LNP sequencialmente (separados por minutos) no dia 0 (“Molde + LNP dia 0”); outro coorte recebeu AAV no dia 0 e LNP no dia 1 (“Molde + LNP dia 1”); e um coorte final recebeu AAV no dia 0 e LNP no dia 7 (“Molde + LNP dia 7”). Em 1 semana, 2 semanas e 6 semanas, o plasma foi coletado para análise de expressão de hFIX.

[0276] Conforme mostrado na Figura 9, hFIX foi detectado em cada coorte em cada tempo examinado, exceto para o ponto no tempo de 1 semana para o coorte que

recebeu a dose de LNP no mesmo dia na semana 1 após entrega de AAV.

EXEMPLO 8- DOSAGEM MÚLTIPLA DE LNP DEPOIS DA ENTREGA DE AAV

[0277] Nesse Exemplo, os efeitos de dosagem de repetição de LNP depois da administração de ssAAV para inserção alvejada no íntron 1 de albumina foram examinados.

[0278] O ssAAV e as LNPs testados nesse Exemplo foram preparados e entregues aos camundongos C57B1/6, conforme descrito no Exemplo 1. A formulação de LNP conteve G000551 e o ssAAV foi derivado de P00147. O AAV e a LNP foram entregues em 3e11 vg/ms e 0,5 mg/kg (com relação ao teor de carga de RNA total), respectivamente (n = 5 para cada grupo). Um coorte de “Molde apenas” recebeu AAV apenas, e um coorte de “PBS” não recebeu AAV ou LNP. Um coorte recebeu AAV e LNP sequencialmente (separados por minutos) no dia 0 sem tratamentos adicionais (“Molde + LNP(1x)” na Figura 10); outro coorte recebeu AAV e LNP sequencialmente (separados por minutos) no dia 0 e uma segunda dose no dia 7 (“Molde + LNP(2x)” na Figura 10); e um coorte final recebeu AAV e LNP sequencialmente (separados por minutos) no dia 0, uma segunda dose de LNP no dia 7 e uma terceira dose de LNP no dia 14 (“Molde + LNP(3x)” na Figura 10). Nas semanas 1, 2, 4 e 6 depois da administração de AAV, o plasma foi coletado para análise de expressão de hFIX.

[0279] Conforme mostrado na Figura 10, hFIX foi detectado em cada coorte em cada tempo examinado, e múltiplas doses subsequentes de LNP não aumentaram significativamente a quantidade de expressão de hFIX.

EXEMPLO 9- DURABILIDADE DE EXPRESSÃO DE HFIX *IN VIVO*

[0280] A durabilidade de expressão de hFIX após a inserção alvejada no íntron 1 de albumina ao longo do tempo em animais tratados foi avaliada nesse Exemplo. Para essa finalidade, hFIX foi medido no soro de animais tratados como parte de um estudo de durabilidade de um ano.

[0281] O ssAAV e as LNPs testados nesse Exemplo foram preparados e entregues aos camundongos C57B1/6, conforme descrito no Exemplo 1. A formulação de LNP conteve G000551 e o ssAAV foi derivado de P00147. O AAV foi entregue em 3e11 vg/ms e a LNP foi entregue em 0,25 ou 1,0 mg/kg (com relação ao teor de carga de RNA total) (n = 5 para cada grupo).

[0282] Conforme mostrado na Figura 11A e na Figura 11B e nas Tabelas 21-22, a expressão de hFIX do íntron 1 de albumina foi prolongada em cada ponto no tempo avaliado para ambos os grupos dentro de 12 semanas. Acredita-se que uma queda nos níveis observada nas 8 semanas ocorre devido à variabilidade do ensaio de ELISA. Níveis de albumina sérica foram medidos por ELISA na semana 2 e semana 41, o que mostra que níveis de albumina de circulação são mantidos ao longo do estudo.

TABELA 21: NÍVEIS DE FIX

Semana	Dose			
	0,25 mpk de LNP		1 mpk de LNP	
	hFIX Médio (ug/ml)	Desvio Padrão de hFIX (ug/ml)	hFIX Médio (ug/ml)	Desvio Padrão de hFIX (ug/ml)
2	0,48	0,21	2,24	1,12
4	0,55	0,18	2,82	1,67
8	0,40	0,17	1,72	0,77
12	0,48	0,20	2,85	1,34
20	0,48	0,27	2,45	1,26
41	0,79	0,49	4,63	0,95

TABELA 22: NÍVEIS DE HFIX

Semana	Dose			
	0,25 mpk de LNP		1 mpk de LNP	
	hFIX Médio (ug/ml)	Desvio Padrão de hFIX (ug/ml)	hFIX Médio (ug/ml)	Desvio Padrão de hFIX (ug/ml)
2	0,87	0,15	4,02	1,75
8	0,99	0,15	4,11	1,41
12	0,93	0,14	4,15	1,35
20	0,83	0,22	4,27	1,54
41	0,83	0,37	4,76	1,62
52	0,82	0,25	4,72	1,54

EXEMPLO 10 - EFEITOS DE DOSES VARIADAS DE AAV E LNP PARA MODULAR EXPRESSÃO DE HFIX *IN VIVO*

[0283] Nesse Exemplo, os efeitos de variação da dose tanto de AAV quanto de LNP para modular a expressão de hFIX após a inserção alvejada no íntron 1 de albumina foram avaliados em camundongos C57B1/6.

[0284] O ssAAV e as LNPs testados nesse Exemplo foram preparados e entregues aos camundongos, conforme descrito no Exemplo 1. A formulação de LNP conteve G000553 e o ssAAV foi derivado de P00147. O AAV foi entregue em 1e11, 3e11, 1e12 ou 3e12 vg/ms e a LNP foi entregue em 0,1, 0,3 ou 1,0 mg/kg (com relação ao teor de carga de RNA total) (n = 5 para cada grupo). Duas semanas depois da dose, os animais sofreram eutanásia, e os soros foram coletados para análise de expressão de hFIX.

[0285] Conforme mostrado na Figura 12A (1 semana) e na Figura 12B (2 semanas) e na Tabela 23, a variação da dose de AAV ou LNP pode modular a quantidade de expressão de hFIX do íntron 1 de albumina *in vivo.*

TABELA 23: HFIX SÉRICO

Ponto no tempo	Dose de RNP (mg/kg)	Dose de AAV (MOI)	FIX Médio (ng/ml)	SD	N
Semana 1	0,1	1E+11	0,08	0,02	2
		3E+11	0,11	0,04	5
		1E+12	0,41	0,15	5
		3E+12	0,61	0,17	5
	0,3	1E+11	0,36	0,14	5
		3E+11	0,67	0,26	5
		1E+12	1,76	0,14	5
		3E+12	4,70	2,40	5
	1,0	1E+11	3,71	0,31	4
		3E+11	8,00	0,51	5
		1E+12	14,17	1,38	5
		3E+12	20,70	2,79	5
	Soro humano 1:1,000		6,62	-	1
Semana 2	0,1	1E+11	0,12	0,01	2
		3E+11	0,26	0,07	5
		1E+12	0,83	0,24	5
		3E+12	1,48	0,35	5
	0,3	1E+11	0,70	0,26	4
		3E+11	1,42	0,37	5
		1E+12	3,53	0,49	5
		3E+12	8,94	4,39	5
	1,0	1E+11	5,40	0,47	4
		3E+11	12,31	2,45	5
		1E+12	17,89	1,95	5
		3E+12	25,52	3,62	5
	Soro humano 1:1,000		4,47	-	1

EXEMPLO 11- EXAME *IN VITRO* DE CONSTRUTOS BIDIRECIONAIS

ATRAVÉS DE SÍTIOS-ALVO EM HEPATÓCITOS CINOMOLGOS PRIMÁRIOS E HUMANOS PRIMÁRIOS

[0286] Nesse Exemplo, vetores de ssAAV que compreendem um construto bidirecional foram testados através de um painel de sítios-alvo que utilizam gRNAs que alvejam íntron 1 de albumina cinomolga (“cino”) e humana em hepatócitos cinos primários (PCH) e humanos primários (PHH), respectivamente.

[0287] Os materiais de entrega de pacote de ssAAV e lipídio testados nesse Exemplo foram preparados e entregues em PCH e PHH, conforme descrito no Exemplo 1. Depois do tratamento, meio de DNA e célula genômico isolado foi coletado para análise de edição e expressão de transgene, respectivamente. Cada um dos vetores compreendeu um repórter que pode ser medido através da detecção de fluorescência à base de luciferase, conforme descrito no Exemplo 1 (derivado de plasmídeo P00415), plotado na Figura 13B e na Figura 14B como unidades de luciferase relativa (“RLU”). Os dados de RLU mostrados na Figura 13B e na Figura 14B graficamente são reproduzidos numericamente na Tabela 3 e Tabela 4 abaixo. Por exemplo, os vetores de AAV contiveram o ORF de NanoLuc (além de GFP). Esquemas dos vetores testados são fornecidos na Figura 13B e na Figura 14B. Os gRNAs testados são mostrados em cada uma das Figuras com o uso de um número encurtado para aqueles listados na Tabela 1 e Tabela 3.

[0288] Conforme mostrado na Figura 13A para PCH e Figura 14A para PHH, níveis variados de edição foram detectados para cada uma das combinações testadas (dados de edição para algumas combinações testadas no experimento de PCH não são relatados na Figura 13A e na Tabela 3 devido à falha de determinados pares iniciadores usados para o sequenciamento à base de amplicon). Os dados de edição mostrados na Figura 13A e na Figura 14A graficamente são reproduzidos numericamente na Tabela 3 e Tabela 4 abaixo. No entanto, conforme mostrado na Figura 13B, na Figura 13C e na Figura 14B e na Figura 14C, um nível significativo de

formação de indel não foi preditivo para a inserção ou expressão dos transgenes, o que indica pouca correlação entre edição e inserção/expressão dos construtos bidirecionais em PCH e PHH, respectivamente. Como uma medida, o valor de R^2 calculado na Figura 13C é de 0,13, e o valor de R^2 da Figura 14D é de 0,22.

TABELA 3: DADOS DE EXPRESSÃO DE TRANSGENE E EDIÇÃO DE ÍNTRON 1 DE ALBUMINA PARA SGRNAS ENTREGUES AOS HEPATÓCITOS CINOMOLGOS PRIMÁRIOS

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médias	Desvio Padrão de RLU
G009867	25,05	0,21	10650,67	1455,97
G009866	18,7	3,96	75556,67	12182,98
G009876	14,85	4,88	27463,33	10833,53
G009875	12,85	2,33	51660,00	6362,36
G009874	28,25	6,01	270433,30	133734,10
G009873	42,65	5,59	178600,00	87607,25
G009865	59,15	0,21	301666,70	18610,03
G009872	48,15	3,46	320233,30	63517,43
G009871	46,5	5,23	211966,70	65852,44
G009864	33,2	8,34	210033,30	61201,33
G009863	54,8	12,45	69853,33	15216,92
G009862	44,6	7,21	508666,70	119876,30
G009861	28,65	0,21	178666,70	15821,93
G009860	33,2	7,07	571333,30	52728,87
G009859	0,05	0,07	258333,30	79052,73
G009858	14,65	1,77	402333,30	25579,94
G009857	23	0,99	312333,30	73036,52
G009856	14,8	0,99	95900,00	21128,42
G009851	1,5	0,42	105766,70	27048,91
G009868	12,15	2,47	43033,33	9141,85
G009850	63,45	13,93	228200,00	101542,10
G009849	57,55	8,27	225400,00	46001,30

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médias	Desvio Padrão de RLU
G009848	33	5,37	156333,30	20647,84
G009847	66,75	7	100866,70	22159,72
G009846	61,85	5,02	31766,67	10107,59
G009845	54,4	7,5	43020,00	11582,23
G009844	47,15	2,05	110466,70	32031,44

TABELA 4: DADOS DE EXPRESSÃO DE TRANSGENE E EDIÇÃO DE ÍNTRON 1 DE ALBUMINA PARA SGRNAS ENTREGUES AOS HEPATÓCITOS HUMANOS PRIMÁRIOS

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médias	Desvio Padrão de RLU
G009844	19,07	2,07	268333,30	80432,17
G009851	0,43	0,35	18033,33	2145,54
G009852	47,20	3,96	18400,00	2251,67
G009857	0,10	0,14	71100,00	14609,24
G009858	8,63	9,16	32000,00	18366,55
G009859	3,07	3,50	59500,00	16014,99
G009860	18,80	4,90	190333,30	54307,76
G009861	10,27	2,51	62233,33	9865,26
G009866	13,60	13,55	96200,00	46573,81
G009867	12,97	3,04	3916,67	1682,03
G009868	0,63	0,32	10176,67	2037,80
G009874	49,13	0,60	318000,00	114118,40
GO 12747	3,83	0,23	51000,00	6161,17
GO 12748	1,30	0,35	17433,33	2709,86
GO 12749	9,77	1,50	75066,67	11809,04
G012750	42,73	4,58	5346,67	2977,35
G012751	7,77	1,16	32066,67	18537,62
G012752	32,93	2,27	402000,00	83144,45
G012753	21,20	2,95	71800,00	32055,73
G012754	0,60	0,10	16933,33	4254,80

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médias	Desvio Padrão de RLU
G012755	1,10	0,10	13833,33	3685,56
G012756	2,17	0,40	35600,00	6055,58
G012757	1,07	0,25	13993,33	6745,08
G012758	0,90	0,10	34900,00	15308,82
G012759	2,60	0,35	30566,67	15287,36
G012760	39,10	6,58	6596,67	2133,13
G012761	36,17	2,43	467666,70	210965,20
G012762	8,50	0,57	217000,00	13000,00
G012763	47,07	3,07	142333,30	37581,02
G012764	44,57	5,83	1423333,00	261023,60
G012765	19,90	1,68	179666,70	57011,69
G012766	8,50	0,28	243333,30	17473,79

[0289] Adicionalmente, vetores de ssAAV que compreendem um construto bidirecional foram testados através de um painel de sítios-alvo que utilizam RNAs guia únicos que alvejam íntron 1 de albumina humana em hepatócitos humanos primários (PHH).

[0290] Os materiais de ssAAV e LNP foram preparados e entregues aos PHH, conforme descrito no Exemplo 1. Depois do tratamento, meio de DNA e célula genômico isolado foi coletado para análise de edição e expressão de transgene, respectivamente. Cada um dos vetores compreendeu um repórter que pode ser medido através da detecção de fluorescência à base de luciferase, conforme descrito no Exemplo 1 (derivado de plasmídeo P00415), plotado na Figura 14D e mostrado na Tabela 24 como unidades de luciferase relativa (“RLU”). Por exemplo, os vetores de AAV contiveram o ORF de NanoLuc (além de GFP). Esquemas dos vetores testados são fornecidos na Figura 13B e na Figura 14B. Os gRNAs testados são mostrados na Figura 14D com o uso de um número encurtado para aqueles listados na Tabela 1 e Tabela 7.

TABELA 24: DADOS DE EXPRESSÃO DE TRANSGENE DE ÍNTRON 1 DE ALBUMINA PARA SGRNAS ENTREGUES AOS HEPATÓCITOS CINOMOLGOS PRIMÁRIOS

Guia	Luciferase Média (RLU)	Desvio Padrão de Luciferase (RLU)
G009844	3700000	509117
G009852	281000	69296
G009857	1550000	127279
G009858	551000	108894
G009859	1425000	77782
G009860	2240000	183848
G009861	663500	238295
G009866	274000	11314
G009867	44700	566
G009874	2865000	431335
GO 12747	651000	59397
GO 12749	867000	93338
G012752	4130000	268701
GO 12753	1145000	162635
GO12757	579000	257387
G012760	129000	36770
GO12761	4045000	728320
GO 12762	2220000	127279
G012763	1155000	205061
GO 12764	11900000	1555635
G012765	1935000	134350
G012766	2050000	169706
LNP	8430	212

EXEMPLO 12- TESTE *IN VIVO* DA INSERÇÃO DE GENE DE FATOR 9 HUMANO EM PRIMATAS NÃO HUMANOS

[0291] Nesse exemplo, um estudo de 8 semanas foi realizado para avaliar a

inserção de gene de Fator 9 humano e a expressão de proteína de hFIX em macacos cinomolgos através da administração de vírus adenoassociado (AAV) e/ou nanopartículas lipídicas (LNP) com vários guias. Esse estudo foi conduzido com formulações de LNP e formulações de AAV preparadas conforme descrito acima. Cada formulação de LNP conteve mRNA de Cas9 e RNA guia (gRNA) com uma razão de mRNA:gRNA de 2:1 em peso. O ssAAV foi derivado de P00147.

[0292] Macacos cinomolgos foram tratados em coortes de n = 3. Animais foram dosados com AAV através de injeção ou infusão de bolo lenta nas doses descritas na Tabela 5. Depois do tratamento de AAV, animais receberam tampão ou LNP, conforme descrito na Tabela 5, através de bolo lento ou infusão.

[0293] Duas semanas depois da dose, espécimes hepáticas foram coletadas através de biópsia percutânea guiada por ultrassom única. Cada espécime de biópsia foi instantaneamente congelado em nitrogênio líquido e armazenado em -86 a -60 °C. Análise de edição dos espécimes hepáticos foi realizada através de sequenciamento NGS, conforme anteriormente descrito.

[0294] Para análise de ELISA de Fator IX, amostras sanguíneas foram coletadas a partir dos animais nos dias 7, 14, 28 e 56 depois da dose. Amostras sanguíneas foram coletadas e processadas em plasma depois da retirada de sangue e armazenadas em -86 a -60 °C até a análise.

[0295] Os níveis de Fator IX humano totais foram determinados a partir de amostras de plasma através de ELISA. Brevemente, a microplaca de 96 poços Reacti-Bind (VWR nº de cat PI15041) foi revestida com anticorpo de captura (mAB de camundongo para anticorpo de Fator IX humano (HTI, nº de cat AHIX-5041)) em uma concentração de 1 µg/ml, então, bloqueada com o uso de 1x PBS com 5% de Albumina de Soro Bovino. Amostras de teste ou padrões de proteína de Fator IX humano purificado (ERL, nº de cat HFIX 1009, nº de lote HFIX4840) diluídos em plasma de macaco cinomolgo foram incubados a seguir em poços individuais. O

anticorpo de detecção (anticorpo policlonal de Fator 9 anti-humano de ovelha, Abeam, nº de cat abl28048) foi adsorvido em uma concentração de 100 ng/ml. O anticorpo secundário (pAbs de IgG anti-ovelha de burro com HRP, Abeam, nº de cat ab97125) foi usado em 100 ng/ml. Conjunto de Reagente de Substrato de TMB (BD OptEIA nº de cat 555214) foi usado para desenvolver a placa. Densidade óptica foi avaliada espectrofotometricamente em 450 nm em um leitor de microplaca (sistema i3 de Dispositivos Moleculares) e analisada com o uso de SoftMax pro 6.4.

[0296] A formação de indel foi detectada, o que confirma que a edição ocorreu. Os dados de NGS mostraram formação de indel eficaz. Expressão de hFIX a partir do lócus de albumina em NHPs foi medida por ELISA e é representada na Tabela 6 e Figura 15. Níveis plasmáticos de hFIX alcançaram níveis anteriormente descritos como terapeuticamente eficazes (George, et al., NEJM 377(23), 2215-27, 2017).

[0297] Conforme medido, níveis de proteína de hFIX de circulação foram prolongados através do estudo de oito semanas (consultar a Figura 15 que mostra os níveis médios do dia 7, 14, 28 e 56 de ~135, ~140, ~150 e ~110 ng/ml, respectivamente), que alcança níveis de proteína que estão na faixa de ~75 ng/ml a ~250 ng/ml. Níveis de hFIX plasmáticos foram calculados com o uso de uma atividade específica de ~8 vezes superior para a variante de hFIX hiperfuncional R338L (Simioni et al., NEJM 361(17), 1671-75, 2009) (que relata uma atividade de proteína específica de hFIX-R338L de 390±28 U por miligrama, e uma atividade de proteína específica para fator IX de tipo selvagem de 45±2,4 U por miligrama). Ao calcular a atividade de Fator IX funcionalmente normalizada para a variante de Fator IX hiperfuncional testada nesse exemplo, o experimento alcançou níveis estáveis de proteína de Fator IX humano na NHPs durante o estudo de 8 semanas que correspondem a cerca de 20-40% de atividade de Fator IX de tipo selvagem (faixa abrange 12-67% de atividade de Fator IX de tipo selvagem).

TABELA 5: EDIÇÃO NO FÍGADO

ID do Animal	ID guia	F9-AAV (vg/kg)	Volume de F9-AAV (ml/kg)	LNP (mg/kg)	Volume de LNP (ml/kg)
4001	G009860	3E+13	1	3	2
4002	G009860	3E+13	1	3	2
4003	G009860	3E+13	1	3	2
5001	TSS	3E+13	1	0	0
5002	TSS	3E+13	1	0	0
5003	TSS	3E+13	1	0	0
6001	G009862	0	0	3	2
6002	G009862	0	0	3	2
6003	G009862	0	0	3	2

TABELA 6: EXPRESSÃO DE HFIX

ID do Animal	Dia 7 Fator IX (ng/ml)	Dia 14 Fator IX (ng/ml)	Dia 28 Fator IX (ng/ml)	Dia 56 Fator IX (ng/ml)
4001	122,84/+- 2,85	94,93/+- 0,56	105,65/+- 1,94	97,31/+-1,49
4002	149,77/+- 13,5	222,92/+- 9,61	252,49/+- 6,46	152,05/+- 7,46
4003	134,06/+- 6,17	107,04/+- 6,46	95,30/+- 3,18	74,23/+- 3,53
5001	ND	ND	ND	ND
5002	ND	ND	ND	ND
5003	ND	ND	ND	ND
6001	ND	ND	ND	ND
6002	ND	ND	ND	ND
6003	ND	ND	ND	ND

EXEMPLO 13- TESTE *IN VIVO* DA INSERÇÃO DE FATOR 9 EM PRIMATAS NÃO HUMANOS

[0298] Nesse exemplo, um estudo foi realizado para avaliar a inserção de gene de Fator 9 e expressão de proteína de hFIX em macacos cinomolgos depois da administração de ssAAV derivado de P00147 e/ou nanopartículas lipídicas de

CRISPR/Cas9 (LNP) com vários guias que incluem G009860 e vários componentes de LNP.

[0299] A formação de indel foi medida através de NGS, o que confirma que a edição ocorreu. Níveis de Fator IX humano totais foram determinados a partir de amostras plasmáticas através de ELISA com o uso de um mAB de camundongo para anticorpo de Fator IX humano (HTI, nº de cat AHIX-5041), anticorpo policlonal de Fator 9 anti-humano de ovelha (Abeam, nº de cat ab128048), e pAbs de IgG anti-ovelha de burro com HRP (Abeam, nº de cat ab97125), conforme descrito no Exemplo 12. Níveis de proteína de FIX humano >3 vezes superiores àqueles alcançados no experimento do Exemplo 12 foram obtidos a partir do modelo bidirecional com o uso de LNP de CRISPR/Cas9 alternativas. No estudo, os resultados de ensaio de ELISA indicam que os níveis de proteína de hFIX de circulação na faixa normal, ou acima da mesma, de níveis de FIX humano (3-5 ug/ml; Amiral et al, Clin. Chem., 30(9), 1512-16, 1984) foram alcançados com o uso de G009860 nas NHPs pelo menos nos pontos no tempo do dia 14 e 28. Dados iniciais indicam níveis de proteína de FIX humano de circulação de ~3-4 µg/ml no dia 14 após uma dose única, com níveis prolongados através dos primeiros 28 dias (~3-5 µg/ml) do estudo. Os níveis de FIX humano foram medidos no fim do estudo pelo mesmo método e os dados são apresentados na Tabela 25.

TABELA 25: NÍVEIS DE PROTEÍNA DE FATOR IX HUMANO EM SORO - MÉTODO ELISA DO EXEMPLO 13

	Dia 7 FIX ng/ml	DESVIO PADRÃO	Dia 14 FIX ng/ml	DESVIO PADRÃO	Dia 28 FIX ng/ml	DESVIO PADRÃO	Dia 42 FIX ng/ml	DESVIO PADRÃO	Dia 56 FIX ng/ml	DESVIO PADRÃO
3001	2532,8	145,6	2562,6	99,0	3011,7	62,7	2936,7	72,4	2748,5	86,0
3002	2211,4	95,8	2958,5	119,2	3350,2	98,4	3049,7	112,7	3036,7	90,6
3003	3195,1	475,6	4433,9	238,7	3367,2	157,7	3746,1	95,6	3925,0	157,4

[0300] Níveis de albumina de circulação foram medidos por ELISA, que indica que os níveis de albumina de linha de base são mantidos em 28 dias. Níveis de albumina testados em animais não tratados variaram ± ~15% no estudo. Nos animais tratados, níveis de albumina de circulação se alteraram minimamente e não caíram da faixa normal, e os níveis se recuperaram até a linha de base dentro de um mês.

[0301] Níveis de proteína de FIX humano de circulação também foram determinados por um imunoensaio de sanduíche com uma faixa dinâmica maior. Brevemente, uma Placa SECTOR de Estreptavidina de 96 poços MSD GOLD (Meso Scale Diagnostics, nº de Cat. L15SA-1) foi bloqueada com 1% de Agente de Bloqueio ECL (Sigma, GERPN2125). Após esvaziar a solução de bloqueio, o anticorpo de captura biotinilado (Sino Biological, 11503-R044) foi imobilizado na placa. Proteína de FIX humano recombinante (Enzyme Research Laboratories, HFIX 1009) foi usada para preparar um padrão de calibragem em 0,5% de Agente de Bloqueio ECL. Depois de uma lavagem, os padrões de calibragem e amostras plasmáticas foram adicionados à placa e incubados. Depois de uma lavagem, um anticorpo de detecção (Haematologic Technologies, AHIX-5041) conjugado com uma marca de sulfo-marcação foi adicionado aos poços e incubado. Após lavar qualquer anticorpo de detecção não ligado, Tampão de Leitura T foi aplicado aos poços. Sem nenhuma incubação adicional, a placa foi imageada com um instrumento MSD Quick Plex SQ120 e os dados foram analisados com o pacote de software Discovery Workbench 4.0 (Meso Scale Discovery). Concentrações são expressas como concentrações médias calculadas em ug/m. Para as amostras, N = 3 a menos que seja indicado com um asterisco, nesse caso N = 2. Expressão de hFIX a partir do lócus de albumina no grupo de estudo tratado, conforme medido por ELISA de MSD, é representada na Tabela 26.

TABELA 26: NÍVEIS DE PROTEÍNA DE FATOR IX HUMANO EM SORO

	Média Calculada Cone. (ug/ml)					
	3001		3002		3003	
Ponto de Tempo	Cone.	Interensaio CV	Cone.	Interensaio CV	Cone.	Interensaio CV
Dia 7	7,85	20%	5,63	14%	11,20	26%
Dia 14	8,65	15%	11,06	18%	14,70	28%
Dia 28	9,14	7%	14,12	7%	10,85	25%
Dia 42	9,03	10%	33,12*	0%	13,22	13%
Dia 56	10,24	13%	16,72	12%	33,84*	4%

EXEMPLO 14 - ANÁLISE FORA DO ALVO DE GUIAS HUMANOS DE ALBUMINA

[0302] Um método bioquímico (Consultar, por exemplo, Cameron et al., *Nature Methods*. 6, 600-606; 2017) foi usado para determinar potenciais sítios genômicos fora do alvo clivados por Cas9 que alveja albumina. Nesse experimento, 13 sgRNA que alvejam albumina humana e dois guias de controle com perfis fora de alvo conhecidos foram examinados com o uso de DNA genômico de HEK293 isolado. O número de potenciais sítios fora do alvo detectados com o uso de uma concentração guia de 16 nM no ensaio bioquímico foram mostrados na Tabela 27. O ensaio identificou potenciais sítios fora do alvo para os sgRNAs testados.

TABELA 27: ANÁLISE FORA DO ALVO

ID de gRNA	Alvo	Sequência Guia (SEQ ID NO:)	Fora do Alvo Contagem de Sítio
GO 12753	Albumina	GACUGAAACUUCACAGAAUA (SEQ ID NO: 20)	62
GO 12761	Albumina	AGUGCAAUGGAUAGGUCUUU (SEQ ID NO: 28)	75
G012752	Albumina	UGACUGAAACUUCACAGAAU (SEQ ID NO: 19)	223

ID de gRNA	Alvo	Sequência Guia (SEQ ID NO:)	Fora do Alvo Contagem de Sítio
GO 12764	Albumina	CCUCACUCUUGUCUGGGCAA (SEQ ID NO: 31)	3985
G012763	Albumina	UGGGCAAGGGAAGAAAAAAA (SEQ ID NO: 30)	5443
G009857	Albumina	AUUUAUGAGAUCAACAGCAC (SEQ ID NO: 5)	131
G009859	Albumina	UUAAAUAAAGCAUAGUGCAA (SEQ ID NO: 7)	91
G009860	Albumina	UAAAGCAUAGUGCAAUGGAU (SEQ ID NO: 8)	133
GO 12762	Albumina	UGAUUCCUACAGAAAAACUC (SEQ ID NO: 29)	68
G009844	Albumina	GAGCAACCUCACUCUUGUCU (SEQ ID NO: 2)	107
G012765	Albumina	ACCUCACUCUUGUCUGGGCA (SEQ ID NO: 32)	41
G012766	Albumina	UGAGCAACCUCACUCUUGUC (SEQ ID NO: 33)	78
G009874	Albumina	UAAUAAAAUUCAAACAUCCU (SEQ ID NO: 13)	53
G000644	EMX1	GAGUCCGAGCAGAAGAAGAA (SEQ ID NO 1129)	304
G000645	VEGFA	GACCCCCUCCACCCCGCCUC (SEQ ID NO 1130)	1641

[0303] Em ensaios de detecção fora do alvo conhecidos, tais como o método bioquímico usado acima, um grande número de potenciais sítios fora do alvo são tipicamente recuperados, em concepção, de modo a “lançar uma rede ampla” para potenciais sítios que podem ser validados em outros contextos, por exemplo, em uma célula primária de interesse. Por exemplo, o método bioquímico sobre-representa, tipicamente, o número de potenciais sítios fora do alvo conforme o ensaio utiliza DNA genômico de alto peso molecular purificado livre do ambiente celular e é dependente da dose de RNP de Cas9 usada. Consequentemente, potenciais sítios fora do alvo

identificados por esse método são validados com o uso de sequenciamento alvejado dos potenciais sítios fora do alvo identificados.

EXEMPLO 15 CONSTRUÇÃO DE CONSTRUTOS PARA A EXPRESSÃO DE PROTEÍNAS SECRETORAS OU NÃO SECRETORAS

[0304] Construtos, tais como construtos bidirecionais, podem ser projetados de modo que os mesmos expressam proteínas secretoras ou não secretoras. Para a produção de uma proteína secretora, um construto pode compreender uma sequência de sinal que auxilia na translocação do polipeptídeo para o lúmen de ER. Alternativamente, um construto pode utilizar a sequência de sinal endógena da célula hospedeira (por exemplo, a sequência de sinal endógena de albumina quando o transgene é integrado em um lócus de albumina da célula hospedeira).

[0305] Em contraste, os construtos para a expressão de proteínas não secretoras podem ser projetados de modo que os mesmos não compreendam uma sequência de sinal e de modo que os mesmos não utilizem a sequência de sinal endógena da célula hospedeira. Alguns métodos pelos quais isso pode ser alcançado incluem a incorporação de uma sequência de sítio de entrada de ribossomo interno (IRES) no construto. Sequências de IRES, tais como IRES de EMCV, permitem a iniciação da tradução a partir de qualquer posição dentro de um mRNA imediatamente a jusante de onde o IRES está localizado. Isso permitiria a expressão de uma proteína que não tem a sequência de sinal endógena da célula hospedeira de um sítio de inserção que contém uma sequência de sinal a montante (por exemplo, a sequência de sinal encontrada no Éxon 1 do lócus de albumina não seria incluída na proteína expressa). Na ausência de uma sequência de sinal, a proteína não seria secretada. Os exemplos de sequências de IRES que podem ser usadas em um construto incluem aquelas de picornavírus (por exemplo, FMDV), vírus da praga (CFFV), poliovírus (PV), vírus da encefalomiocardite (ECMV), vírus da doença pé e boca (FMDV), vírus da hepatite C (HCV), vírus da febre suína clássica (CSFV), vírus da leucemia murina

(MLV), vírus da imunodeficiência símia (SIV) ou vírus da paralisia do grilo (CrPV).

[0306] Uma abordagem alternativa para expressar proteínas não secretoras deve incluir um ou mais peptídeos de autoclivagem a montante do polipeptídeo de interesse no construto. Um peptídeo de autoclivagem, tal como sequências do tipo 2A ou 2A, servem como sinais de salto de ribossomo para produzir múltiplas proteínas individuais a partir de um único transcrito de mRNA. Conforme mostrado no ID de Plasmídeo P00415 da Tabela 11, um peptídeo de autoclivagem (por exemplo, P2A) pode ser usado para gerar um vetor bicistrônico que expressa dois transgenes (por exemplo, nanoluciferase e GFP). Alternativamente, um peptídeo de autoclivagem pode ser usado para expressar uma proteína que não tem a sequência de sinal endógena da célula hospedeira (por exemplo, a sequência 2A localizada a montante da proteína de interesse resultaria na clivagem entre a sequência de sinal endógena de albumina e a proteína de interesse). Os peptídeos 2A representativos que poderiam ser utilizados são mostrados na Tabela 28. Adicionalmente, os resíduos (GSG) podem ser adicionados à extremidade 5' do peptídeo para melhorar a eficiência de clivagem conforme mostrado na Tabela 12.

Tabela 28: Peptídeos de autoclivagem para uso em construtos

Peptídeo	Sequência de Aminoácidos
T2A (SEQ ID NO: 1131)	EGRGSLLTCGDVEENPGP
P2A (SEQ ID NO: 1132)	ATNFSLLKQAGDVEENPGP
E2A (SEQ ID NO: 1133)	QCTNYALLKLAGDVESNPGP
F2A (SEQ ID NO: 1134)	VKQTLNFDLLKLAGDVESNPGP
T2A com resíduos GSG (SEQ ID NO: 1135)	GSGEGRGSLLTCGDVEENPGP
P2A com resíduos GSG (SEQ ID NO: 1136)	GSGATNFSLLKQAGDVEENPGP
E2A com resíduos GSG (SEQ ID NO: 1137)	GSGQCTNYALLKLAGDVESNPGP
F2A com resíduos GSG (SEQ ID NO: 1138)	GSGVKQTLNFDLLKLAGDVESNPGP

EXEMPLO 16. USO DE CAMUNDONGOS DE ALBUMINA HUMANIZADA PARA EXAMINAR RNAS GUIA PARA INSERÇÃO DE *F9* HUMANO *IN VIVO*

[0307] Os inventores buscaram identificar RNAs guia eficazes para inserção de *hF9* no lócus de albumina humana. Para essa finalidade, os inventores utilizaram camundongos nos quais o lócus de albumina de camundongo foi substituído pela sequência genômica de albumina humana correspondente, que inclui o primeiro íntron (camundongos $ALB^{hu/hu}$). Isso permitiu aos inventores testar a eficiência de inserção de RNAs guia que alvejam o primeiro íntron de albumina humana no contexto de um fígado adulto *in vivo*. Dois experimentos de camundongo separados foram montados com o uso dos camundongos $ALB^{hu/hu}$ para examinar um total de 11 RNAs guia, em que cada um alveja o primeiro íntron do lócus de albumina humana. Todos os camundongos foram ponderados e injetados por meio da veia caudal no dia 0 do experimento. O sangue foi coletado nas semanas 1, 3, 4 e 6 por meio de sangramento caudal, e plasma foi separado. Camundongos foram finalizados na semana 7. O sangue foi coletado por meio da veia cava, e plasma foi separado. Fígados e baços foram dissecados também.

[0308] No primeiro experimento, 6 LNPs que compreendem mRNA de Cas9 e os seguintes guias foram preparados como no Exemplo 1 e testados: G009852, G009859, G009860, G009864, G009874 e G012764. LNPs foram diluídas em 0,3 mg/kg (com o uso de um peso médio de 30 gramas) e coinjetadas com AAV8 acondicionado com o modelo de inserção de hF9 bidirecional em uma dose de genomas virais 3E11 por camundongo. Cinco camundongos machos $ALB^{hu/hu}$ entre 12 e 14 semanas de idade foram injetados por grupo. Cinco camundongos do mesmo coorte foram injetados com AAV8 acondicionado com um promotor CAGG operacionalmente ligado ao hF9, que resulta em expressão episssômica de hF9 (em genomas virais 3E11 por camundongo). Houveram três grupos de controle negativo com três camundongos por grupo que foram injetados com tampão apenas, AAV8

acondicionado com o modelo de inserção de hF9 bidirecional apenas ou LNP-G009874 apenas.

[0309] No experimento, as seguintes LNPs que compreendem mRNA de Cas9 e os seguintes guias foram preparados como no Exemplo 1 e testados: G009860, G012764, G009844, G009857, G012752, G012753 e G012761. Todas foram diluídas em 0,3 mg/kg (com o uso de um peso médio de 40 gramas) e coinjetadas com AAV8 acondicionado com o modelo de inserção de hF9 bidirecional em uma dose de genomas virais 3E11 por camundongo. Cinco camundongos machos $ALB^{hu/hu}$ de 30 semanas de idade foram injetados por grupo. Cinco camundongos do mesmo coorte foram injetados com AAV8 acondicionado com um promotor CAGG operacionalmente ligado ao hF9, que resulta em expressão episssômica de hF9 (em genomas virais 3E11 por camundongo). Houveram três grupos de controle negativo com três camundongos por grupo que foram injetados com tampão apenas, AAV8 acondicionado com o modelo de inserção de hF9 bidirecional apenas ou LNP-G009874 apenas.

[0310] Para análise, um ELISA foi realizado para medir níveis de hFIX que circula nos camundongos em cada ponto no tempo. Kits de ELISA de Fator IX Humano (ab188393) foram usados com essa finalidade, e todas as placas foram executadas com plasma normal agrupado humano da George King Bio-Medical como um controle de ensaio positivo. Níveis de expressão de Fator IX Humano nas amostras plasmáticas em cada grupo na semana 6 depois da injeção são mostrados na Figura 16A e Figura 16B. Consistente com os dados de inserção *in vitro*, níveis de soro de Fator IX de baixos a nulos foram detectados quando o RNA guia G009852 foi usado. Consistente com a falta de uma sequência de PAM adjacente em albumina humana, níveis de soro de Fator IX não foram detectáveis quando o RNA guia G009864 foi usado. Expressão de Fator IX no soro foi observada para os grupos que usam RNAs guia G009859, G009860, G009874 e G0012764.

[0311] Baços e uma porção do lóbulo lateral esquerdo de todos os fígados

foram submetidos à análise de sequenciamento da próxima geração (NGS). NGS foi usado para avaliar a porcentagem de células hepáticas com inserções/deleções (indels) no lócus de albumina humanizada na semana 7 depois da injeção com doador de *hF9* de AAV e CRISPR/Cas9 de LNP. Consistente com a falta de uma sequência de PAM adjacente em albumina humana, nenhuma edição foi detectável no fígado quando o RNA guia G009864 foi usado. A edição no fígado foi observada para os grupos que usam RNAs guia G009859, G009860, G009874 e G012764 (dados não mostrados).

[0312] O fígado restante foi corrigido por 24 horas em formalina tamponada 10% neutra e, então, transferido para 70% de etanol. Quatro a cinco amostras de lóbulos separados foram cortados e enviados para HistoWisz e foram processadas e incorporadas em blocos de parafina. Seções de cinco mícrons foram, então, cortadas a partir de cada bloco de parafina, e BASESCOPE™ foi realizado no Ventana Ultra Discovery (Roche) com o uso do procedimento BASESCOPE™ universal e reagentes pela Advanced Cell Diagnostics e uma sonda projetada personalizada que alveja a junção de mRNA única formada entre a sequência de sinal de albumina humana a partir do primeiro íntron do lócus de albumina $ALB^{hu/hu}$ e do transgene de hF9 quando a integração e transcrição bem-sucedida é alcançada. Software de imageamento HALO (Indica Labs) foi, então, usado para quantificar a porcentagem de células positivas em cada amostra. A média da porcentagem de células positivas através dos múltiplos lóbulos para cada animal estava, então, correlacionada aos níveis de hFIX no soro na semana 7. Os resultados são mostrados na Figura 17 e Tabela 29. Os níveis de soro da semana 7 e a % de células positivas para o mRNA de hALB-hFIX se correlacionaram fortemente ($r = 0,89$; $R^2 = 0,79$).

TABELA 29. HFIX DA SEMANA 7 E DADOS BASESCOPE™.

Camundongo	Guia	hFIX ug/ml (Semana 7)	% de Sonda de mRNA (Seções 4-5)	% Padrão de Sonda de mRNA	Células Totais Contadas
1	Buffer	ND	0,09	0,03	152833

Camundongo	Guia	hFIX ug/ml (Semana 7)	% de Sonda de mRNA (Seções 4-5)	% Padrão de Sonda de mRNA	Células Totais Contadas
4	AAV Apenas	ND	0,53	0,67	351084
7	LNP Apenas	ND	0,48	0,33	75160
10	CAGF9	211,8	0,20	0,22	190277
15	G009852	ND	0,30	0,09	144518
20	G009859	0,5	0,82	0,45	143817
21	G009859	0,5	0,88	0,43	160172
22	G009859	2,3	1,71	1,54	26015
23	G009859	3,8	2,74	0,59	183085
24	G009859	0,6	2,78	1,96	152424
25	G009860	5,6	12,46	5,80	78935
26	G009860	10,6	13,76	5,32	112252
27	G009860	9,7	14,80	5,45	201592
28	G009860	2,1	3,32	0,76	84710
29	G009860	3,0	1,52	0,35	203277
30	G009864	ND	1,94	1,78	145807
35	G009874	1,7	2,42	1,14	126665
36	G009874	1,5	1,08	0,53	195861
37	G009874	2,1	1,02	1,29	181679
38	G009874	5,5	0,40	0,43	175359
39	G009874	1,5	0,44	0,18	205417
40	G012764	15,7	28,85	7,11	167824
41	G012764	19,6	19,17	8,23	70081
42	G012764	1,9	1,95	1,79	154742
43	G012764	7,7	4,38	0,68	114060
44	G012764	3,0	1,64	1,04	238623
43	DapB (-)	—	0,12	0,07	144730

ÍNTRON 1 DE ALBUMINA HUMANA: (SEQ ID NO: 1)

GTAAGAAATCCATTTTTCTATTGTTCAACTTTTATTCTATTTTCCCAGTAA

AATAAAGTTTTAGTAAACTCTGCATCTTTAAAGAATTATTTTGGCATTTATTTCTAA
AATGGCATAGTATTTTGTATTTGTGAAGTCTTACAAGGTTATCTTATTAATAAAAT
TCAAACATCCTAGGTAAAAAAAAAAAAAGGTCAGAATTGTTTAGTGACTGTAATTT
TCTTTTGCGCACTAAGGAAAGTGCAAAGTAACTTAGAGTGACTGAAACTTCACAG
AATAGGGTTGAAGATTGAATTCATAACTATCCCAAAGACCTATCCATTGCACTAT
GCTTTATTTAAAAACCACAAAACCTGTGCTGTTGATCTCATAAATAGAACTTGTAT
TTATATTTATTTTCATTTTAGTCTGTCTTCTTGGTTGCTGTTGATAGACACTAAAA
GAGTATTAGATATTATCTAAGTTTGAATATAAGGCTATAAATATTTAATAATTTTTA
AAATAGTATTCTTGGTAATTGAATTATTCTTCTGTTTAAAGGCAGAAGAAATAATT
GAACATCATCCTGAGTTTTTCTGTAGGAATCAGAGCCCAATATTTTGAAACAAAT
GCATAATCTAAGTCAAATGGAAAGAAATATAAAAAGTAACATTATTACTTCTTGTT
TTCTTCAGTATTTAACAATCCTTTTTTTTCTTCCCTTGCCCAG

TABELA 7. RNA GUIA DE ALBUMINA DE CAMUNDONGO

ID guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCCUU	chr5:90461148-90461168	98
G000552	AUCGGGAACUGGCAUCUUCA	chr5:90461590-90461610	99
G000553	GUUACAGGAAAAUCUGAAGG	chr5:90461569-90461589	100
G000554	GAUCGGGAACUGGCAUCUUC	chr5:90461589-90461609	101
G000555	UGCAUCUGAGAACCCUUAGG	chr5:90461151-90461171	102
G000666	CACUCUUGUCUGUGGAAACA	chr5:90461709-90461729	103
G000667	AUCGUUACAGGAAAAUCUGA	chr5:90461572-90461592	104
G000668	GCAUCUUCAGGGAGUAGCUU	chr5:90461601-90461621	105
G000669	CAAUCUUUAAAUAUGUUGUG	chr5:90461674-90461694	106
G000670	UCACUCUUGUCUGUGGAAAC	chr5:90461710-90461730	107
GO 11722	UGCUUGUAUUUUUCUAGUAA	chr5:90461039-90461059	108
G011723	GUAAAUAUCUACUAAGACAA	chr5:90461425-90461445	109
GO 11724	UUUUUCUAGUAAUGGAAGCC	chr5:90461047-90461067	110
G011725	UUAUAUUAUUGAUAUAUUUU	chr5:90461174-90461194	111

ID guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
GO 11726	GCACAGAUAUAAACACUUAA	chr5:90461480-90461500	112
GO 11727	CACAGAUAUAAACACUUAAC	chr5:90461481-90461501	113
G011728	GGUUUUAAAAAUAAUAAUGU	chr5:90461502-90461522	114
GO 11729	UCAGAUUUUCCUGUAACGAU	chr5:90461572-90461592	115
G011730	CAGAUUUUCCUGUAACGAUC	chr5:90461573-90461593	116
G011731	CAAUGGUAAAUAAGAAAUAA	chr5:90461408-90461428	117
G013018	GGAAAAUCUGAAGGUGGCAA	chr5:90461563-90461583	118
G013019	GGCGAUCUCACUCUUGUCU G	chr5:90461717-90461737	119

TABELA 8. SGRNAS DE ALBUMINA DE CAMUNDONGO E PADRÃO DE MODIFICAÇÃO

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCCUUGU UUUAGAGCUAGAAAUAGCAAGU UAAAAUAAGGCUAGUCCGUUAU CAACUUGAAAAAGUGGCACCGA GUCGGUGCUUUU	120	mA*mU*mU*UGCAUCUGAGAACC CUUGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	142
G000552	AUCGGGAACUGGCAUCUUCA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	121	mA*mU*mC*GGGAACUGGCAUCU UCAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	143

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
GOOO553	GUUACAGGAAAAUCUGAAGG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	122	mG*mU*mU*ACAGGAAAAUCUGA AGGGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	144
G000554	GAUCGGGAACUGGCAUCUUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	123	mG*mA*mU*CGGGAACUGGCAUC UUCGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	145
G000555	UGCAUCUGAGAACCCUUAGG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	124	mU*mG*mC*AUCUGAGAACCCUU AGGGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	146
G000666	CACUCUUGUCUGUGGAAACA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	125	mC*mA*mC*UCUUGUCUGUGGAA ACAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	147
G000667	AUCGUUACAGGAAAAUCUGA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	126	mA*mU*mC*GUUACAGGAAAAUC UGAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	148

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G000668	GCAUCUUCAGGGAGUAGCUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	127	mG*mC*mA*UCUUCAGGGAGUAG CUUGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	149
G000669	CAAUCUUUAAAUAUGUUGUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	128	mC*mA*mA*UCUUUAAAUAUGUU GUGGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	150
G000670	UCACUCUUGUCUGUGGAAAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	129	mU*mC*mA*CUCUUGUCUGUGGA AACGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	151
GO 11722	UGCUUGUAUUUUUCUAGUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	130	mU*mG*mC*UUGUAUUUUUCUAG UAAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	152
G011723	GUAAAUAUCUACUAAGACAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	131	mG*mU*mA*AAUAUCUACUAAGA CAAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	153

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
GO 11724	UUUUUCUAGUAAUGGAAGCC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	132	mU*mU*mU*UUCUAGUAAUGGAA GCCGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	154
G011725	UUAUAUUAUUGAUAUAUUUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	133	mU*mU*mA*UAUUAUUGAUAUAU UUUGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	155
GO 11726	GCACAGAUAUAAACACUUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	134	mG*mC*mA*CAGAUAUAAACACU UAAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	156
GO 11727	CACAGAUAUAAACACUUAAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	135	mC*mA*mC*AGAUAUAAACACUU AACGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	157
G011728	GGUUUUAAAAAUAAUAAUGU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	136	mG*mG*mU*UUUAAAAAUAAUAA UGUGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	158

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G011729	UCAGAUUUUCCUGUAACGAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	137	mU*mC*mA*GAUUUUCCUGUAAC GAUGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	159
G011730	CAGAUUUUCCUGUAACGAUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	138	mC*mA*mG*AUUUUCCUGUAACG AUCGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	160
G011731	CAAUGGUAAAUAAGAAAUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	139	mC*mA*mA*UGGUAAAUAAGAAA UAAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	161
G013018	GGAAAAUCUGAAGGUGGCAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	140	mG*mG*mA*AAAUCUGAAGGUGG CAAGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	162
G013019	GGCGAUCUCACUCUUGUCUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	141	mG*mG*mC*GAUCUCACUCUUGU CUGGUUUUAGAmGmCmUmAmGm AmAmAmUmAmGmCAAGUUAAAA UAAGGCUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAmGmUm GmGmCmAmCmCmGmAmGmUmCm GmGmUmGmCmU*mU*mU*mU	163

TABELA 9. RNA GUIA DE ALBUMINA CINO

ID guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU	chr5:61198711-61198731	164
G009845	AGCAACCUCACUCUUGUCUG	chr5:61198712-61198732	165
G009846	ACCUCACUCUUGUCUGGGGA	chr5:61198716-61198736	166
G009847	CCUCACUCUUGUCUGGGGAA	chr5:61198717-61198737	167
G009848	CUCACUCUUGUCUGGGGAAG	chr5:61198718-61198738	168
G009849	GGGGAAGGGGAGAAAAAAAA	chr5:61198731-61198751	169
G009850	GGGAAGGGGAGAAAAAAAAA	chr5:61198732-61198752	170
G009851	AUGCAUUUGUUUCAAAAUAU	chr5:61198825-61198845	171
G009852	UGCAUUUGUUUCAAAAUAUU	chr5:61198826-61198846	172
G009853	UGAUUCCUACAGAAAAAGUC	chr5:61198852-61198872	173
G009854	UACAGAAAAAGUCAGGAUAA	chr5:61198859-61198879	174
G009855	UUUCUUCUGCCUUUAAACAG	chr5:61198889-61198909	175
G009856	UUAUAGUUUUAUAUUCAAAC	chr5:61198957-61198977	176
G009857	AUUUAUGAGAUCAACAGCAC	chr5:61199062-61199082	177
G009858	GAUCAACAGCACAGGUUUUG	chr5:61199070-61199090	178
G009859	UUAAAUAAAGCAUAGUGCAA	chr5:61199096-61199116	179
G009860	UAAAGCAUAGUGCAAUGGAU	chr5:61199101-61199121	180
G009861	UAGUGCAAUGGAUAGGUCUU	chr5:61199108-61199128	181
G009862	AGUGCAAUGGAUAGGUCUUA	chr5:61199109-61199129	182

ID guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
ID guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G009863	UUACUUUGCACUUUCCUUAG	chr5:61199186-61199206	183
G009864	UACUUUGCACUUUCCUUAGU	chr5:61199187-61199207	184
G009865	UCUGACCUUUUAUUUUACCU	chr5:61199238-61199258	185
G009866	UACUAAAACUUUAUUUUACU	chr5:61199367-61199387	186
G009867	AAAGUUGAACAAUAGAAAAA	chr5:61199401-61199421	187
G009868	AAUGCAUAAUCUAAGUCAAA	chr5:61198812-61198832	188
G009869	AUUAUCCUGACUUUUUCUGU	chr5:61198860-61198880	189
G009870	UGAAUUAUUCCUCUGUUUAA	chr5:61198901-61198921	190
G009871	UAAUUUUCUUUUGCCCACUA	chr5:61199203-61199223	191
G009872	AAAAGGUCAGAAUUGUUUAG	chr5:61199229-61199249	192
G009873	AACAUCCUAGGUAAAAUAAA	chr5:61199246-61199266	193
G009874	UAAUAAAAUUCAAACAUCCU	chr5:61199258-61199278	194
G009875	UUGUCAUGUAUUUCUAAAAU	chr5:61199322-61199342	195
G009876	UUUGUCAUGUAUUUCUAAAA	chr5:61199323-61199343	196

TABELA 10: SGRNA CINO E PADRÕES DE MODIFICAÇÃO

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	197	mG*mA*mG*CAACCUCACUCUU GUCUGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCm GmAmGmUmCmGmGmUmGmCm U*mU*mU*mU	230
G009845	AGCAACCUCACUCUUGUCUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	198	mA*mG*mC*AACCUCACUCUUC UCUGGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCm GmAmGmUmCmGmGmUmGmCm U*mU*mU*mU	231
G009846	ACCUCACUCUUGUCUGGGGA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	199	mA*mC*mC*UCACUCUUGUCUC GGGAGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAA GUUAAAAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCm CmGmAmGmUmCmGmGmUmGm CmU*mU*mU*mU	232
G009847	CCUCACUCUUGUCUGGGGAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	200	mC*mC*mU*CACUCUUGUCUGG GGAAGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCm GmAmGmUmCmGmGmUmGmCm U*mU*mU*mU	233

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009848	CUCACUCUUGUCUGGGGAAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	201	mC*mU*mC*ACUCUUGUCUGG GGAAGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAA GUUAAAAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCmAmCm CmGmAmGmUmCmGmGmUmGm CmU*mU*mU*mU	234
G009849	GGGGAAGGGGAGAAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	202	mG*mG*mG*GAAGGGGAGAAA AAAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	235
G009850	GGGAAGGGGAGAAAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	203	mG*mG*mG*AAGGGGAGAAAA AAAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	236
G009851	AUGCAUUUGUUUCAAAAUAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	204	mA*mU*mG*CAUUUGUUUCAA AAUAUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	237

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009852	UGCAUUUGUUUCAAAAUAUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	205	mU*mG*mC*AUUUGUUUCAAA AUAUUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	238
G009853	UGAUUCCUACAGAAAAAGUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	206	mU*mG*mA*UUCCUACAGAAA AAGUCGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	239
G009854	UACAGAAAAAGUCAGGAUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	207	mU*mA*mC*AGAAAAAGUCAG GAUAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	240
G009855	UUUCUUCUGCCUUUAAACAG GUUUUAGAGCUAGAAAUAGC	208	mU*mU*mU*CUUCUGCCUUUA AACAGGUUUUAGAmGmCmUm	241
	AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU		AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009856	UUAUAGUUUUAUAUUCAAAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	209	mU*mU*mA*UAGUUUUAUAUU CAAACGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	242
G009857	AUUUAUGAGAUCAACAGCAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	210	mA*mU*mU*UAUGAGAUCAAC AGCACGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	243
G009858	GAUCAACAGCACAGGUUUUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	211	mG*mA*mU*CAACAGCACAGG UUUUGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	244
G009859	UUAAAUAAAGCAUAGUGCAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	212	mU*mU*mA*AAUAAAGCAUAG UGCAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	245

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009860	UAAAGCAUAGUGCAAUGGAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	213	mU*mA*mA*AGCAUAGUGCAA UGGAUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	246
G009861	UAGUGCAAUGGAUAGGUCUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	214	mU*mA*mG*UGCAAUGGAUAG GUCUUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	247
G009862	AGUGCAAUGGAUAGGUCUUA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	215	mA*mG*mU*GCAAUGGAUAGG UCUUAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	248
G009863	UUACUUUGCACUUUCCUUAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	216	mU*mU*mA*CUUUGCACUUUC CUUAGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	249
G009864	UACUUUGCACUUUCCUUAGU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC	217	mU*mA*mC*UUUGCACUUUCC UUAGUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAC	250

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU		UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	
G009865	UCUGACCUUUUAUUUUACCU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	218	mU*mC*mU*GACCUUUUAUUU UACCUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	251
G009866	UACUAAAACUUUAUUUUACU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	219	mU*mA*mC*UAAAACUUUAUU UUACUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	252
G009867	AAAGUUGAACAAUAGAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	220	mA*mA*mA*GUUGAACAAUAG AAAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	253

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009868	AAUGCAUAAUCUAAGUCAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	221	mA*mA*mU*GCAUAAUCUAAG UCAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	254
G009869	AUUAUCCUGACUUUUUCUGU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	222	mA*mU*mU*AUCCUGACUUUU UCUGUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	255
G009870	UGAAUUAUUCCUCUGUUUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	223	mU*mG*mA*AUUAUUCCUCUG UUUAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	256
G009871	UAAUUUUCUUUUGCCCACUA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	224	mU*mA*mA*UUUUCUUUUGCC CACUAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	257

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009872	AAAAGGUCAGAAUUGUUUAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	225	mA*mA*mA*AGGUCAGAAUUG UUUAGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	258
G009873	AACAUCCUAGGUAAAAUAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	226	mA*mA*mC*AUCCUAGGUAAA AUAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	259
G009874	UAAUAAAAUUCAAACAUCCU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	227	mU*mA*mA*UAAAAUUCAAAC AUCCUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	260
G009875	UUGUCAUGUAUUUCUAAAAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	228	mU*mU*mG*UCAUGUAUUUCU AAAAUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	261

ID guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009876	UUUGUCAUGUAUUUCUAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	229	mU*mU*mU*GUCAUGUAUUUC UAAAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	262

TABELA 11: COMPONENTES DE VETOR E SEQUÊNCIAS

ID de Plasmídeo	ITR 5’	Aceitante de Splice (1ª orientação)	Transgene (1ª orientação)	Poli-A (1ª orientação)	Polo-A (2ª orientação)	Transgene (2ª orientação)	Aceitante de Splice (2ª orientação)	ITR 3’
P00147	(SEQ ID NO: 263)	Aceitante de Splice de Albumina de Camundongo (SEQ ID NO: 264)	Fator IX Humano (R338L) (SEQ ID NO: 265)	SEQ ID NO: 266	SEQ ID NO: 267	Fator IX Humano (R338L) (SEQ ID NO: 268)	Aceitante de Splice de Albumina de Camundongo (SEQ ID NO: 269)	(SEQ ID NO: 270)
P00411	(SEQ ID NO: 263)	Aceitante de Splice de Fator IX Humano (SEQ ID NO: 271)	Fator IX Humano (R338L)-HiBit (SEQ ID NO: 272)	SEQ ID NO: 266	SEQ ID NO: 267	Fator IX Humano (R338L)-HiBit (SEQ ID NO: 273)	Aceitante de Splice de Fator IX Humano (SEQ ID NO: 274)	(SEQ ID NO: 270)
P00415	(SEQ ID NO: 263)	Aceitante de Splice de Albumina de Camundongo (SEQ ID	Nluc-P2A-GFP (SEQ ID NO: 275)	SEQ ID NO: 266	SEQ ID NO: 267	Nluc-P2A-GFP (SEQ ID NO: 276)	Aceitante de Splice de Albumina de Camundongo (SEQ ID	(SEQ ID NO: 270)

ID de Plasmídeo	ITR 5’	Aceitante de Splice (1ª orientação)	Transgene (1ª orientação)	Poli-A (1ª orientação)	Polo-A (2ª orientação)	Transgene (2ª orientação)	Aceitante de Splice (2ª orientação)	ITR 3’
		NO: 264)					NO: 269)	
P00418	(SEQ ID NO: 263)	Aceitante de Splice de Albumina de Camundongo (SEQ ID NO: 264)	Fator IX Humano (R338L)-HiBit (SEQ ID NO: 272)	SEQ ID NO: 266	SEQ ID NO: 267	Fator IX Humano (R338L)-HiBit (SEQ ID NO: 273)	Aceitante de Splice de Albumina de Camundongo (SEQ ID NO: 269)	(SEQ ID NO: 270)

Sequência de ITR de 5’ (SEQ ID NO: 263):

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCA GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT

Aceitante de Splice de Albumina de Camundongo (1ª orientação) (SEQ ID NO: 264):

TAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTC TTCAT CAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAG

Fator IX Humano (R338L) 1ª Orientação (SEQ ID NO: 265):

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTA TAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATG GAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAA CAACTGAATTTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCAT GTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCC CTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAAT GGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCC TGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTG

CCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTG
AGACTGTTTTTCCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTT
GGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGT
GGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAA
GTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTG
CCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATAT
TGAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCA
CCACAACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAA
CTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACA
AGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGG
GAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCC
ACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAAC
ATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGT
GGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATT
AGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTA
TCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAA

Poli-A (1ª orientação) (SEQ ID NO: 266):

CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA
TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGG
GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT
GGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCT
CTAGGGGGTATC CCC

Poli-A (2ª orientação) (SEQ ID NO: 267):

AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCA
ATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAG
CATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGT

CCAAA CTCATCAATGTATCTTATCATGTCTG

Fator IX Humano (R338L) 2ª Orientação (SEQ ID NO: 268):

TTAGGTGAGCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCT
TCGTATAGATGCCATATTTCCCCTTCATCGCACATTCCTCCCCCCAACTTATTAT
CCCGGTCAAGAAACTTGTTCCTTCGACTTCAGTGACGTGTGGTCCACCTGAATC
ACCTTGGCATGAGTCGCGACCGCCCTCGTGAAACCCAGCACAAAACATGTTATT
GTAAATCGTAAATTTCGTGGACAGAAGACAGGTCGCTCTATCGACCAACGGGAC
GCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTGTGGAAGACCCGCCCCC
ACCCACTCACATATCCGCTCCCAAATTTCAAGAAGATATTTGTATATTCTTTATCG
GCTATACAAATCGGGGTAACATAGGAGTTAAGTACGAGTGGCTCGTCCAGCTCC
AGGAGGGCTATATCATGGTTGTACTTGTTTATAGCGGCATTATAATTGTGATGGG
GTATGATCCTGATAACATTCCTTTTCTGTTCAGTATGCTCAGTTTCTTCAATGTTG
TGTTCGCCAGCCACGACCGTAATCTTAACCCCCGTCTCGACACAGTGTGCGGC
CGTTACAATCCACTTTTCATTGACTATGGAGCCCCCACAAAACGCGTCGACTTTT
CCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCCTCGCCCCC
GACAACCCTAGTAAAGTCATTAAATGACTGTGTGGATTGTGTTATATTATCAAGA
ATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCTCG
GCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGG
CACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCA
ACTACACACAACTTTGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCA
TTTTTAATGTTGCAGGTGACGTCCAACTCGCAGTTTTTTCCTTCAAAACCAAAAG
GGCACCAACACTCGTAGGAATTTATATCGTCTTTACAACTCCCCCCATTCAGACA
TGGATTAGATTCGCATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGGT
CCGTTCTGTATTCTCAAACACCTCGCGCGCTTCTTCAAAACTGCATTTTTCCTCC
ATACACTCTCGCTCCAAGTTCCCTTGCACGAATTCTTCAAGCTTTCCTGAGTTAT
ACCTTTTAGGCCGGTTAAGTATCTTATTCGCGTTTTCGTGGTCCAGAAA

Aceitante de Splice de Albumina de Camundongo (2ª orientação) (SEQ ID NO:

269):

CTGTGGAAACAGGGAGAGAAAAACCACACAACATATTTAAAGATTGATG
AAGAC AACTAACTGTAATATGCTGCTTTTTGTTCTTCTCTTCACTGACCTA

Sequência de ITR de 3’ (SEQ ID NO: 270):

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCT
CGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCG
CCCGGCCTC AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Aceitante de Splice de Fator IX Humano (1ª Orientação) (SEQ ID NO: 271):

GATTATTTGGATTAAAAACAAAGACTTTCTTAAGAGATGTAAAATTTTCAT
GATG TTTTCTTTTTTGCTAAAACTAAAGAATTATTCTTTTACATTTCAG

Fator IX Humano (R338L)-HiBit (1ª Orientação) (SEQ ID NO: 272):

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTA
TAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATG
GAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAA
CAACTGAATTTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCAT
GTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCC
CTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAAT
GGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCC
TGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTG
CCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTG
AGACTGTTTTTCCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTT
GGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGT
GGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAA
GTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTG
CCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATAT
TGAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCA
CCACAACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAA

CTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACA
AGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGG
GAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCC
ACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAAC
ATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGT
GGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATT
AGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTC
TCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTGTCAGCGGATGG
AGACTGTT CAAGAAGATCAGCTAA

Fator IX Humano (R338L)-HiBit (2ª Orientação) (SEQ ID NO: 273):

TTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTC
TTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCC
CCTTCATCGCACATTCCTCCCCCCAACTTATTATCCCGGTCAAGAAACTTGTTCC
TTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACC
GCCCTCGTGAAACCCAGCACAAAACATGTTATTGTAAATCGTAAATTTCGTGGAC
AGAAGACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAG
GGCTGATCGACCTTTGTGGAAGACCCGCCCCCACCCACTCACATATCCGCTCC
CAAATTTCAAGAAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACA
TAGGAGTTAAGTACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTG
TACTTGTTTATAGCGGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCC
TTTTCTGTTCAGTATGCTCAGTTTCTTCAATGTTGTGTTCGCCAGCCACGACCGT
AATCTTAACCCCCGTCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTG
ACTATGGAGCCCCCACAAAACGCGTCGACTTTTCCGTTGAGCACCACCTGCCAT
GGAAATTGGCCAGGTTTAGCGTCCTCGCCCCCGACAACCCTAGTAAAGTCATTA
AATGACTGTGTGGATTGTGTTATATTATCAAGAATCGTTTCGGCTTCAGTAGAGT
TAACGTAGTCCACATCGGGAAAAACTGTCTCGGCCCTTGTCAACTTTGATGTCT
GGGACACACTTACCCGACCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTT

TGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACACAACTTTGTTGTCG
GCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGTGACGT
CCAACTCGCAGTTTTTTCCTTCAAAACCAAAAGGGCACCAACACTCGTAGGAATT
TATATCGTCTTTACAACTCCCCCCATTCAGACATGGATTAGATTCGCATTGGTCC
CCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCT
CGCGCGCTTCTTCAAAACTGCATTTTTCCTCCATACACTCTCGCTCCAAGTTCCC
TTGCACGAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATC
TTATTCGCGTTTTCGTGGTCC AGAAA

Aceitante de Splice de Fator IX Humano (2ª Orientação) (SEQ ID NO: 274):

CTGAAATGTAAAAGAATAATTCTTTAGTTTTAGCAAAAAAGAAAACATCAT
GAA AATTTTACATCTCTTAAGAAAGTCTTTGTTTTTAATCCAAATAATC

Nluc-P2A-GFP (1ª Orientação) (SEQ ID NO: 275):

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTA
TAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATG
GAAGAAAAGTGTAGTTTTGAAGAAGCAGTATTCACTTTGGAGGACTTTGTCGGT
GACTGGAGGCAAACCGCTGGTTATAATCTCGACCAAGTACTGGAACAGGGCGG
GGTAAGTTCCCTCTTTCAGAATTTGGGTGTAAGCGTCACACCAATCCAGCGGAT
TGTGTTGTCTGGAGAGAACGGACTCAAAATTGACATCCATGTTATCATTCCATAT
GAAGGTCTCAGTGGAGACCAAATGGGGCAGATCGAGAAGATTTTCAAGGTAGTT
TACCCAGTCGACGATCACCACTTCAAAGTCATTCTCCACTATGGCACACTTGTTA
TCGACGGAGTAACTCCTAATATGATTGATTACTTTGGTCGCCCGTATGAGGGCA
TCGCAGTGTTTGATGGCAAAAAGATCACCGTAACAGGAACGTTGTGGAATGGGA
ACAAGATAATCGACGAGAGATTGATAAATCCAGACGGGTCACTCCTGTTCAGGG
TTACAATTAACGGCGTCACAGGATGGAGACTCTGTGAACGAATACTGGCCACAA
ATTTTTCACTCCTGAAGCAGGCCGGAGACGTGGAGGAAAACCCAGGGCCCGTG
AGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG
ACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGA

TGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGC
CCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTC
AGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
CGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACA
AGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGA
GCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACG
GCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG
CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT
GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT
CTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGCCCGAAGAAGAAGAGAA
AGGTCTAA

Nluc-P2A-GFP (2ª Orientação) (SEQ ID NO: 276):

TTACACCTTCCTCTTCTTCTTGGGGCTGCCGCCGCCCTTGTACAGCTCG
TCCATGCCCAGGGTGATGCCGGCGGCGGTCACGAACTCCAGCAGCACCATGTG
GTCCCTCTTCTCGTTGGGGTCCTTGCTCAGGGCGCTCTGGGTGCTCAGGTAGT
GGTTGTCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTA
GTGGTCGGCCAGCTGCACGCTGCCGTCCTCGATGTTGTGCCTGATCTTGAAGT
TCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATGTACACGTTGTGGCTGT
TGTAGTTGTACTCCAGCTTGTGGCCCAGGATGTTGCCGTCCTCCTTGAAGTCGA
TGCCCTTCAGCTCGATCCTGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGG
CCCTGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTCCTCTCCTGCACGT
AGCCCTCGGGCATGGCGCTCTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGG
TACCTGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACCAGGGTGGGCCA
GGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGT
AGGTGGCGTCGCCCTCGCCCTCGCCGCTCACGCTGAACTTGTGGCCGTTCACG

TCGCCGTCCAGCTCCACCAGGATGGGCACCACGCCGGTGAACAGCTCCTCGCC
CTTGCTCACGGGGCCGGGGTTCTCCTCCACGTCGCCGGCCTGCTTCAGCAGGC
TGAAGTTGGTGGCCAGGATCCTCTCGCACAGCCTCCAGCCGGTCACGCCGTTG
ATGGTCACCCTGAACAGCAGGCTGCCGTCGGGGTTGATCAGCCTCTCGTCGAT
GATCTTGTTGCCGTTCCACAGGGTGCCGGTCACGGTGATCTTCTTGCCGTCGAA
CACGGCGATGCCCTCGTAGGGCCTGCCGAAGTAGTCGATCATGTTGGGGGTCA
CGCCGTCGATCACCAGGGTGCCGTAGTGCAGGATCACCTTGAAGTGGTGGTCG
TCCACGGGGTACACCACCTTGAAAATCTTCTCGATCTGGCCCATCTGGTCGCCG
CTCAGGCCCTCGTAGGGGATGATCACGTGGATGTCGATCTTCAGGCCGTTCTC
GCCGCTCAGCACGATCCTCTGGATGGGGGTCACGCTCACGCCCAGGTTCTGGA
ACAGGCTGCTCACGCCGCCCTGCTCCAGCACCTGGTCCAGGTTGTAGCCGGCG
GTCTGCCTCCAGTCGCCCACGAAGTCCTCCAGGGTGAACACGGCCTCCTCGAA
GCTGCACTTCTCCTCCATGCACTCCCTCTCCAGGTTGCCCTGCACGAACTCCTC
CAGCTTGCCGCTGTTGTACCTCTTGGGCCTGTTCAGGATCTTGTTGGCGTTCTC
GTGGTCCAGGAA

Sequência completa de P00147 (de ITR para ITR): (SEQ ID NO: 277)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTCT
TAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATC
AATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGAT
CATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAAT
TGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTA
GTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAATTTTG
GAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAA
GGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAG

CAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGAT
ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTG
GAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCC
TGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACT
CAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCC
AAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCT
GTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGTGTTG
AAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGA
ACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTACAAT
GCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGACGAACCC
TTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGA
ACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCC
ACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCG
AGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCT
GGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCA
TGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGA
AGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGT
CAACTGGATTAAGGAAAAAACAAAGCTCACTTAACCTCGACTGTGCCTTCTAGTT
GCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTG
CCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAG
TAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG
GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTC
TGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCAAAAAACCTCCC
ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTG
TTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAA
TAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTAT
CTTATCATGTCTGTTAGGTGAGCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGA

GAGACCTTCGTATAGATGCCATATTTCCCCTTCATCGCACATTCCTCCCCCCAAC
TTATTATCCCGGTCAAGAAACTTGTTCCTTCGACTTCAGTGACGTGTGGTCCACC
TGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAACCCAGCACAAAACAT
GTTATTGTAAATCGTAAATTTCGTGGACAGAAGACAGGTCGCTCTATCGACCAAC
GGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTGTGGAAGACCCG
CCCCCACCCACTCACATATCCGCTCCCAAATTTCAAGAAGATATTTGTATATTCT
TTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGTACGAGTGGCTCGTCC
AGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGCGGCATTATAATTGT
GATGGGGTATGATCCTGATAACATTCCTTTTCTGTTCAGTATGCTCAGTTTCTTC
AATGTTGTGTTCGCCAGCCACGACCGTAATCTTAACCCCCGTCTCGACACAGTG
TGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCCACAAAACGCGTC
GACTTTTCCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCCTC
GCCCCCGACAACCCTAGTAAAGTCATTAAATGACTGTGTGGATTGTGTTATATTA
TCAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACT
GTCTCGGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGG
GAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCT
CAGTGCAACTACACACAACTTTGTTGTCGGCGGAATTTTTACAGAATTGCTCGCA
TCGTCCATTTTTAATGTTGCAGGTGACGTCCAACTCGCAGTTTTTTCCTTCAAAA
CCAAAAGGGCACCAACACTCGTAGGAATTTATATCGTCTTTACAACTCCCCCCAT
TCAGACATGGATTAGATTCGCATTGGTCCCCATCGACATATTGCTTCCAGAACTC
AGTGGTCCGTTCTGTATTCTCAAACACCTCGCGCGCTTCTTCAAAACTGCATTTT
TCCTCCATACACTCTCGCTCCAAGTTCCCTTGCACGAATTCTTCAAGCTTTCCTG
AGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCGTTTTCGTGGTCCAGAAA
AACTGTGGAAACAGGGAGAGAAAAACCACACAACATATTTAAAGATTGATGAAG
ACAACTAACTGTAATATGCTGCTTTTTGTTCTTCTCTTCACTGACCTAAGAGATCT
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC
ACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGG

CCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Sequência completa de P00411 (de ITR para ITR): (SEQ ID NO: 278)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTCT
GATTATTTGGATTAAAAACAAAGACTTTCTTAAGAGATGTAAAATTTTCATGATGT
TTTCTTTTTTGCTAAAACTAAAGAATTATTCTTTTACATTTCAGTTTTTCTTGATCA
TGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTG
GAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGT
TTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAATTTTGGA
AGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCA
GTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG
AAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCA
GTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATAT
CGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGA
AGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTG
ATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCA
AAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAA
ACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGT
GGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGTGTTGAA
ACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGAA
CATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTACAATG
CAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGACGAACCCTT
AGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAAC
ATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCAC
AAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAG
CCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGG

CTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATG
TTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAG
AGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGTCA
ACTGGATTAAGGAAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAAGA
AGATCAGCTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTA
ATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGG
GGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGC
ATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCAAAAAACCTCCCACACCTCCCCCTGAACCTGAAAC
ATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTAC
AAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTC
TAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTTAGGAAATCT
TCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTCTTTTCTTTTATCCAATT
CACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCCTTCATCGCACATTC
CTCCCCCCAACTTATTATCCCGGTCAAGAAACTTGTTCCTTCGACTTCAGTGACG
TGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAACCC
AGCACAAAACATGTTATTGTAAATCGTAAATTTCGTGGACAGAAGACAGGTCGCT
CTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTT
GTGGAAGACCCGCCCCCACCCACTCACATATCCGCTCCCAAATTTCAAGAAGAT
ATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGTACG
AGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGCG
GCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTCAGTAT
GCTCAGTTTCTTCAATGTTGTGTTCGCCAGCCACGACCGTAATCTTAACCCCCG
TCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCC
CACAAAACGCGTCGACTTTTCCGTTGAGCACCACCTGCCATGGAAATTGGCCAG
GTTTAGCGTCCTCGCCCCCGACAACCCTAGTAAAGTCATTAAATGACTGTGTGG

ATTGTGTTATATTATCAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCAC
ATCGGGAAAAACTGTCTCGGCCCTTGTCAACTTTGATGTCTGGGACACACTTAC
CCGACCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGA
GCCGGTAGCCCTCAGTGCAACTACACACAACTTTGTTGTCGGCGGAATTTTTAC
AGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGTGACGTCCAACTCGCAGTT
TTTTCCTTCAAAACCAAAAGGGCACCAACACTCGTAGGAATTTATATCGTCTTTA
CAACTCCCCCCATTCAGACATGGATTAGATTCGCATTGGTCCCCATCGACATATT
GCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGCGCGCTTCTT
CAAAACTGCATTTTTCCTCCATACACTCTCGCTCCAAGTTCCCTTGCACGAATTC
TTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCGTTTT
CGTGGTCCAGAAAAACTGAAATGTAAAAGAATAATTCTTTAGTTTTAGCAAAAAA
GAAAACATCATGAAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTTAATCCAAAT
AATCAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCG
CTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTT
TGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGG GAGTGGCCAA

Sequência completa de P00415 (de ITR para ITR): (SEQ ID NO: 279)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTCT
TAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATC
AATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGAT
CATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAAT
TGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTA
GTTTTGAAGAAGCAGTATTCACTTTGGAGGACTTTGTCGGTGACTGGAGGCAAA
CCGCTGGTTATAATCTCGACCAAGTACTGGAACAGGGCGGGGTAAGTTCCCTCT
TTCAGAATTTGGGTGTAAGCGTCACACCAATCCAGCGGATTGTGTTGTCTGGAG
AGAACGGACTCAAAATTGACATCCATGTTATCATTCCATATGAAGGTCTCAGTGG

AGACCAAATGGGGCAGATCGAGAAGATTTTCAAGGTAGTTTACCCAGTCGACGA
TCACCACTTCAAAGTCATTCTCCACTATGGCACACTTGTTATCGACGGAGTAACT
CCTAATATGATTGATTACTTTGGTCGCCCGTATGAGGGCATCGCAGTGTTTGAT
GGCAAAAAGATCACCGTAACAGGAACGTTGTGGAATGGGAACAAGATAATCGAC
GAGAGATTGATAAATCCAGACGGGTCACTCCTGTTCAGGGTTACAATTAACGGC
GTCACAGGATGGAGACTCTGTGAACGAATACTGGCCACAAATTTTTCACTCCTG
AAGCAGGCCGGAGACGTGGAGGAAAACCCAGGGCCCGTGAGCAAGGGCGAGG
AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAAC
GGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA
AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC
ACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGA
CCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC
AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG
GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACA
GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACT
TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTAC
CAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACA
TGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG
CTGTACAAGGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGTCTAACCTCGACT
GTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTG
ACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCA
TCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGA
CAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTG
GGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCC
CCAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG

TTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATC
ACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA
ACTCATCAATGTATCTTATCATGTCTGTTACACCTTCCTCTTCTTCTTGGGGCTG
CCGCCGCCCTTGTACAGCTCGTCCATGCCCAGGGTGATGCCGGCGGCGGTCA
CGAACTCCAGCAGCACCATGTGGTCCCTCTTCTCGTTGGGGTCCTTGCTCAGG
GCGCTCTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGTCGC
CGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGCTGCACGCTGCCGTCCTCG
ATGTTGTGCCTGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCA
TGATGTACACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGGCCCAGGATGT
TGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATCCTGTTCACCAGGGTGT
CGCCCTCGAACTTCACCTCGGCCCTGGTCTTGTAGTTGCCGTCGTCCTTGAAGA
AGATGGTCCTCTCCTGCACGTAGCCCTCGGGCATGGCGCTCTTGAAGAAGTCG
TGCTGCTTCATGTGGTCGGGGTACCTGCTGAAGCACTGCACGCCGTAGGTCAG
GGTGGTCACCAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATG
AACTTCAGGGTCAGCTTGCCGTAGGTGGCGTCGCCCTCGCCCTCGCCGCTCAC
GCTGAACTTGTGGCCGTTCACGTCGCCGTCCAGCTCCACCAGGATGGGCACCA
CGCCGGTGAACAGCTCCTCGCCCTTGCTCACGGGGCCGGGGTTCTCCTCCACG
TCGCCGGCCTGCTTCAGCAGGCTGAAGTTGGTGGCCAGGATCCTCTCGCACAG
CCTCCAGCCGGTCACGCCGTTGATGGTCACCCTGAACAGCAGGCTGCCGTCGG
GGTTGATCAGCCTCTCGTCGATGATCTTGTTGCCGTTCCACAGGGTGCCGGTCA
CGGTGATCTTCTTGCCGTCGAACACGGCGATGCCCTCGTAGGGCCTGCCGAAG
TAGTCGATCATGTTGGGGGTCACGCCGTCGATCACCAGGGTGCCGTAGTGCAG
GATCACCTTGAAGTGGTGGTCGTCCACGGGGTACACCACCTTGAAAATCTTCTC
GATCTGGCCCATCTGGTCGCCGCTCAGGCCCTCGTAGGGGATGATCACGTGGA
TGTCGATCTTCAGGCCGTTCTCGCCGCTCAGCACGATCCTCTGGATGGGGGTC
ACGCTCACGCCCAGGTTCTGGAACAGGCTGCTCACGCCGCCCTGCTCCAGCAC
CTGGTCCAGGTTGTAGCCGGCGGTCTGCCTCCAGTCGCCCACGAAGTCCTCCA

GGGTGAACACGGCCTCCTCGAAGCTGCACTTCTCCTCCATGCACTCCCTCTCCA
GGTTGCCCTGCACGAACTCCTCCAGCTTGCCGCTGTTGTACCTCTTGGGCCTGT
TCAGGATCTTGTTGGCGTTCTCGTGGTCCAG GAA

Sequência completa de P00418 (de ITR para ITR): (SEQ ID NO: 280)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTCT
TAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATC
AATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGAT
CATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAAT
TGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTA
GTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAATTTTG
GAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAA
GGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAG
CAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGAT
ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTG
GAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCC
TGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACT
CAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCC
AAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCT
GTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGTGTTG
AAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGA
ACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTACAAT
GCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGACGAACCC
TTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGA
ACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCC

ACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCG
AGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCT
GGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCA
TGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGA
AGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGT
CAACTGGATTAAGGAAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAA
GAAGATCAGCTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC
CCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC
TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGG
GGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
GCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCT
GGGGCTCTAGGGGGTATCCCCAAAAAACCTCCCACACCTCCCCCTGAACCTGA
AACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGT
TACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCA
TTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTTAGGAAA
TCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTCTTTTCTTTTATCCA
ATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCCTTCATCGCACAT
TCCTCCCCCCAACTTATTATCCCGGTCAAGAAACTTGTTCCTTCGACTTCAGTGA
CGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAAC
CCAGCACAAAACATGTTATTGTAAATCGTAAATTTCGTGGACAGAAGACAGGTC
GCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACC
TTTGTGGAAGACCCGCCCCCACCCACTCACATATCCGCTCCCAAATTTCAAGAA
GATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGT
ACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATA
GCGGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTCAG
TATGCTCAGTTTCTTCAATGTTGTGTTCGCCAGCCACGACCGTAATCTTAACCCC
CGTCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCC

CCCACAAAACGCGTCGACTTTTCCGTTGAGCACCACCTGCCATGGAAATTGGCC
AGGTTTAGCGTCCTCGCCCCCGACAACCCTAGTAAAGTCATTAAATGACTGTGT
GGATTGTGTTATATTATCAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCC
ACATCGGGAAAAACTGTCTCGGCCCTTGTCAACTTTGATGTCTGGGACACACTT
ACCCGACCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGC
GAGCCGGTAGCCCTCAGTGCAACTACACACAACTTTGTTGTCGGCGGAATTTTT
ACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGTGACGTCCAACTCGCA
GTTTTTTCCTTCAAAACCAAAAGGGCACCAACACTCGTAGGAATTTATATCGTCT
TTACAACTCCCCCCATTCAGACATGGATTAGATTCGCATTGGTCCCCATCGACAT
ATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGCGCGCTT
CTTCAAAACTGCATTTTTCCTCCATACACTCTCGCTCCAAGTTCCCTTGCACGAA
TTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCGT
TTTCGTGGTCCAGAAAAACTGTGGAAACAGGGAGAGAAAAACCACACAACATAT
TTAAAGATTGATGAAGACAACTAACTGTAATATGCTGCTTTTTGTTCTTCTCTTCA
CTGACCTAAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTG
CGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGA
CCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGC
CAA

Sequência completa de P00123 (de ITR para ITR): (SEQ ID NO: 281)

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGAC
CAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCG
AGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGGAGGGGT
GGAGTCGTGATAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAG
TTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAG
TTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAA
TTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGA
AGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACA

ACTGAATTTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTT
TAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTT
TGGATTTGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGC
AGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTA
CTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCAT
TTCCATGTGGAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGAC
TGTTTTTCCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATA
ACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGA
AGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGAT
GCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACT
GTGTTGAAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGA
GACAGAACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAA
CTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGAC
GAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAAT
ACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAG
TCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGT
TGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTC
TGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGG
ACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTG
GGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCG
GTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAACCTCGACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGG
AAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTG
TCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTAT
GGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACTAGT
CCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC

GCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC
AGAGAGGGA

Sequência completa de P00204 (de ITR para ITR): (SEQ ID NO: 282)

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGAC
CAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCG
AGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGGAGGGGT
GGAGTCGTGACCTAGGTCGTCTCCGGCTCTGCTTTTTCCAGGGGTGTGTTTCGC
CGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTTCTAGTAA
TGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCTAGATT
ATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACTAGGTCA
GTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTT
AAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAA
ACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAG
AGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTG
AAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCA
GTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTG
CAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAG
AACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTT
GTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGAC
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAG
TTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGT
GGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGC
ACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCA
GGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGA
GGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGTGTTGAAACTG
GTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAC
AGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTACAATGCAGCT

ATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGC
TAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTT
CCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGG
GAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACA
TGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCC
ATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACT
GAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGT
GCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTCAACTGG
ATTAAGGAAAAAACAAAGCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCC
ATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCC
CACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
CATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGG
AAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCG
GAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTATTG
ATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTA
AATTTCTTTTAATTAAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAA
AAATTAAAAGTGTACTAGTCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAG
GCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAG
TGAGCGAGCGAGCGCGCAGAGAGGGA

Sequência completa de P00353 (de ITR para ITR): (SEQ ID NO: 283)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTGA
TTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAAAAAAAAATGCTAGGCAGAATG
ACTCAAATTACGTTGGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTG
CTCGACCATGCTATACTAAAAATTAAAAGTGTGTGTTACTAATTTTATAAATGGAG

TTTCCATTTATATTTACCTTTATTTCTTATTTACCATTGTCTTAGTAGATATTTACAA
ACATGACAGAAACACTAAATCTTGAGTTTGAATGCACAGATATAAACACTTAACG
GGTTTTAAAAATAATAATGTTGGTGAAAAAATATAACTTTGAGTGTAGCAGAGAG
GAACCATTGCCACCTTCAGATTTTCCTGTAACGATCGGGAACTGGCATCTTCAG
GGAGTAGCTTAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGT
TGTCTTCATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGT
TTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTTTCTT
GATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGT
AAATTGGAAGAGTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACTGAAT
TTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGG
CGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTT
GAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGC
GAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAG
GGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCA
TGTGGAAGAGTTTCTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTT
TTCCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACAT
CACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGA
TGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGC
ATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACTGCTGCCCACTGT
GTTGAAACTGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGA
CAGAACATACAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACAACTA
CAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGGACGAA
CCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACA
CGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCT
TCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGA
CCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGT

GCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACC
CCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGG
TGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTA
TGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAACCTCGACTGTGCCTTCTA
GTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAG
GTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCT
GAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGG
GAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGC
TTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCGTGAGATCG
CCCATCGGTATAATGATTTGGGAGAACAACATTTCAAAGGCCTGTAAGTTATAAT
GCTGAAAGCCCACTTAATATTTCTGGTAGTATTAGTTAAAGTTTTAAAACACCTTT
TTCCACCTTGAGTGTGAGAATTGTAGAGCAGTGCTGTCCAGTAGAAATGTGTGC
ATTGACAGAAAGACTGTGGATCTGTGCTGAGCAATGTGGCAGCCAGAGATCACA
AGGCTATCAAGCACTTTGCACATGGCAAGTGTAACTGAGAAGCACACATTCAAA
TAATAGTTAATTTTAATTGAATGTATCTAGCCATGTGTGGCTAGTAGCTCCTTTCC
TGGAGAGAGAATCTGGAGCCCACATCTAACTTGTTAAGTCTGGAATCTTATTTTT
TATTTCTGGAAAGGTCTATGAACTATAGTTTTGGGGGCAGCTCACTTACTAACTT
TTAATGCAATAAGAATCTCATGGTATCTTGAGAACATTATTTTGTCTCTTTGTAGA
TCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCG
CTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCC
CGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Sequência completa de P00354 (de ITR para ITR): (SEQ ID NO: 284)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTTA
GCCTCTGGCAAAATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGC
TCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTT

TTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAA
ATTTTCCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTATTATTGT
CATTATTTGCATCTGAGAACCCTTAGGTGGTTATATTATTGATATATTTTTGGTAT
CTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTA
AAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAATTTATT
ACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAAATTAAAAGTGT
GTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTATTTCTTATTTA
CCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTTGAGTTTGA
ATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTGAAAAAAT
ATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCTGTAAC
GATCGGGAACTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAACA
AAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGT
TTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCT
GAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAA
CCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGT
TTTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCAGTATGTTGATGGAGAT
CAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAAT
TCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATG
TAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGA
TAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAA
GTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACT
TCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTA
CTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGA
CTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCA
GGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAA
AAATGGATTGTAACTGCTGCCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTG
TCGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATG

TGATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACA
CCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTG
GCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTC
TTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAA
GTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGA
TTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCA
GTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAA
GCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCT
CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGG
GTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCGTGAGATCGCCCATCGGTATAATGATTTGGGAGA
ACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTTCTG
GTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGTA
GAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGT
GCTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATG
GCAAGTGTAACTGAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTAT
CTAGCCATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCCCACA
TCTAACTTGTTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTA
TAGTTTTGGGGGCAGCTCACTTACTAACTTTTAATGCAATAAGAATCTCATGGTA
TCTTGAGAACATTATTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTA
AGGGGTCTATACTTAAGTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAA
AAATGAGTAATTAATTTATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAAC
AGAGGTATATGGTTAGAAAAGAAACAATTCAAAGGACTTATATAATATCTAGCCT

TGACAATGAATAAATTTAGAGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAATCT
ATTGACACATATGTGCATCTGCACTTCAGCATGGTAGAAGTCCATATTCAGATCT
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC
ACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGG
CCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

P00350: O construto de F9 de HA de 300/600 pb (para G551) (SEQ ID NO: 285)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTAA
GTATATTAGAGCGAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTA
GCCTCTGGCAAAATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGC
TCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTT
TTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAA
ATTTTCCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTATTATTGT
CATTATTTGCATCTGAGAACCTTTTTCTTGATCATGAAAACGCCAACAAAATTCTG
AATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAAC
CTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTT
TTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCAGTATGTTGATGGAGATC
AGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATT
CCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGT
AACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGAT
AACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAG
TCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTT
CTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTAC
TGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGAC
TTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAG

GTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAA
AATGGATTGTAACTGCTGCCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGT
CGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATGT
GATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACA
CCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTG
GCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTC
TTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAA
GTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGA
TTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCA
GTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAA
GCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCT
CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGG
GTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTATTGATATATTTTTGGTATC
TTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAA
AAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAATTTATTA
CGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAAATTAAAAGTGTG
TGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTATTTCTTATTTAC
CATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAAGATCTAGGAACC
CCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGG
CCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGT
GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

P00356: O construto de F9 de HA de 300/2000 pb (para G551) (SEQ ID NO:

286)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTAA
GTATATTAGAGCGAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTA
GCCTCTGGCAAAATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGC
TCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTT
TTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAA
ATTTTCCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTATTATTGT
CATTATTTGCATCTGAGAACCTTTTTCTTGATCATGAAAACGCCAACAAAATTCTG
AATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAAC
CTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTT
TTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCAGTATGTTGATGGAGATC
AGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATT
CCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGT
AACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGAT
AACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAG
TCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTT
CTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTAC
TGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGAC
TTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAG
GTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAA
AATGGATTGTAACTGCTGCCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGT
CGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATGT
GATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACA
CCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTG

GCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTC
TTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAA
GTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGA
TTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCA
GTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAA
GCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCT
CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGG
GTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTATTGATATATTTTTGGTATC
TTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAA
AAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAATTTATTA
CGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAAATTAAAAGTGTG
TGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTATTTCTTATTTAC
CATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTTGAGTTTGAA
TGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTGAAAAAATA
TAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCTGTAACG
ATCGGGAACTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAACAAA
AAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTT
TTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG
GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATAT
TTCTGGTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAA
TTGTAGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGA
TCTGTGCTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCA
CATGGCAAGTGTAACTGAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAAT

GTATCTAGCCATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCC
CACATCTAACTTGTTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGA
ACTATAGTTTTGGGGGCAGCTCACTTACTAACTTTTAATGCAATAAGATCCATGG
TATCTTGAGAACATTATTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAG
TAAGGGGTCTATACTTAAGTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAA
AAAAATGAGTAATTAATTTATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATA
ACAGAGGTATATGGTTAGAAAAGAAACAATTCAAAGGACTTATATAATATCTAGC
CTTGACAATGAATAAATTTAGAGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAAT
CTATTGACACATATGTGCATCTGCACTTCAGCATGGTAGAAGTCCATATTCCTTT
GCTTGGAAAGGCAGGTGTTCCCATTACGCCTCAGAGAATAGCTGACGGGAAGA
GGCTTTCTAGATAGTTGTATGAAAGATATACAAAATCTCGCAGGTATACACAGGC
ATGATTTGCTGGTTGGGAGAGCCACTTGCCTCATACTGAGGTTTTTGTGTCTGC
TTTTCAGAGTCCTGATTGCCTTTTCCCAGTATCTCCAGAAATGCTCATACGATGA
GCATGCCAAATTAGTGCAGGAAGTAACAGACTTTGCAAAGACGTGTGTTGCCGA
TGAGTCTGCCGCCAACTGTGACAAATCCCTTGTGAGTACCTTCTGATTTTGTGG
ATCTACTTTCCTGCTTTCTGGAACTCTGTTTCAAAGCCAATCATGACTCCATCAC
TTAAGGCCCCGGGAACACTGTGGCAGAGGGCAGCAGAGAGATTGATAAAGCCA
GGGTGATGGGAATTTTCTGTGGGACTCCATTTCATAGTAATTGCAGAAGCTACA
ATACACTCAAAAAGTCTCACCACATGACTGCCCAAATGGGAGCTTGACAGTGAC
AGTGACAGTAGATATGCCAAAGTGGATGAGGGAAAGACCACAAGAGCTAAACC
CTGTAAAAAGAACTGTAGGCAACTAAGGAATGCAGAGAGAAAGATCTAGGAACC
CCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGG
CCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGT
GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

P00362: O construto de F9 de HA de 300/1500 pb (para G551) (SEQ ID NO: 287)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA

CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC
GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTAA
GTATATTAGAGCGAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTA
GCCTCTGGCAAAATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGC
TCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTT
TTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAA
ATTTTCCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTATTATTGT
CATTATTTGCATCTGAGAACCTTTTTCTTGATCATGAAAACGCCAACAAAATTCTG
AATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAAC
CTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTT
TTTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCAGTATGTTGATGGAGATC
AGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATT
CCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGT
AACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGAT
AACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAG
TCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTT
CTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTAC
TGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGAC
TTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAG
GTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAA
AATGGATTGTAACTGCTGCCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGT
CGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATGT
GATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACA
CCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTG
GCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTC
TTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAA

GTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGA
TTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCA
GTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAA
GCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCT
CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGG
GTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA
TGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTATTGATATATTTTTGGTATC
TTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAA
AAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAATTTATTA
CGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAAATTAAAAGTGTG
TGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTATTTCTTATTTAC
CATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTTGAGTTTGAA
TGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTGAAAAAATA
TAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCTGTAACG
ATCGGGAACTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAACAAA
AAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTT
TTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG
GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATAT
TTCTGGTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAA
TTGTAGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGA
TCTGTGCTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCA
CATGGCAAGTGTAACTGAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAAT
GTATCTAGCCATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCC
CACATCTAACTTGTTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGA

ACTATAGTTTTGGGGGCAGCTCACTTACTAACTTTTAATGCAATAAGATCCATGG
TATCTTGAGAACATTATTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAG
TAAGGGGTCTATACTTAAGTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAA
AAAAATGAGTAATTAATTTATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATA
ACAGAGGTATATGGTTAGAAAAGAAACAATTCAAAGGACTTATATAATATCTAGC
CTTGACAATGAATAAATTTAGAGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAAT
CTATTGACACATATGTGCATCTGCACTTCAGCATGGTAGAAGTCCATATTCCTTT
GCTTGGAAAGGCAGGTGTTCCCATTACGCCTCAGAGAATAGCTGACGGGAAGA
GGCTTTCTAGATAGTTGTATGAAAGATATACAAAATCTCGCAGGTATACACAGGC
ATGATTTGCTGGTTGGGAGAGCCACTTAGATCTAGGAACCCCTAGTGATGGAGT
TGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAG
CCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGC
GCAGAGAGGGAGTGGCCAA

ORF de Cas9 (SEQ ID NO: 703)

ATGGATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGG
GTTGGGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCC
TGGGGAACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGT
TTGACTCCGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGCG
ACGCTACACCCGGCGGAAGAATCGCATCTGCTATCTGCAAGAGATCTTTTCGAA
CGAAATGGCAAAGGTCGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCT
GGTGGAGGAGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTCG
ACGAAGTGGCGTACCACGAAAAGTACCCGACCATCTACCATCTGCGGAAGAAG
TTGGTTGACTCAACTGACAAGGCCGACCTCAGATTGATCTACTTGGCCCTCGCC
CATATGATCAAATTCCGCGGACACTTCCTGATCGAAGGCGATCTGAACCCTGAT
AACTCCGACGTGGATAAGCTTTTCATTCAACTGGTGCAGACCTACAACCAACTG
TTCGAAGAAAACCCAATCAATGCTAGCGGCGTCGATGCCAAGGCCATCCTGTCC
GCCCGGCTGTCGAAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGG

GAGAGAAAAAGAACGGACTTTTCGGCAACTTGATCGCTCTCTCACTGGGACTCA
CTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGACGCGAAGCTGCAACTCT
CAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGCACAAATTGGCGATC
AGTACGCGGATCTGTTCCTTGCCGCTAAGAACCTTTCGGACGCAATCTTGCTGT
CCGATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAGCGCCTCG
ATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTC
GTGAGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTCCAAG
AATGGGTACGCAGGGTACATCGATGGAGGCGCTAGCCAGGAAGAGTTCTATAA
GTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGAAGAACTGCTGGTCAA
GCTGAACAGGGAGGATCTGCTCCGGAAACAGAGAACCTTTGACAACGGATCCA
TTCCCCACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAG
GACTTTTACCCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACG
TTCCGCATCCCGTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGC
GTGGATGACTAGAAAATCAGAGGAAACCATCACTCCTTGGAATTTCGAGGAAGT
TGTGGATAAGGGAGCTTCGGCACAAAGCTTCATCGAACGAATGACCAACTTCGA
CAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTACGAATA
CTTCACTGTCTACAACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAG
GAAGCCGGCCTTTCTGTCCGGAGAACAGAAGAAAGCAATTGTCGATCTGCTGTT
CAAGACCAACCGCAAGGTGACCGTCAAGCAGCTTAAAGAGGACTACTTCAAGAA
GATCGAGTGTTTCGACTCAGTGGAAATCAGCGGGGTGGAGGACAGATTCAACG
CTTCGCTGGGAACCTATCATGATCTCCTGAAGATCATCAAGGACAAGGACTTCC
TTGACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGACCC
TTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCT
TCGACGATAAGGTCATGAAACAACTCAAGCGCCGCCGGTACACTGGTTGGGGC
CGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGATAAACAGAGCGGTAAAACT
ATCCTGGATTTCCTCAAATCGGATGGCTTCGCTAATCGTAACTTCATGCAATTGA
TCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAAAAAGCACAAGTGTCCG

GACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTCGCCGGCG
ATTAAGAAGGGAATTCTGCAAACTGTGAAGGTGGTCGACGAGCTGGTGAAGGT
CATGGGACGGCACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACC
AGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATGAAGCGGATCGAA
GAAGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAA
CACGCAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGA
CATGTACGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGA
CCACATCGTTCCACAGTCCTTTCTGAAGGATGACTCGATCGATAACAAGGTGTT
GACTCGCAGCGACAAGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGG
TCGTGAAGAAGATGAAGAATTACTGGCGGCAGCTCCTGAATGCGAAGCTGATTA
CCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAG
CTGGATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAGATTAC
CAAGCACGTGGCGCAGATCTTGGACTCCCGCATGAACACTAAATACGACGAGA
ACGATAAGCTCATCCGGGAAGTGAAGGTGATTACCCTGAAAAGCAAACTTGTGT
CGGACTTTCGGAAGGACTTTCAGTTTTACAAAGTGAGAGAAATCAACAACTACCA
TCACGCGCATGACGCATACCTCAACGCTGTGGTCGGTACCGCCCTGATCAAAAA
GTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTACAAGGTCTACGACGT
GAGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGAAAT
ACTTCTTTTACTCAAACATCATGAACTTTTTCAAGACTGAAATTACGCTGGCCAAT
GGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGGAGAAACGGGCGAAAT
CGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTCGCAAAGTGCTCTCTATGC
CGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGATTTTCAAAGG
AATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAAGAAAGACT
GGGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTC
CTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAAGCTCAAATCCGTCAA
AGAGCTGCTGGGGATTACCATCATGGAACGATCCTCGTTCGAGAAGAACCCGAT
TGATTTCCTCGAGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAA

ACTCCCCAAGTACTCACTGTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGC
TTCGGCCGGAGAACTCCAAAAAGGAAATGAGCTGGCCTTGCCTAGCAAGTACG
TCAACTTCCTCTATCTTGCTTCGCACTACGAAAAACTCAAAGGGTCACCGGAAG
ATAACGAACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAA
TCATCGAACAAATCTCCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACC
TCGACAAAGTCCTGTCGGCCTACAATAAGCATAGAGATAAGCCGATCAGAGAAC
AGGCCGAGAACATTATCCACTTGTTCACCCTGACTAACCTGGGAGCCCCAGCC
GCCTTCAAGTACTTCGATACTACTATCGATCGCAAAAGATACACGTCCACCAAG
GAAGTTCTGGACGCGACCCTGATCCACCAAAGCATCACTGGACTCTACGAAACT
AGGATCGATCTGTCGCAGCTGGGTGGCGAT

ORF de U-dep Cas9 (SEQ ID NO: 704)

ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTC
GGATGGGCAGTCATCACAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGT
CCTGGGAAACACAGACAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGC
TGTTCGACAGCGGAGAAACAGCAGAAGCAACAAGACTGAAGAGAACAGCAAGA
AGAAGATACACAAGAAGAAAGAACAGAATCTGCTACCTGCAGGAAATCTTCAGC
AACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACAGACTGGAAGAAAGCTTC
CTGGTCGAAGAAGACAAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGT
CGACGAAGTCGCATACCACGAAAAGTACCCGACAATCTACCACCTGAGAAAGAA
GCTGGTCGACAGCACAGACAAGGCAGACCTGAGACTGATCTACCTGGCACTGG
CACACATGATCAAGTTCAGAGGACACTTCCTGATCGAAGGAGACCTGAACCCGG
ACAACAGCGACGTCGACAAGCTGTTCATCCAGCTGGTCCAGACATACAACCAGC
TGTTCGAAGAAAACCCGATCAACGCAAGCGGAGTCGACGCAAAGGCAATCCTG
AGCGCAAGACTGAGCAAGAGCAGAAGACTGGAAAACCTGATCGCACAGCTGCC
GGGAGAAAAGAAGAACGGACTGTTCGGAAACCTGATCGCACTGAGCCTGGGAC
TGACACCGAACTTCAAGAGCAACTTCGACCTGGCAGAAGACGCAAAGCTGCAG
CTGAGCAAGGACACATACGACGACGACCTGGACAACCTGCTGGCACAGATCGG

AGACCAGTACGCAGACCTGTTCCTGGCAGCAAAGAACCTGAGCGACGCAATCC
TGCTGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGGCACCGCTGAGC
GCAAGCATGATCAAGAGATACGACGAACACCACCAGGACCTGACACTGCTGAA
GGCACTGGTCAGACAGCAGCTGCCGGAAAAGTACAAGGAAATCTTCTTCGACC
AGAGCAAGAACGGATACGCAGGATACATCGACGGAGGAGCAAGCCAGGAAGAA
TTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAAGAACTG
CTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTCGACAA
CGGAAGCATCCCGCACCAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAA
GACAGGAAGACTTCTACCCGTTCCTGAAGGACAACAGAGAAAAGATCGAAAAGA
TCCTGACATTCAGAATCCCGTACTACGTCGGACCGCTGGCAAGAGGAAACAGC
AGATTCGCATGGATGACAAGAAAGAGCGAAGAAACAATCACACCGTGGAACTTC
GAAGAAGTCGTCGACAAGGGAGCAAGCGCACAGAGCTTCATCGAAAGAATGAC
AAACTTCGACAAGAACCTGCCGAACGAAAAGGTCCTGCCGAAGCACAGCCTGC
TGTACGAATACTTCACAGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAG
AAGGAATGAGAAAGCCGGCATTCCTGAGCGGAGAACAGAAGAAGGCAATCGTC
GACCTGCTGTTCAAGACAAACAGAAAGGTCACAGTCAAGCAGCTGAAGGAAGA
CTACTTCAAGAAGATCGAATGCTTCGACAGCGTCGAAATCAGCGGAGTCGAAGA
CAGATTCAACGCAAGCCTGGGAACATACCACGACCTGCTGAAGATCATCAAGGA
CAAGGACTTCCTGGACAACGAAGAAAACGAAGACATCCTGGAAGACATCGTCCT
GACACTGACACTGTTCGAAGACAGAGAAATGATCGAAGAAAGACTGAAGACATA
CGCACACCTGTTCGACGACAAGGTCATGAAGCAGCTGAAGAGAAGAAGATACA
CAGGATGGGGAAGACTGAGCAGAAAGCTGATCAACGGAATCAGAGACAAGCAG
AGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGACGGATTCGCAAACAGAAA
CTTCATGCAGCTGATCCACGACGACAGCCTGACATTCAAGGAAGACATCCAGAA
GGCACAGGTCAGCGGACAGGGAGACAGCCTGCACGAACACATCGCAAACCTG
GCAGGAAGCCCGGCAATCAAGAAGGGAATCCTGCAGACAGTCAAGGTCGTCGA
CGAACTGGTCAAGGTCATGGGAAGACACAAGCCGGAAAACATCGTCATCGAAAT

GGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAA
TGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGCCAGATCCTGAAGGAA
CACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTACCTG
CAGAACGGAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAG
CGACTACGACGTCGACCACATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCA
TCGACAACAAGGTCCTGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAAC
GTCCCGAGCGAAGAAGTCGTCAAGAAGATGAAGAACTACTGGAGACAGCTGCT
GAACGCAAAGCTGATCACACAGAGAAAGTTCGACAACCTGACAAAGGCAGAGA
GAGGAGGACTGAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTC
GAAACAAGACAGATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAA
CACAAAGTACGACGAAAACGACAAGCTGATCAGAGAAGTCAAGGTCATCACACT
GAAGAGCAAGCTGGTCAGCGACTTCAGAAAGGACTTCCAGTTCTACAAGGTCAG
AGAAATCAACAACTACCACCACGCACACGACGCATACCTGAACGCAGTCGTCG
GAACAGCACTGATCAAGAAGTACCCGAAGCTGGAAAGCGAATTCGTCTACGGA
GACTACAAGGTCTACGACGTCAGAAAGATGATCGCAAAGAGCGAACAGGAAATC
GGAAAGGCAACAGCAAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGA
CAGAAATCACACTGGCAAACGGAGAAATCAGAAAGAGACCGCTGATCGAAACAA
ACGGAGAAACAGGAGAAATCGTCTGGGACAAGGGAAGAGACTTCGCAACAGTC
AGAAAGGTCCTGAGCATGCCGCAGGTCAACATCGTCAAGAAGACAGAAGTCCA
GACAGGAGGATTCAGCAAGGAAAGCATCCTGCCGAAGAGAAACAGCGACAAGC
TGATCGCAAGAAAGAAGGACTGGGACCCGAAGAAGTACGGAGGATTCGACAGC
CCGACAGTCGCATACAGCGTCCTGGTCGTCGCAAAGGTCGAAAAGGGAAAGAG
CAAGAAGCTGAAGAGCGTCAAGGAACTGCTGGGAATCACAATCATGGAAAGAA
GCAGCTTCGAAAAGAACCCGATCGACTTCCTGGAAGCAAAGGGATACAAGGAA
GTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCTGTTCGAACTGGAA
AACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGAAACGA
ACTGGCACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTACGA

AAAGCTGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAAC
AGCACAAGCACTACCTGGACGAAATCATCGAACAGATCAGCGAATTCAGCAAGA
GAGTCATCCTGGCAGACGCAAACCTGGACAAGGTCCTGAGCGCATACAACAAG
CACAGAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCCACCTGTTCACA
CTGACAAACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGAC
AGAAAGAGATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCA
GAGCATCACAGGACTGTACGAAACAAGAATCGACCTGAGCCAGCTGGGAGGAG
ACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGTCTAG

mRNA que compreende U dep Cas9 (SEQ ID NO: 705)

GGGUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUG
UCGUUGCAGGCCUUAUUCGGAUCCGCCACCAUGGACAAGAAGUACAGCAUCG
GACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUA
CAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGC
AUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAG
AAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAAC
AGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACG
ACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAA
GCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCAC
GAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGA
CAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUC
AGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCG
ACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAA
CCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUG
AGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGA
AGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAA
CUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAG
GACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGU

ACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAG
CGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGC
AUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCAC
UGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAG
CAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUC
UACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGC
UGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAA
CGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGA
AGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAA
GAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAAC
AGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAA
CUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGA
AUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACA
GCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUA
CGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAG
GCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGC
UGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAG
CGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUG
AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCU
GGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAA
GAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGC
UGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAA
CGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGC
GACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGA
CAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCU
GCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUC
CUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACA

AGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAG
GGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGA
ACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG
AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCG
ACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGU
CCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGA
AGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCA
AGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACA
GAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUG
GACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAA
GCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAAC
GACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCA
GCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUAC
CACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCA
AGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUA
CGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACA
GCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCAC
ACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAA
CAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGU
CCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGA
GGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCG
CAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGAC
AGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAG
AAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCA
GCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGU
CAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAA
AACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACG

AACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUA
CGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUC
GAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAG
CAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUAC
AACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCU
GUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACA
ACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACU
GAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAG
CUGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAGCUAG
CCAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUG
AAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAAC
ACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUG
CUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

REIVINDICAÇÕES

1. Método para inserir um ácido nucleico que codifica um polipeptídeo heterólogo em um lócus de albumina de uma célula hospedeira ou população de células **CARACTERIZADO** pelo fato de que compreende administrar:

i) um gRNA que compreende uma sequência escolhida de:

a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97;

d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2 a 33;

e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33;

f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34 a 97;

g) uma sequência que é complementar a 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2 a 33;

h) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 98 a 119;

i) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 98 a 119; e

j) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 120 a 163;

ii) um agente de ligação de DNA guiado por RNA; e

iii) um construto que compreende um ácido nucleico que codifica o polipeptídeo heterólogo,

assim inserindo o ácido nucleico que codifica o polipeptídeo heterólogo em um lócus de albumina da célula hospedeira ou população de células.

2. Método para expressar um polipeptídeo heterólogo de um lócus de albumina de uma célula hospedeira ou população de células **CARACTERIZADO** pelo fato de que compreende administrar:

i) um gRNA que compreende uma sequência escolhida de:

a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97;

d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2 a 33;

e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33;

f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34 a 97; e

g) uma sequência que compreende 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2 a 33;

ii) um agente de ligação de DNA guiado por RNA; e

iii) um construto que compreende uma sequência de codificação para o polipeptídeo heterólogo,

assim expressando o polipeptídeo heterólogo na célula hospedeira ou

população de células.

3. Método para expressar um agente terapêutico em um tipo de célula sem divisão ou população de células **CARACTERIZADO** pelo fato de que compreende administrar:

i) um gRNA que compreende uma sequência escolhida de:

a) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID Nos: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

b) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2, 8, 13, 19, 28, 29, 31, 32 e 33;

c) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 e 97;

d) uma sequência que é pelo menos 95%, 90%, 85%, 80% ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOS: 2 a 33;

e) pelo menos 17, 18, 19 ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33;

f) uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 34 a 97; e

g) uma sequência que compreende 15 nucleotídeos consecutivos +/- 10 nucleotídeos das coordenadas genômicas listadas para as SEQ ID NOs: 2 a 33;

ii) um agente de ligação de DNA guiado por RNA; e

iii) um construto que compreende uma sequência de codificação para um polipeptídeo heterólogo, assim expressando o agente terapêutico no tipo de célula sem divisão ou população de células.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, **CARACTERIZADO** pelo fato de que o gRNA compreende uma sequência guia selecionada do grupo que consiste na SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4,

SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 e SEQ ID NO: 33.

5. Método, de acordo com qualquer uma das reivindicações 1 a 4, sendo que o método é **CARACTERIZADO** pelo fato de que é realizado in vivo.

6. Método, de acordo com qualquer uma das reivindicações 1 a 4, sendo que o método é **CARACTERIZADO** pelo fato de que é realizado in vitro.

7. Método, de acordo com qualquer uma das reivindicações 1 a 6, **CARACTERIZADO** pelo fato de que o gRNA se liga a uma região a montante de um motivo adjacente de protoespaçador (PAM).

8. Método, de acordo com a reivindicação 7, **CARACTERIZADO** pelo fato de que o PAM é escolhido de NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT e NNNNRYAC.

9. Método, de acordo com qualquer uma das reivindicações 1 a 8, **CARACTERIZADO** pelo fato de que o gRNA é um gRNA duplo (dgRNA).

10. Método, de acordo com qualquer uma das reivindicações 1 a 8, **CARACTERIZADO** pelo fato de que o gRNA é um gRNA único (sgRNA).

11. Método, de acordo com a reivindicação 10, **CARACTERIZADO** pelo fato de que o sgRNA e compreende um ou mais nucleosídeos modificados.

12. Método, de acordo com qualquer uma das reivindicações 1 a 11, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é uma Cas9 ou um ácido nucleico que codifica uma Cas9.

13. Método, de acordo com qualquer uma das reivindicações 1 a 12, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é

um ácido nucleico que codifica o agente de ligação de DNA guiado por RNA.

14. Método, de acordo com a reivindicação 13, **CARACTERIZADO** pelo fato de que o ácido nucleico que codifica o agente de ligação de DNA guiado por RNA é um mRNA.

15. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que o mRNA é um mRNA modificado.

16. Método, de acordo com qualquer uma das reivindicações 1 a 15, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é uma Cas nuclease ou um ácido nucleico que codifica a Cas nuclease.

17. Método, de acordo com a reivindicação 16, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma Cas nuclease de classe 2.

18. Método, de acordo com a reivindicação 16 ou 17, **CARACTERIZADO** pelo fato de que a Cas nuclease é selecionada do grupo que consiste em nuclease de *S. pyogenes*, nuclease de *S. aureus*, nuclease de C. *jejuni*, nuclease de *S. thermophilus*, nuclease de *N. meningitidis* e variantes das mesmas.

19. Método, de acordo com qualquer uma das reivindicações 16 a 18, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma Cas9.

20. Método, de acordo com a reivindicação 19, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma Cas9 nuclease de *S. pyogenes*.

21. Método, de acordo com qualquer uma das reivindicações 16 a 20, **CARACTERIZADO** pelo fato de que a Cas nuclease tem atividade de ligação de DNA sítio-específica.

22. Método, de acordo com qualquer uma das reivindicações 16 a 21, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma nickase.

23. Método, de acordo com qualquer uma das reivindicações 16 a 21, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma clivase.

24. Método, de acordo com qualquer uma das reivindicações 19 a 21,

CARACTERIZADO pelo fato de que a Cas nuclease não tem atividade de nickase ou clivase.

25. Método, de acordo com qualquer uma das reivindicações 19 a 24, **CARACTERIZADO** pelo fato de que o construto de ácido nucleico é um construto doador independente de homologia.

26. Método, de acordo com qualquer uma das reivindicações 1 a 25, **CARACTERIZADO** pelo fato de que o construto é um construto de ácido nucleico bidirecional.

27. Método, de acordo com a reivindicação 26, **CARACTERIZADO** pelo fato de que o construto compreende:

i. um primeiro segmento que compreende uma sequência de codificação para um polipeptídeo heterólogo; e

ii. um segundo segmento que compreende um complemento reverso de uma sequência de codificação do polipeptídeo heterólogo.

28. Método, de acordo com qualquer uma das reivindicações 1 a 27, **CARACTERIZADO** pelo fato de que o construto compreende uma sequência sinal de poliadenilação.

29. Método, de acordo com qualquer uma das reivindicações 1 a 28, **CARACTERIZADO** pelo fato de que o construto compreende um sítio aceitante de splicing.

30. Método, de acordo com qualquer uma das reivindicações 1 a 29, **CARACTERIZADO** pelo fato de que o construto não compreende um braço de homologia.

31. Método, de acordo com qualquer uma das reivindicações 1 a 30, **CARACTERIZADO** pelo fato de que o gRNA é administrado em um vetor e/ou uma nanopartícula lipídica.

32. Método, de acordo com qualquer uma das reivindicações 1 a 31,

CARACTERIZADO pelo fato de que o agente de ligação de DNA guiado por RNA é administrado em um vetor e/ou uma nanopartícula lipídica.

33. Método, de acordo com qualquer uma das reivindicações 1 a 32, **CARACTERIZADO** pelo fato de que o construto que compreende o gene heterólogo é administrado em um vetor e/ou uma nanopartícula lipídica.

34. Método, de acordo com qualquer uma das reivindicações 31 a 33, **CARACTERIZADO** pelo fato de que o vetor é um vetor viral.

35. Método, de acordo com a reivindicação 34, **CARACTERIZADO** pelo fato de que o vetor viral é selecionado do grupo que consiste em um vetor viral adenoassociado (AAV), vetor de adenovírus, vetor de retrovírus e vetor de lentivírus.

36. Vetor, conforme definido na reivindicação 35, **CARACTERIZADO** pelo fato de que o vetor AAV é selecionado do grupo que consiste em AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV 12, rh10 e híbridos dos mesmos.

37. Método, conforme definido em qualquer uma das reivindicações 1 a 36, **CARACTERIZADO** pelo fato de que o gRNA, o agente de ligação de DNA guiado por RNA e o construto compreendem uma sequência de codificação para o polipeptídeo heterólogo, individualmente ou em qualquer combinação, são administrados simultaneamente.

38. Método, conforme definido em qualquer uma das reivindicações 1 a 36, **CARACTERIZADO** pelo fato de que o gRNA, o agente de ligação de DNA guiado por RNA e o construto que compreende uma sequência de codificação para o polipeptídeo heterólogo são administrados sequencialmente, em qualquer ordem e/ou em qualquer combinação.

39. Método, conforme definido em qualquer uma das reivindicações 1 a 36 e 38, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA,

ou agente de ligação de DNA guiado por RNA e gRNA em combinação, é administrado antes de fornecer o construto.

40. Método, conforme definido em qualquer uma das reivindicações 1 a 36 e 38, **CARACTERIZADO** pelo fato de que o construto que compreende uma sequência de codificação para o polipeptídeo heterólogo é administrado antes do gRNA e/ou agente de ligação de DNA guiado por RNA.

41. Método, conforme definido em qualquer uma das reivindicações 1 a 40, **CARACTERIZADO** pelo fato de que o construto de ácido nucleico bidirecional, agente de ligação de DNA guiado por RNA e gRNA, em qualquer combinação, são administrados com uma hora de intervalo entre os mesmos.

42. Método, conforme definido em qualquer uma das reivindicações 1 a 41, **CARACTERIZADO** pelo fato de que o polipeptídeo heterólogo é um polipeptídeo secretado.

43. Método, conforme definido em qualquer uma das reivindicações 1 a 41, **CARACTERIZADO** pelo fato de que o polipeptídeo heterólogo é um polipeptídeo intracelular.

44. Método, conforme definido em qualquer uma das reivindicações 1 a 43, **CARACTERIZADO** pelo fato de que a célula é uma célula de fígado.

45. Método, conforme definido em a reivindicação 44, **CARACTERIZADO** pelo fato de que a célula de fígado é um hepatócito.

46. Método, conforme definido em qualquer uma das reivindicações 1 a 45, **CARACTERIZADO** pelo fato de que a expressão do polipeptídeo heterólogo na célula hospedeira é aumentada em pelo menos cerca de 2%, cerca de 3%, cerca de 4%, cerca de 5%, cerca de 6%, cerca de 7%, cerca de 8%, cerca de 9%, cerca de 10% ou mais, em relação a um nível na célula antes de administrar o gRNA, agente de ligação de DNA guiado por RNA e construto que compreende uma sequência de codificação para o polipeptídeo heterólogo.

47. Método, conforme definido em qualquer uma das reivindicações 1 a 46, **CARACTERIZADO** pelo fato de que o gRNA compreende a SEQ ID NO: 301 ou SEQ ID NO: 302.

48. Método, conforme definido em qualquer uma das reivindicações 1 a 47, **CARACTERIZADO** pelo fato de que o gRNA media o corte específico de alvo por um agente de ligação de DNA guiado por RNA, resultando na inserção da sequência de codificação para o polipeptídeo heterólogo no íntron 1 de um gene de albumina.

49. Método, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADO** pelo fato de que o corte resulta em uma taxa de pelo menos cerca de 2%, cerca de 5%, ou cerca de 10% de inserção de um ácido nucleico heterólogo na população de células.

50. Método, de acordo com a reivindicação 49, **CARACTERIZADO** pelo fato de que o corte resulta em uma taxa entre cerca de 30 e 35%, cerca de 35 e 40%, cerca de 40 e 45%, cerca de 45 e 50%, cerca de 50 e 55%, cerca de 55 e 60%, cerca de 60 e 65%, cerca de 65 e 70%, cerca de 70 e 75%, cerca de 75 e 80%, cerca de 80 e 85%, cerca de 85 e 90%, cerca de 90 e 95%, ou cerca de 95 e 99% de inserção da sequência de codificação para o polipeptídeo heterólogo.

51. Método, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADO** pelo fato de que a proteína de ligação de DNA guiada por RNA é uma Cas9 nuclease de *S. pyogenes*.

52. Método, de acordo com a reivindicação 51, **CARACTERIZADO** pelo fato de que a nuclease é uma clivase ou uma nickase.

53. Método, conforme definido em qualquer uma das reivindicações 1 a 52, **CARACTERIZADO** pelo fato de que compreende adicionalmente administrar uma LNP que compreende o gRNA.

54. Método, conforme definido em qualquer uma das reivindicações 1 a 53, **CARACTERIZADO** pelo fato de que compreende adicionalmente administrar uma

LNP que compreende um mRNA que codifica o agente de ligação de DNA guiado por RNA.

55. Método, de acordo com a reivindicação 54, **CARACTERIZADO** pelo fato de que a LNP compreende o gRNA e o mRNA que codificam o agente de ligação de DNA guiado por RNA.

56. Método, conforme definido em qualquer uma das reivindicações 1 a 55, **CARACTERIZADO** pelo fato de que o gRNA e a proteína de ligação de DNA guiado por RNA são administrados como um RNP.

57. Método, conforme definido em qualquer uma das reivindicações 53 a 56, **CARACTERIZADO** pelo fato de que o construto é administrado por meio de um vetor.

58. Célula hospedeira **CARACTERIZADA** pelo fato de que é produzida pelo método, conforme definido em qualquer uma das reivindicações 1 a 57.

59. Célula hospedeira **CARACTERIZADA** pelo fato de que compreende um construto de ácido nucleico bidirecional que codifica um polipeptídeo heterólogo integrado a um íntron 1 de um lócus de albumina de uma célula hospedeira.

60. Célula hospedeira, de acordo com a reivindicação 58 ou 59, sendo que a célula hospedeira é **CARACTERIZADA** pelo fato de que é uma célula de fígado.

61. Célula hospedeira, de acordo com qualquer uma das reivindicações 58 a 60, **CARACTERIZADA** pelo fato de que a célula de fígado é um hepatócito.

62. Método ou célula hospedeira, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o agente de ligação de DNA guiado por RNA é um ácido nucleico que codifica um agente de ligação de DNA guiado por RNA.

63. Método ou célula hospedeira, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o agente de ligação de DNA guiado por RNA é uma Cas nuclease.

64. Método ou célula hospedeira, conforme definido em qualquer uma das

reivindicações anteriores, **CARACTERIZADA** pelo fato de que o agente de ligação de DNA guiado por RNA é um ácido nucleico que codifica uma Cas nuclease.

65. Método ou célula hospedeira, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o agente de ligação de DNA guiado por RNA é um mRNA que codifica uma Cas nuclease.

66. Método ou célula hospedeira, de acordo com a reivindicação 65, **CARACTERIZADA** pelo fato de que o mRNA é um mRNA modificado.

67. Método ou célula hospedeira, de acordo com qualquer uma das reivindicações 63 a 66, **CARACTERIZADA** pelo fato de que a Cas nuclease é uma Cas nuclease de classe 2.

68. Método ou composição, conforme definido em qualquer uma das reivindicações 63 a 67, **CARACTERIZADA** pelo fato de que a Cas nuclease é uma Cas9.

69. Método ou composição, conforme definido em qualquer uma das reivindicações 63 a 68, **CARACTERIZADA** pelo fato de que a Cas nuclease é selecionada do grupo que consiste em nuclease de *S. pyogenes*, nuclease de *S. aureus*, nuclease de *C. jejuni*, nuclease de *S. thermophilus*, *nuclease de N. meningitidis* e variantes das mesmas.

70. Método ou composição, de acordo com a reivindicação 69, **CARACTERIZADA** pelo fato de que a Cas nuclease é uma Cas9 nuclease de *S. pyogenes* ou uma variante da mesma.

71. Método ou composição, conforme definido em qualquer uma das reivindicações 63 a 70, **CARACTERIZADA** pelo fato de que a Cas nuclease tem atividade de ligação de DNA sítio-específica.

72. Método ou composição, conforme definido em qualquer uma das reivindicações 63 a 71, **CARACTERIZADA** pelo fato de que a Cas nuclease é uma nickase.

73. Método ou composição, conforme definido em qualquer uma das reivindicações 63 a 72, **CARACTERIZADA** pelo fato de que a Cas nuclease é uma clivase.

74. Método, de acordo com qualquer uma das reivindicações 63 a 73, **CARACTERIZADO** pelo fato de que a Cas nuclease não tem atividade de nickase ou clivase.

75. Método, conforme definido em qualquer uma das reivindicações 1 a 74, **CARACTERIZADO** pelo fato de que compreende alcançar a atividade de polipeptídeo heterólogo ou níveis de polipeptídeo heterólogo de pelo menos cerca de 1% do normal, por exemplo, pelo menos cerca de 5% do normal.

76. Método, conforme definido em qualquer uma das reivindicações 1 a 75, **CARACTERIZADO** pelo fato de que a atividade de polipeptídeo heterólogo ou nível de polipeptídeo heterólogo é menor do que cerca de 500% do normal.

77. Método, conforme definido em qualquer uma das reivindicações 1 a 76, **CARACTERIZADO** pelo fato de que compreende alcançar a atividade de polipeptídeo heterólogo ou níveis de polipeptídeo heterólogo de pelo menos cerca de 1% a 300% do normal.

78. Método in vivo, conforme definido em qualquer uma das reivindicações 5 a 77, **CARACTERIZADO** pelo fato de que compreende adicionalmente alcançar um efeito durável, por exemplo, um efeito de pelo menos 1 mês, 2 meses, 6 meses, 1 ano ou 2 anos no indivíduo.

79. Método in vivo, conforme definido em qualquer uma das reivindicações 5 a 78, **CARACTERIZADO** pelo fato de que os níveis de albumina de circulação do indivíduo são normais em pelo menos 1 mês, 2 meses, 6 meses ou 1 ano após administrar o construto de ácido nucleico.

80. Método in vivo, conforme definido em qualquer uma das reivindicações 5 a 79, **CARACTERIZADO** pelo fato de que os níveis de albumina de circulação do

indivíduo são mantidos em 4 semanas após administrar o construto de ácido nucleico bidirecional.

81. Método in vivo, conforme definido em qualquer uma das reivindicações 5 a 80, **CARACTERIZADO** pelo fato de que os níveis de albumina de circulação do indivíduo reduzem de modo transiente, então, retornam ao normal.

82. Método, conforme definido em qualquer uma das reivindicações 1 a 81, **CARACTERIZADO** pelo fato de que o RNA guia compreende pelo menos 17, 18, 19, ou 20 nucleotídeos contíguos de uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33.

83. Método, conforme definido em qualquer uma das reivindicações 1 a 82, **CARACTERIZADO** pelo fato de que o RNA guia compreende uma sequência que é pelo menos 95%, 90%, 85%, 80%, ou 75% idêntica a uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33.

84. Método, conforme definido em qualquer uma das reivindicações 1 a 83, **CARACTERIZADO** pelo fato de que o RNA guia compreende uma sequência selecionada do grupo que consiste nas SEQ ID NOs: 2 a 33.

85. Método ou célula hospedeira, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende a SEQ ID NO: 301 ou 302.

86. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 2.

87. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 3.

88. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia

compreende uma sequência de ácido nucleico da SEQ ID NO: 4.

89. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 5.

90. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 6.

91. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 7.

92. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 8.

93. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 9.

94. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 10.

95. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 11.

96. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 12.

97. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia

compreende uma sequência de ácido nucleico da SEQ ID NO: 13.

98. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 14.

99. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 15.

100. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 16.

101. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 17.

102. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 18.

103. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 19.

104. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 20.

105. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 21.

106. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia

compreende uma sequência de ácido nucleico da SEQ ID NO: 22.

107. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 23.

108. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 24.

109. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 25.

110. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 26.

111. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 27.

112. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 28.

113. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 29.

114. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 30.

115. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia

compreende uma sequência de ácido nucleico da SEQ ID NO: 31.

116. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 32.

117. Método e composição, conforme definido em qualquer uma das reivindicações anteriores, **CARACTERIZADA** pelo fato de que o RNA guia compreende uma sequência de ácido nucleico da SEQ ID NO: 33.

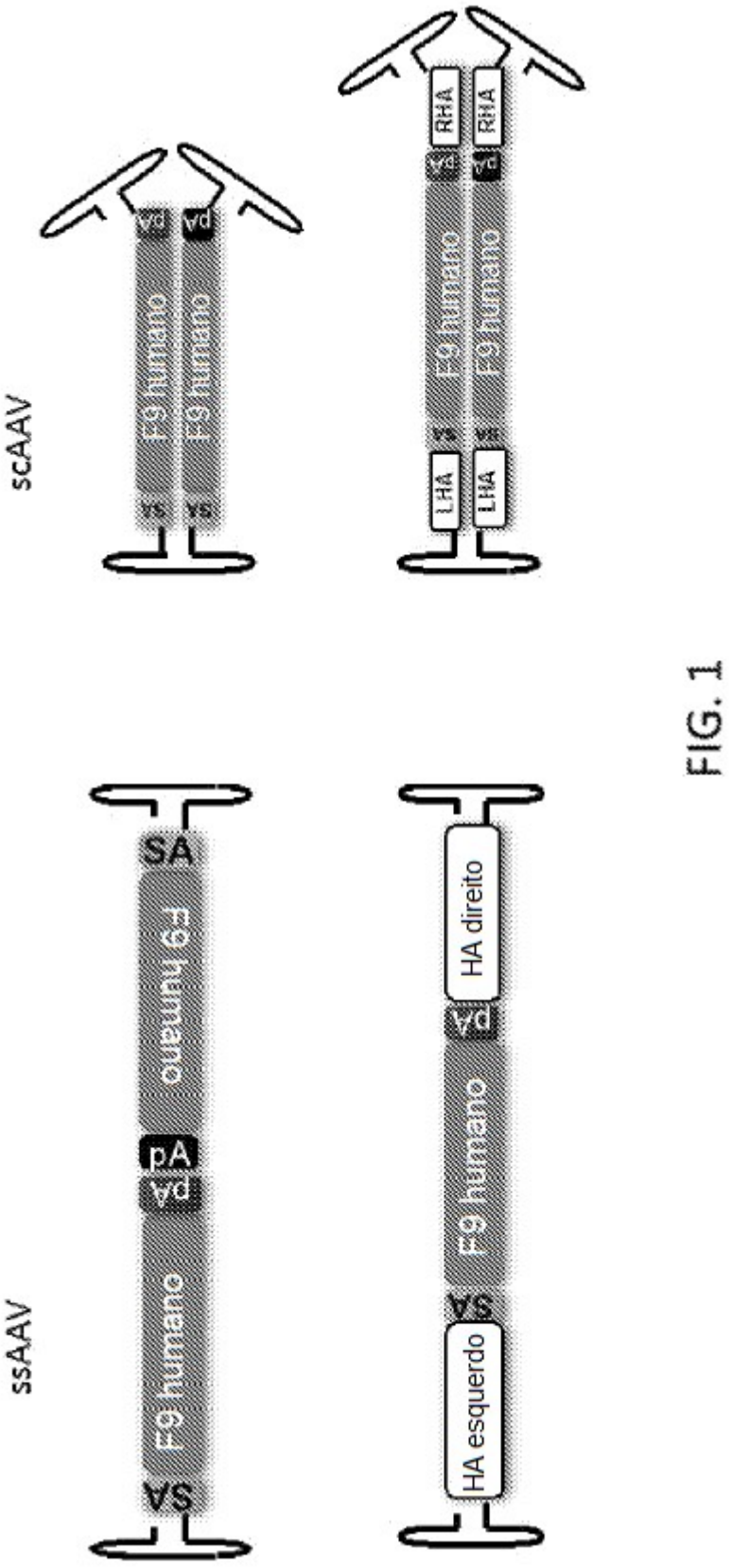


FIG. 1

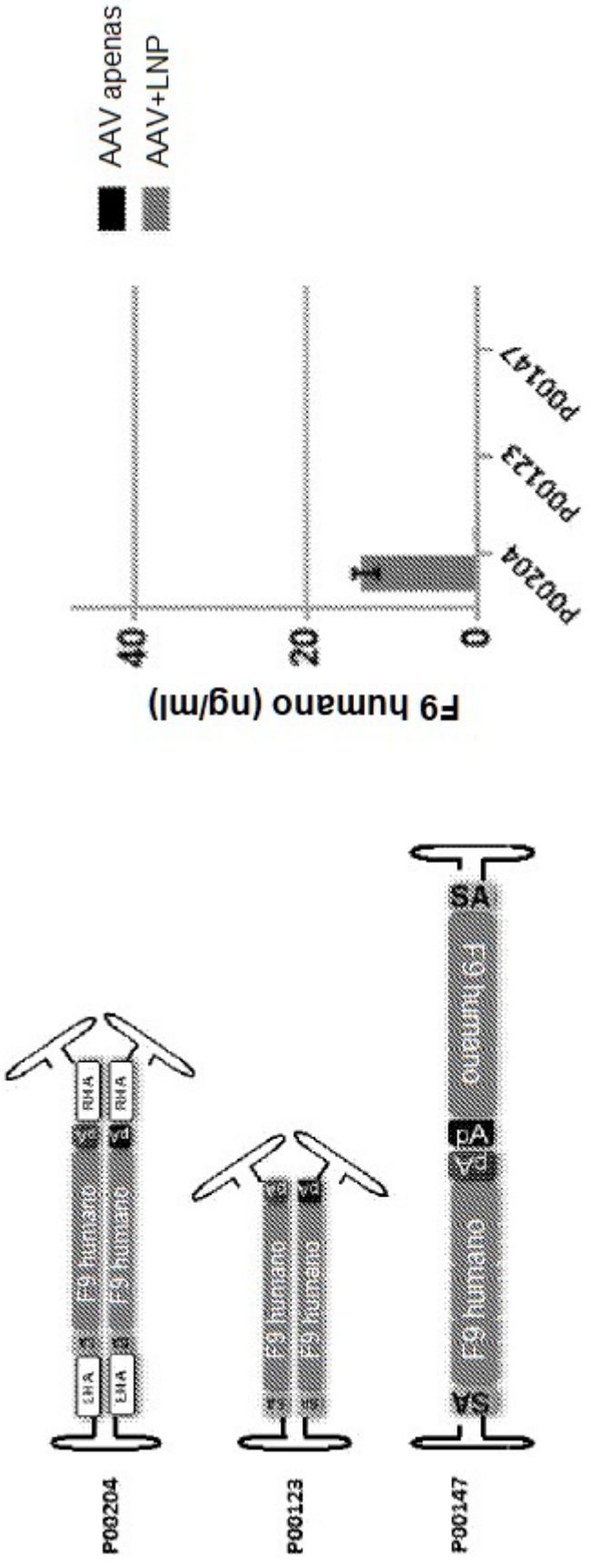

AAV apenas
AAV+LNP
F9 humano (ng/ml)
0
20
40
P00204
P00123
P00147
LHA
RHA
F9 humano
SA
pA
P00204
P00123
P00147

FIG. 2

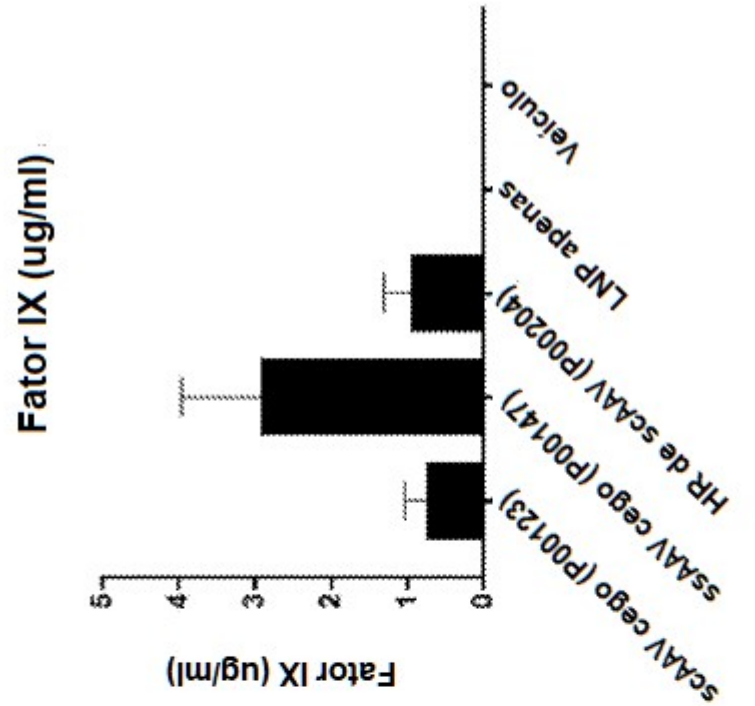

Fator IX (ug/ml)
Fator IX (ug/ml)
0
1
2
3
4
5
scAAV cego (P00123)
ssAAV cego (P00147)
HR de scAAV (P00204)
LNP apenas
Veículo

FIG. 3B

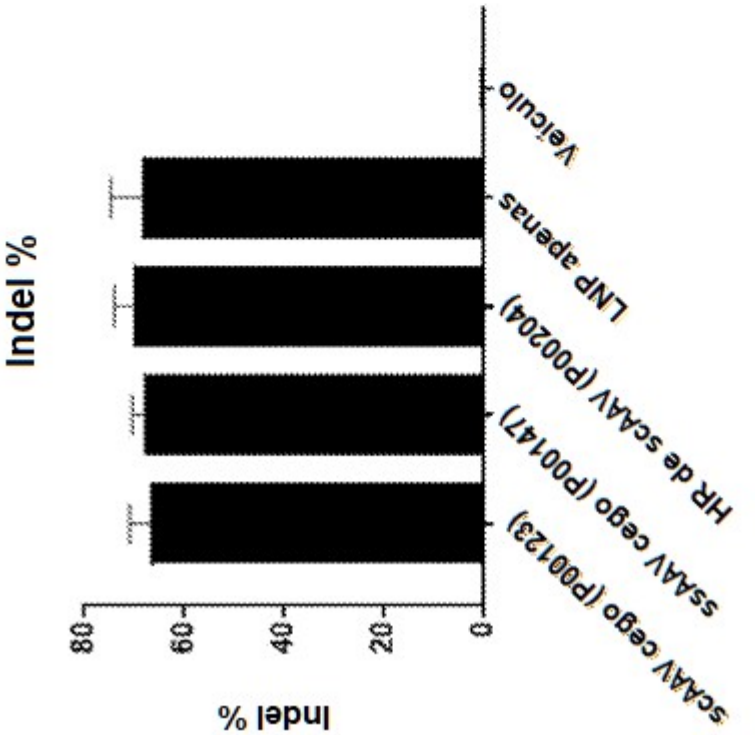

Indel %
Indel %
0
20
40
60
80
scAAV cego (P00123)
ssAAV cego (P00147)
HR de scAAV (P00204)
LNP apenas
Veículo

FIG. 3A

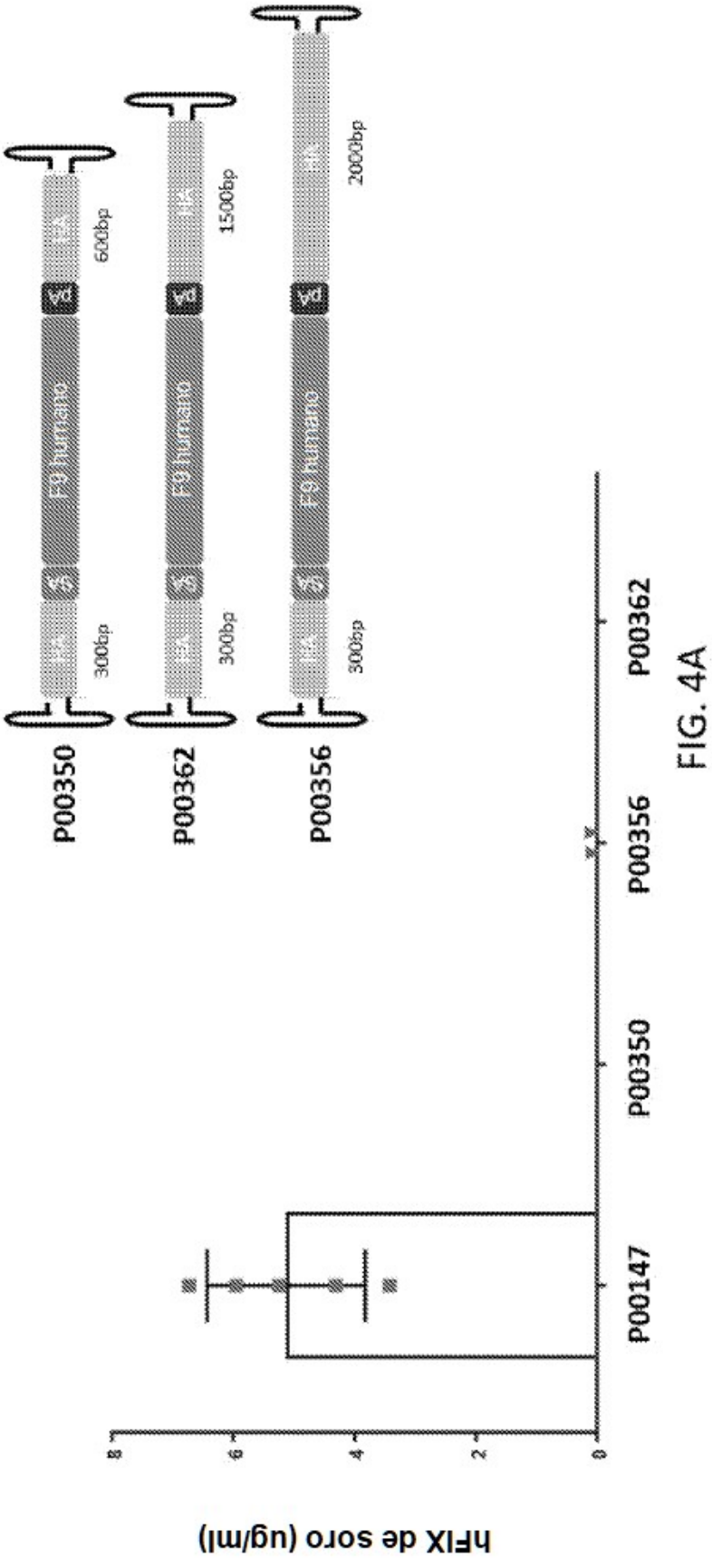

P00350
TFA
300bp
YS
F9 humano
YD
TFA
600bp
P00362
TFA
300bp
YS
F9 humano
YD
TFA
1500bp
P00356
TFA
300bp
YS
F9 humano
YD
TFA
2000bp
P00147
P00350
P00356
P00362
hFIX de soro (ug/ml)
0
2
4
6
8

FIG. 4A

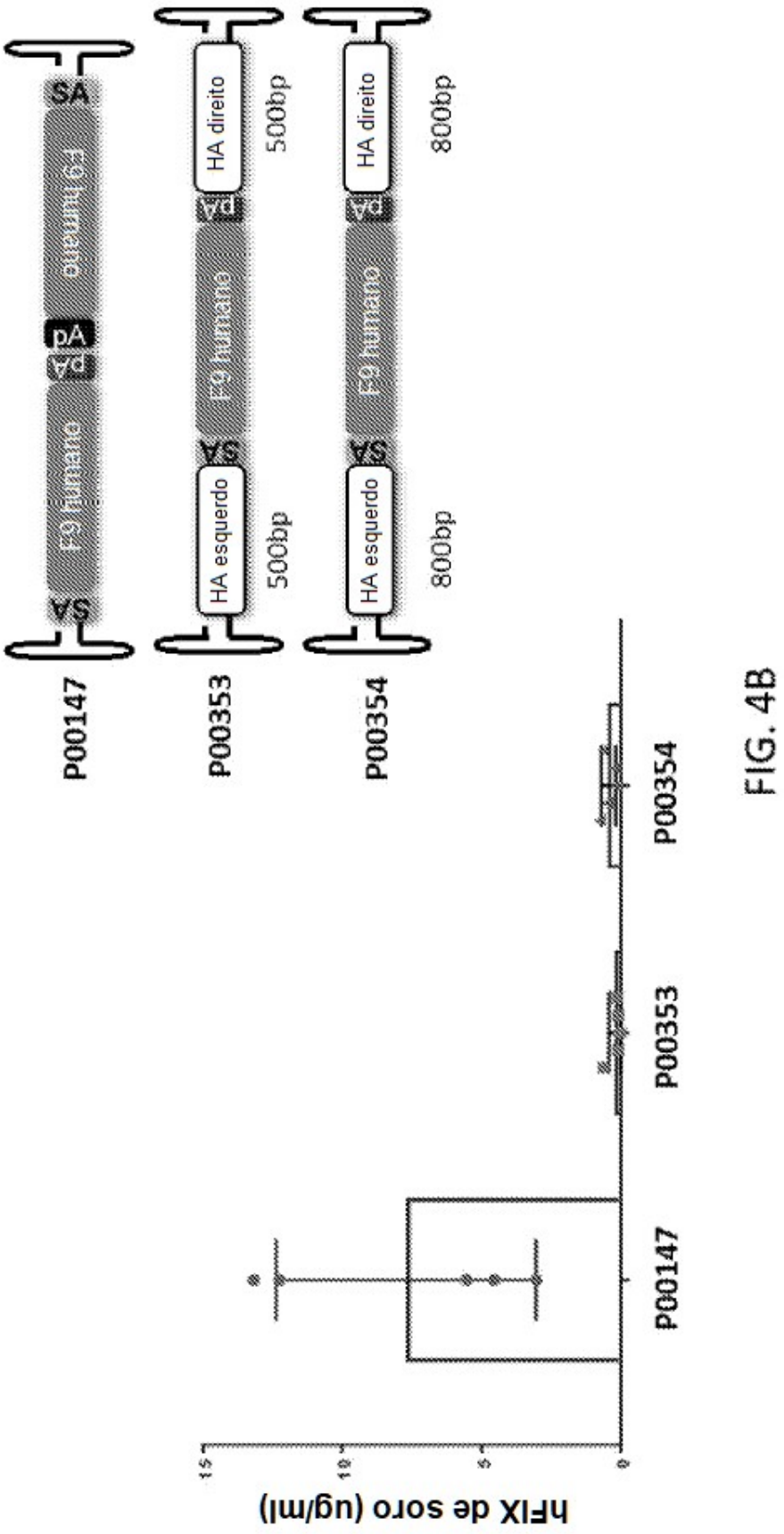

P00147
SA
F9 humano
pA
pA
SA
P00353
HA esquerdo
500bp
SA
F9 humano
pA
HA direito
500bp
P00354
HA esquerdo
800bp
SA
F9 humano
pA
HA direito
800bp
hFIX de soro (ug/ml)
0
5
10
15
P00147
P00353
P00354

FIG. 4B

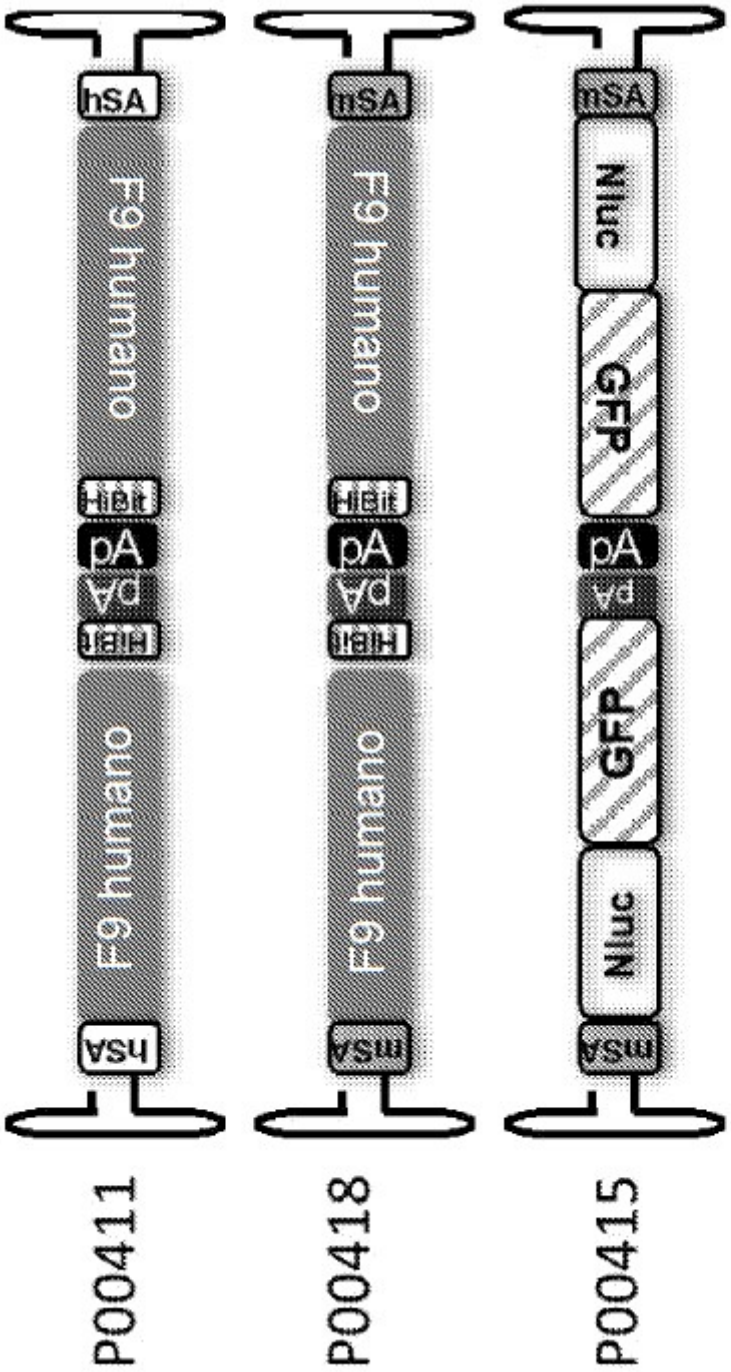

hSA
F9 humano
HiBit
pA
pA
HiBit
hSA
P00411
mSA
F9 humano
HiBit
pA
pA
HiBit
mSA
P00418
mSA
Nluc
GFP
pA
pA
GFP
Nluc
mSA
P00415

FIG. 5A

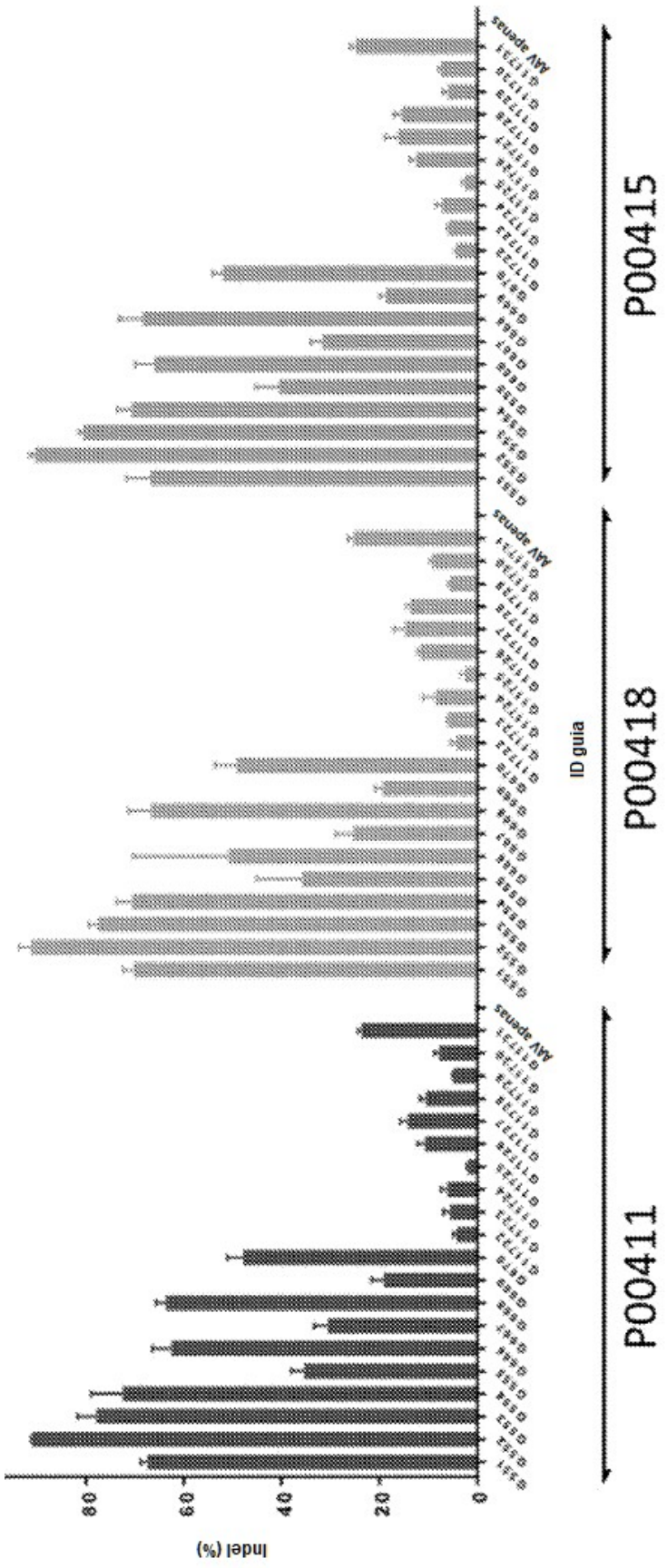


FIG. 5B

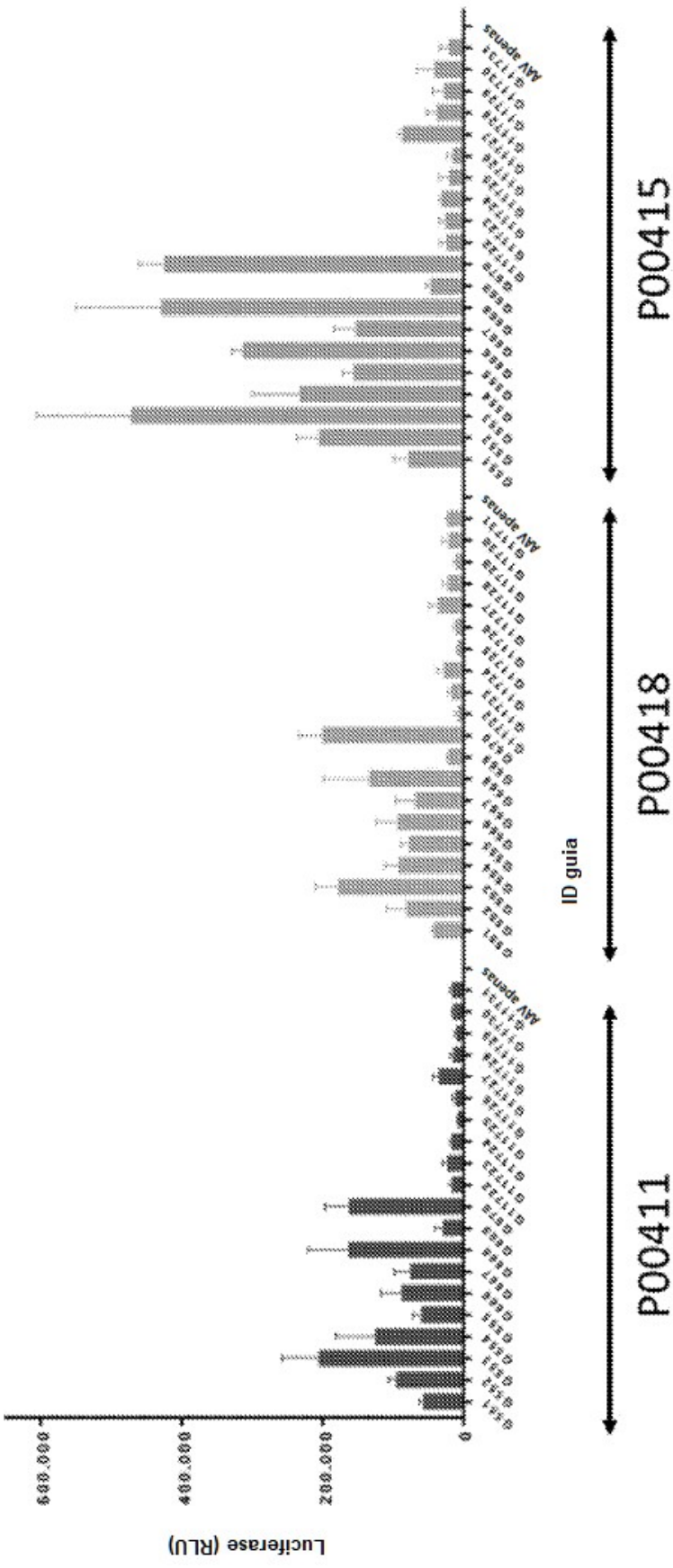

Luciferase (RLU)
600.000
400.000
200.000
0
ID guia
P00411
P00418
P00415

FIG. 5C

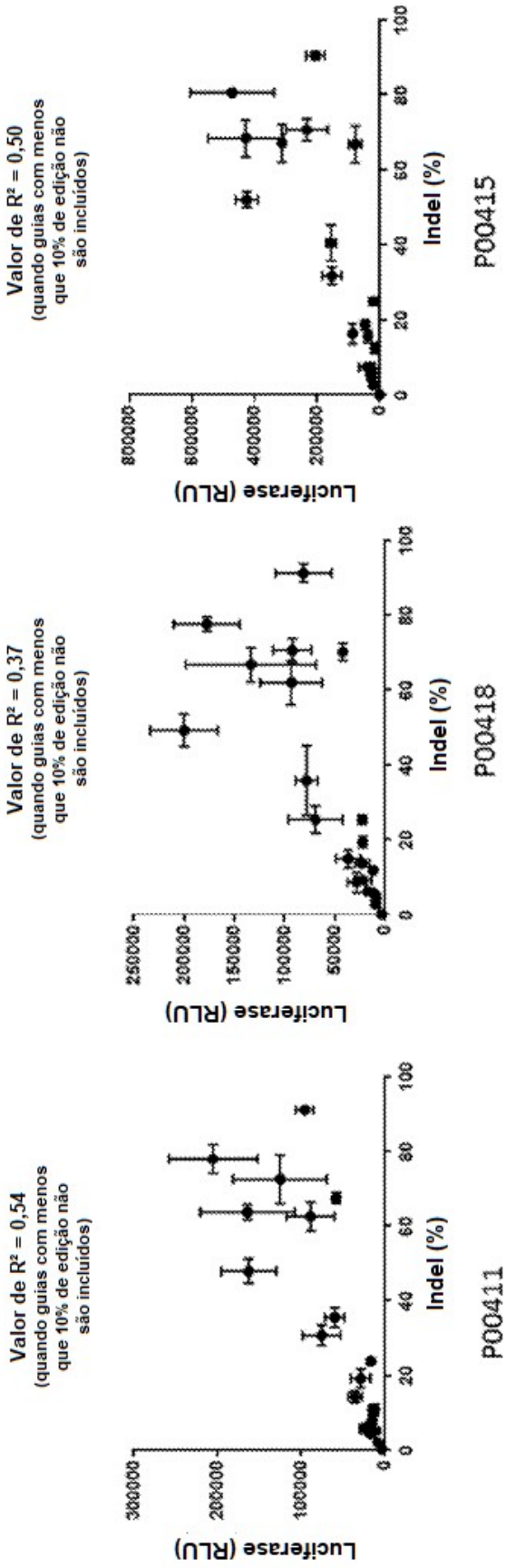
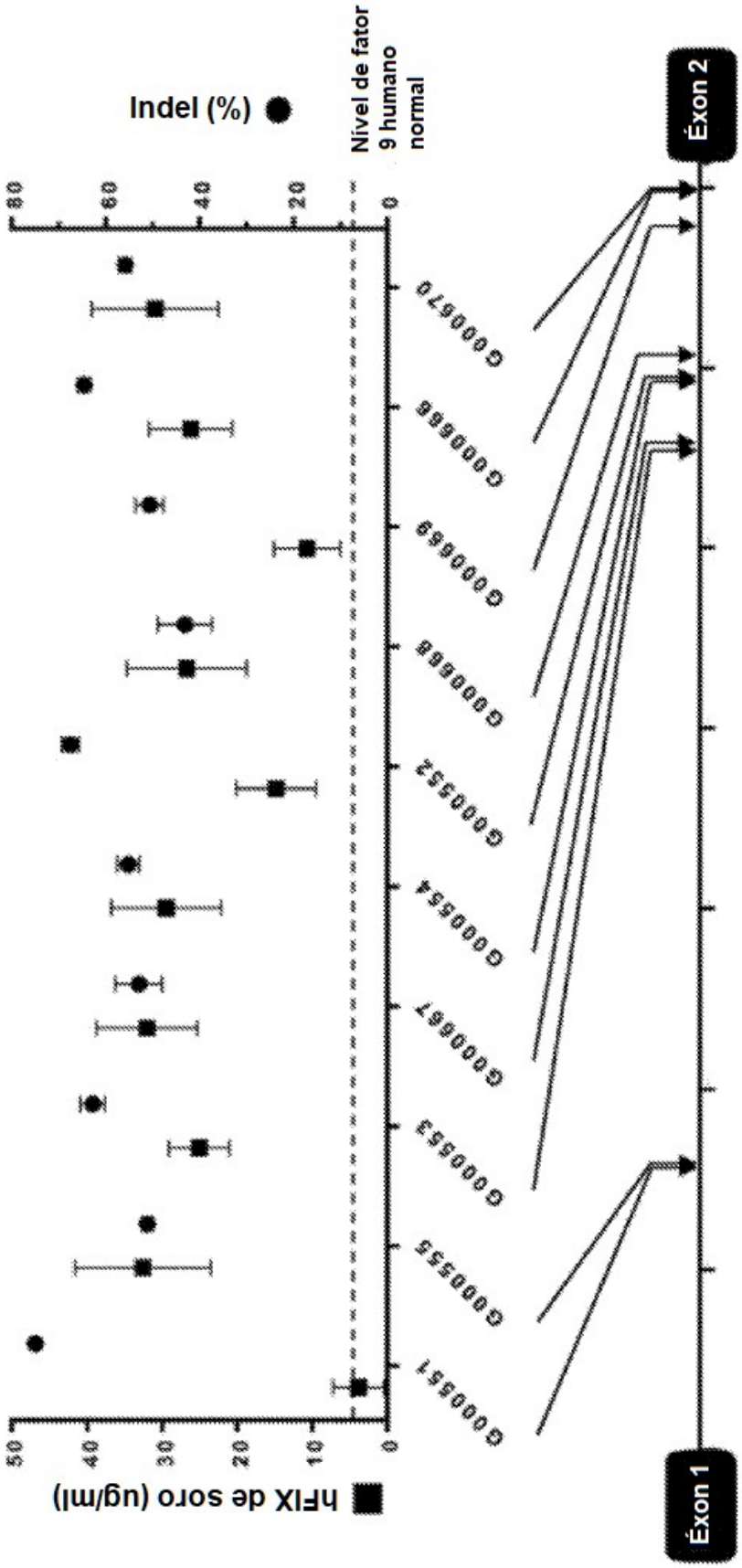

Valor de R² = 0,54
(quando guias com menos que 10% de edição não são incluídos)
Luciferase (RLU)
Indel (%)
P00411
Valor de R² = 0,37
(quando guias com menos que 10% de edição não são incluídos)
Luciferase (RLU)
Indel (%)
P00418
Valor de R² = 0,50
(quando guias com menos que 10% de edição não são incluídos)
Luciferase (RLU)
Indel (%)
P00415

FIG. 5D

FIG. 6


Indel (%)
Nível de fator 9 humano normal
hFIX de soro (ug/ml)
0
10
20
30
40
50
0
20
40
60
80
G000571
G000553
G000553
G000567
G000554
G000552
G000569
G000568
G000566
G000570
Éxon 1
Éxon 2

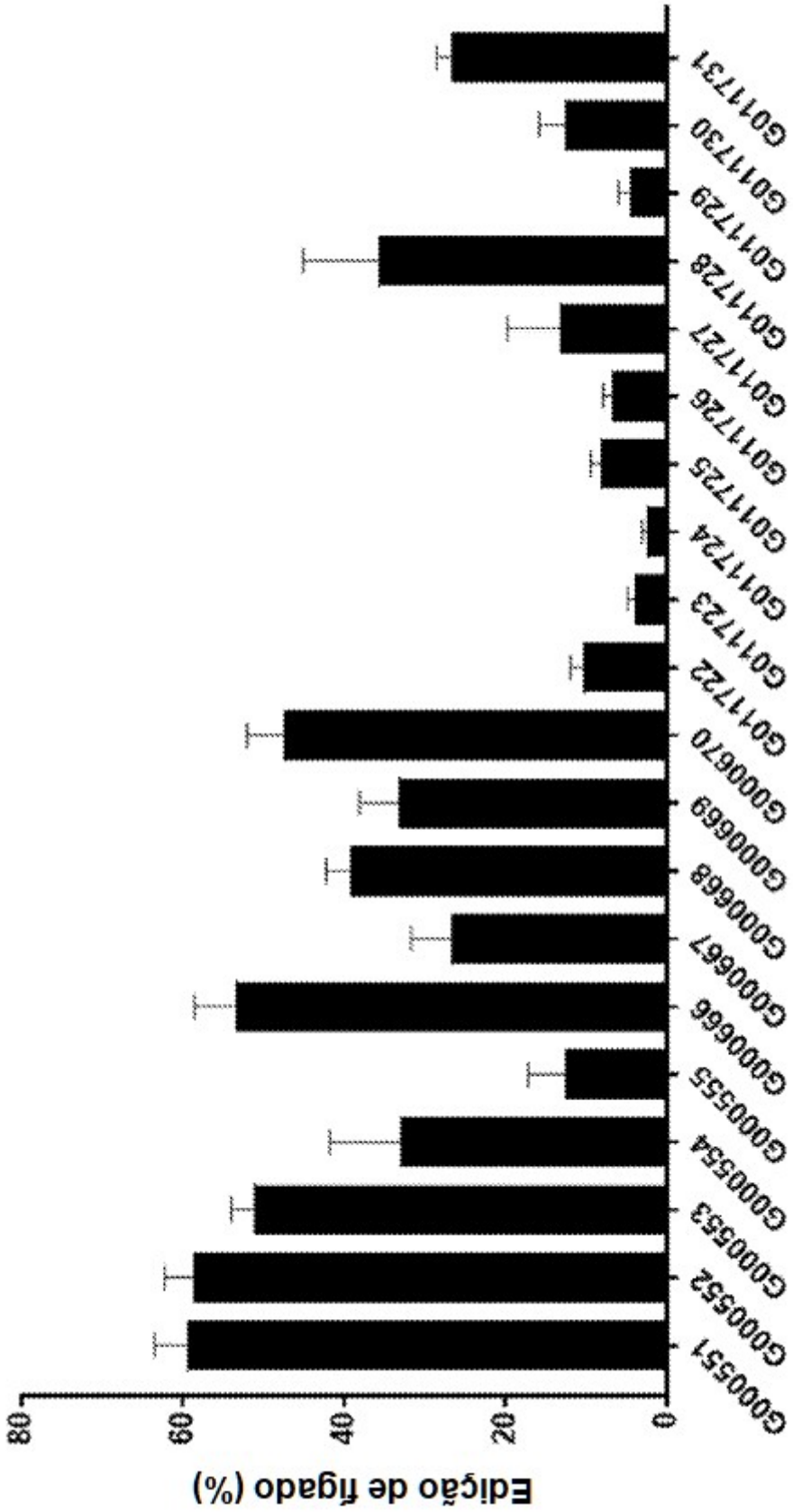

Edição de fígado (%)
0
20
40
60
80
G000551
G000552
G000553
G000554
G000555
G000666
G000667
G000668
G000669
G000670
G011722
G011723
G011724
G011725
G011726
G011727
G011728
G011729
G011730
G011731

FIG. 7A

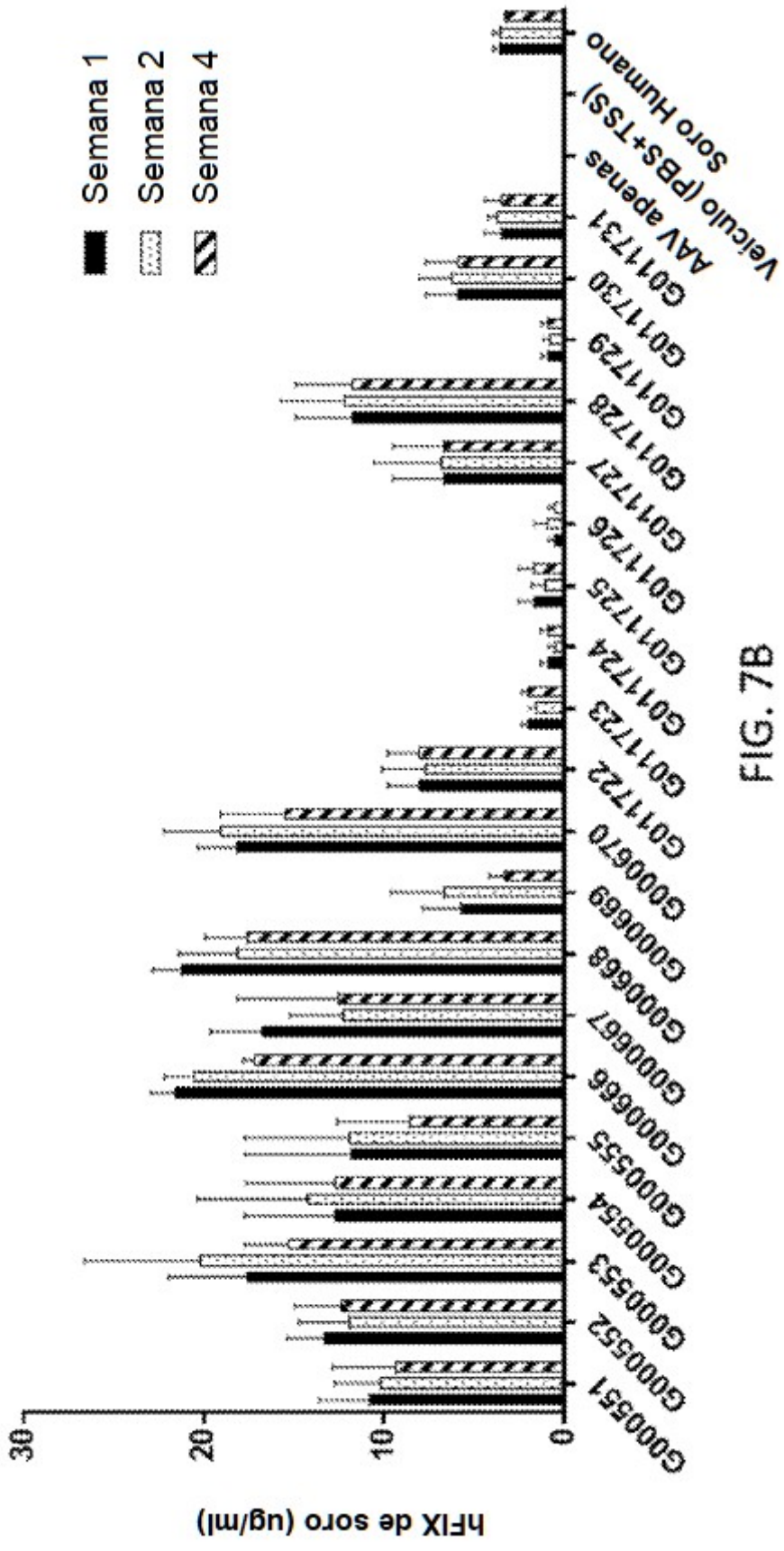

Semana 1
Semana 2
Semana 4
hFIX de soro (ug/ml)
30
20
10
0
G0005S1
G0005S2
G0005S3
G0005S4
G0005S5
G0006A6
G0006A7
G0006A8
G0006A9
G0006C0
G0117C2
G0117C3
G0117C4
G0117C5
G0117C6
G0117C7
G0117C8
G0117C9
G0117D0
G0117D1
AAV apenas
Veículo (PBS+TSS)
soro Humano

FIG. 7B

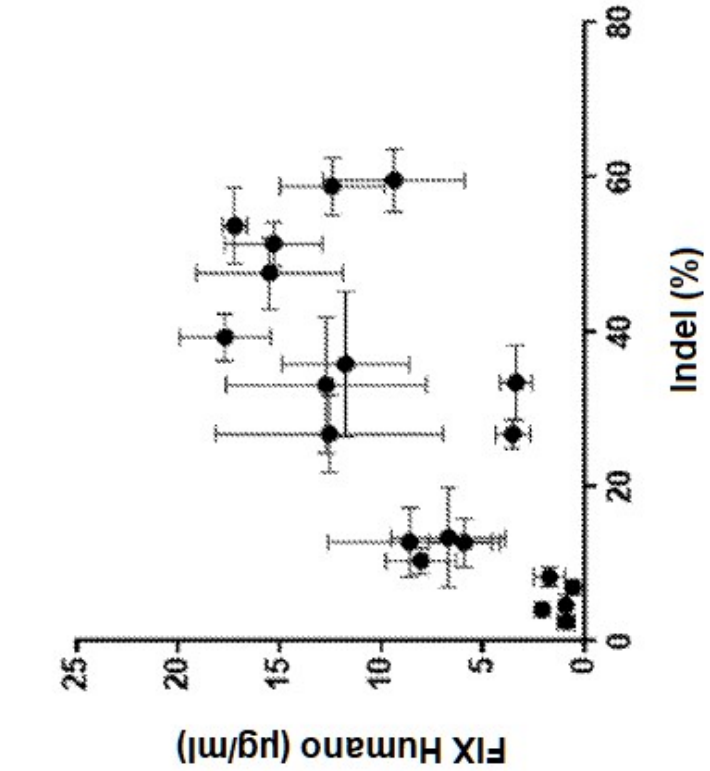

FIX Humano (µg/ml)
Indel (%)
0
5
10
15
20
25
0
20
40
60
80

FIG. 7D

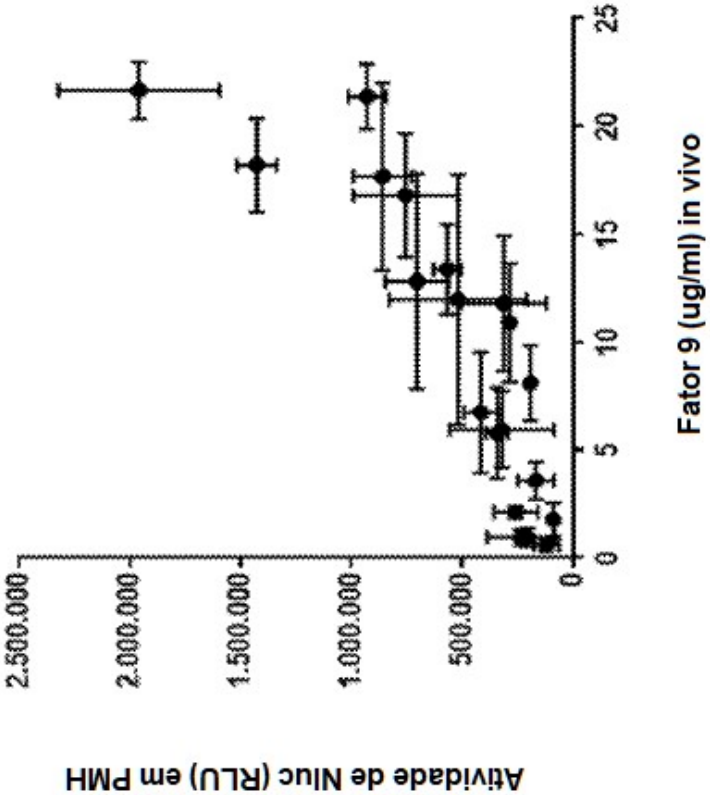

Atividade de Nluc (RLU) em PMH
Fator 9 (ug/ml) in vivo
0
500.000
1.000.000
1.500.000
2.000.000
2.500.000
0
5
10
15
20
25

FIG. 7C

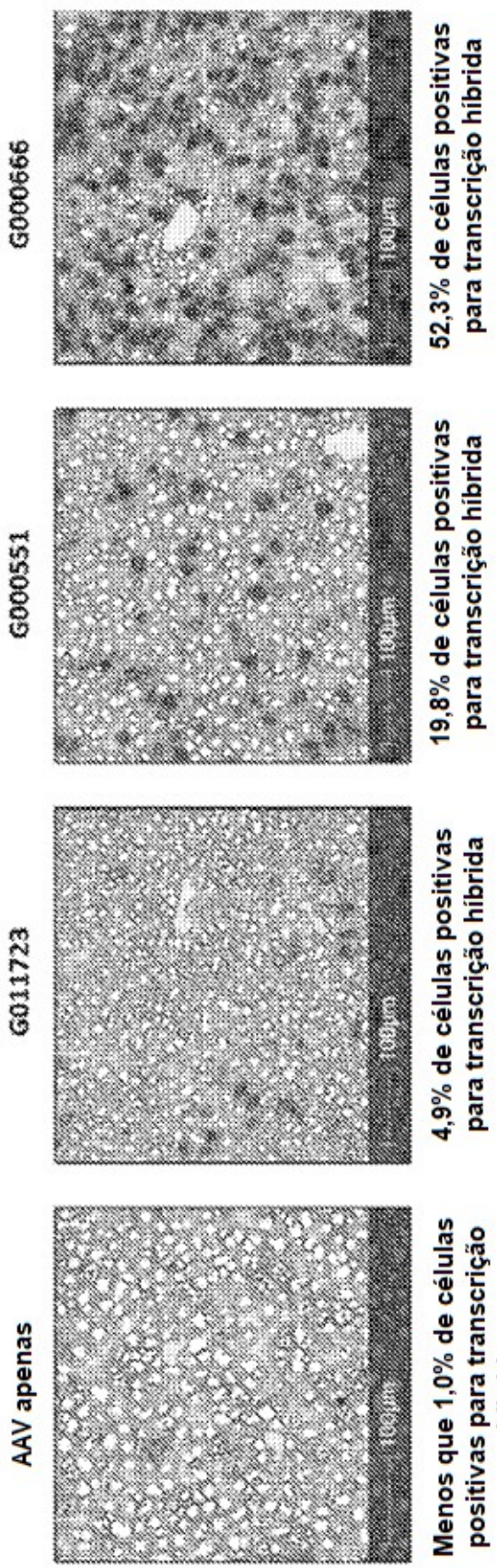
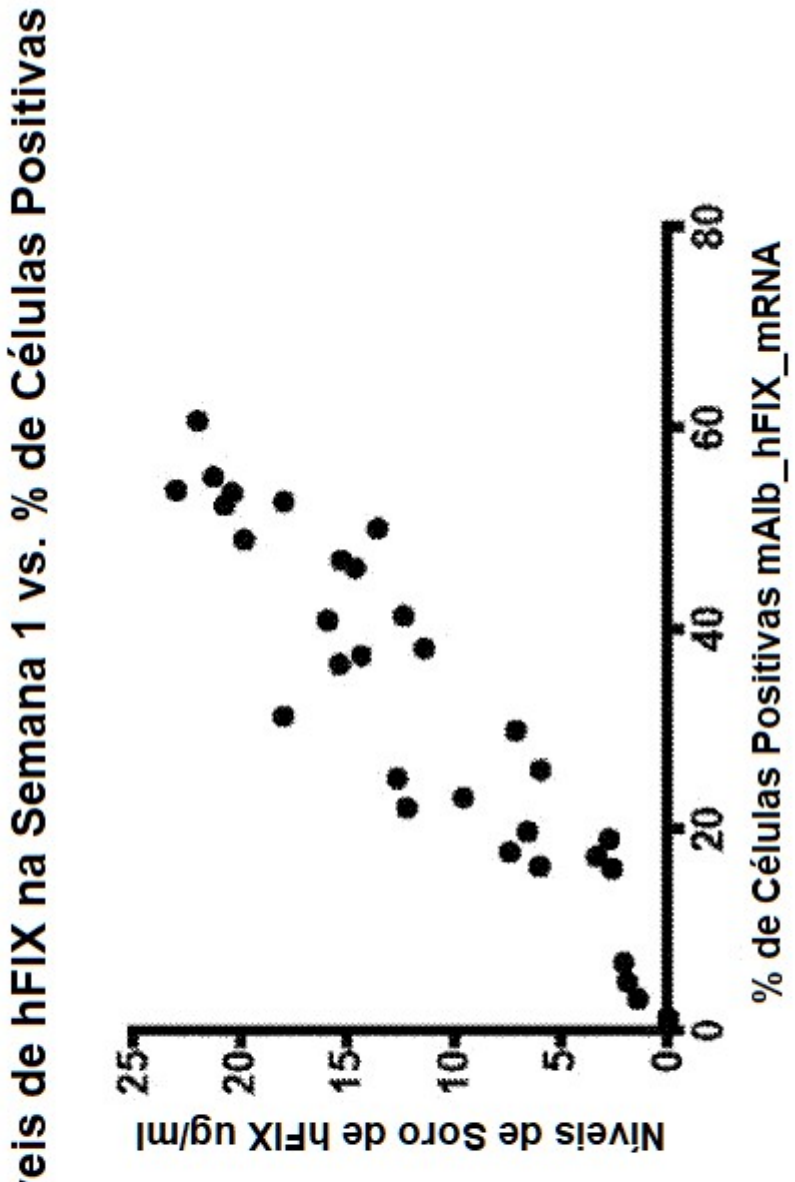

AAV apenas
Menos que 1,0% de células positivas para transcrição híbrida
G011723
4,9% de células positivas para transcrição híbrida
G000551
19,8% de células positivas para transcrição híbrida
G000666
52,3% de células positivas para transcrição híbrida

FIG. 8A

FIG. 8B


Níveis de hFIX na Semana 1 vs. % de Células Positivas
Níveis de Soro de hFIX ug/ml
0
5
10
15
20
25
0
20
40
60
80
% de Células Positivas mAlb_hFIX_mRNA

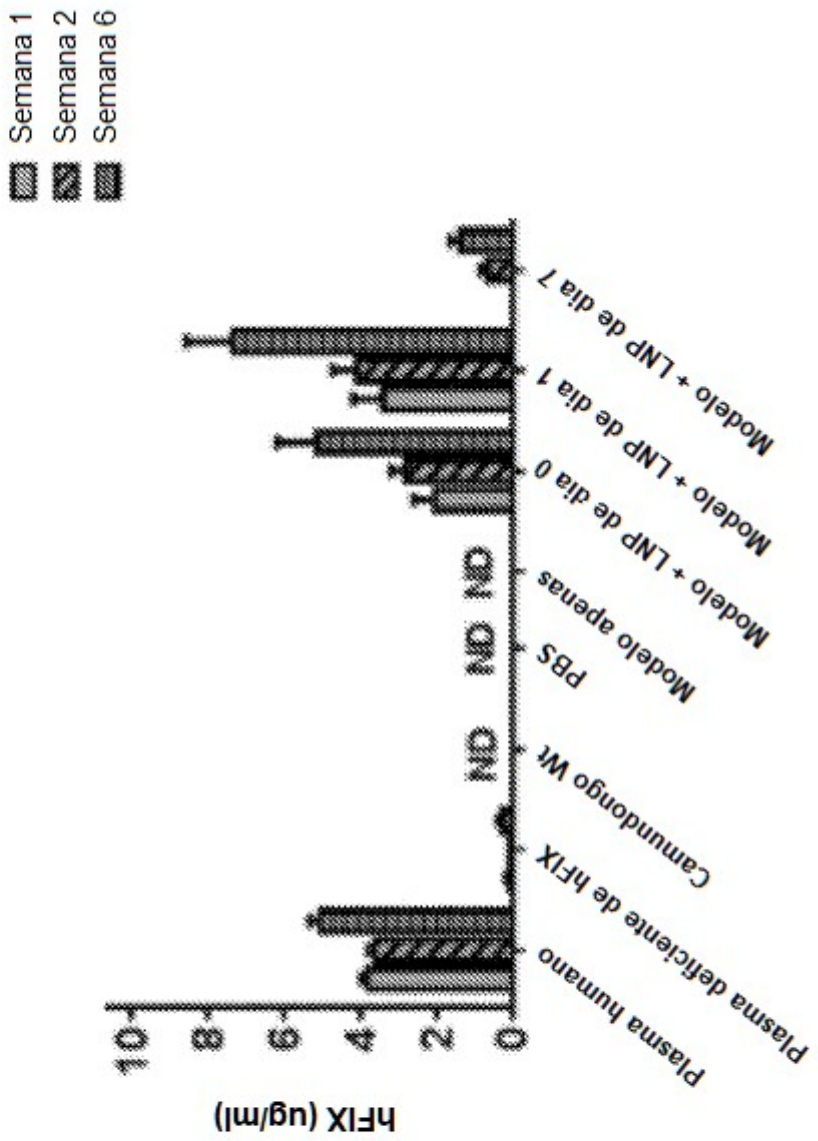

Semana 1
Semana 2
Semana 6
hFIX (ug/ml)
0
2
4
6
8
10
ND
ND
ND
Plasma humano
Plasma deficiente de hFIX
Camundongo Wt
PBS
Modelo apenas
Modelo + LNP de dia 0
Modelo + LNP de dia 1
Modelo + LNP de dia 7

FIG. 9

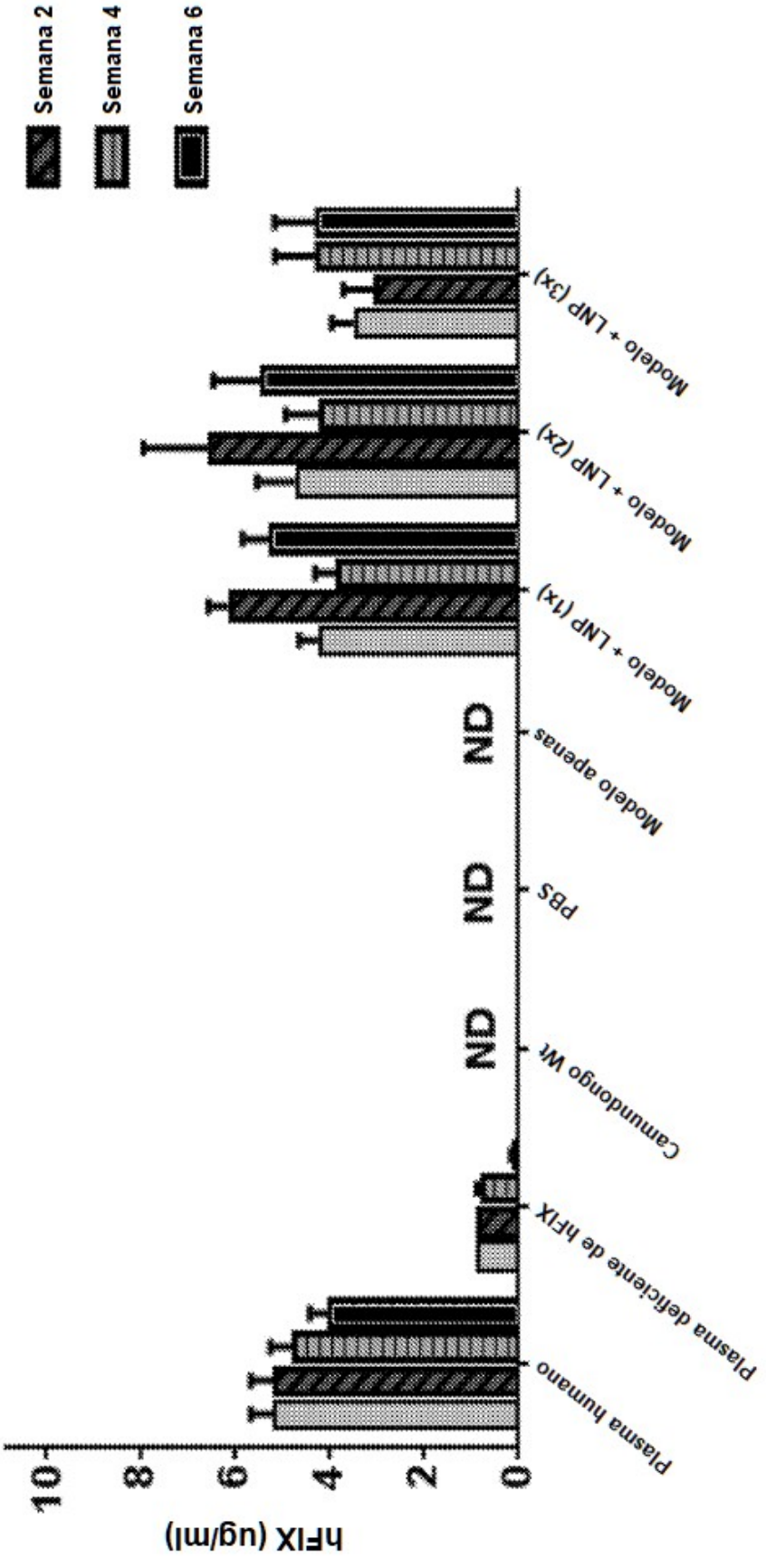
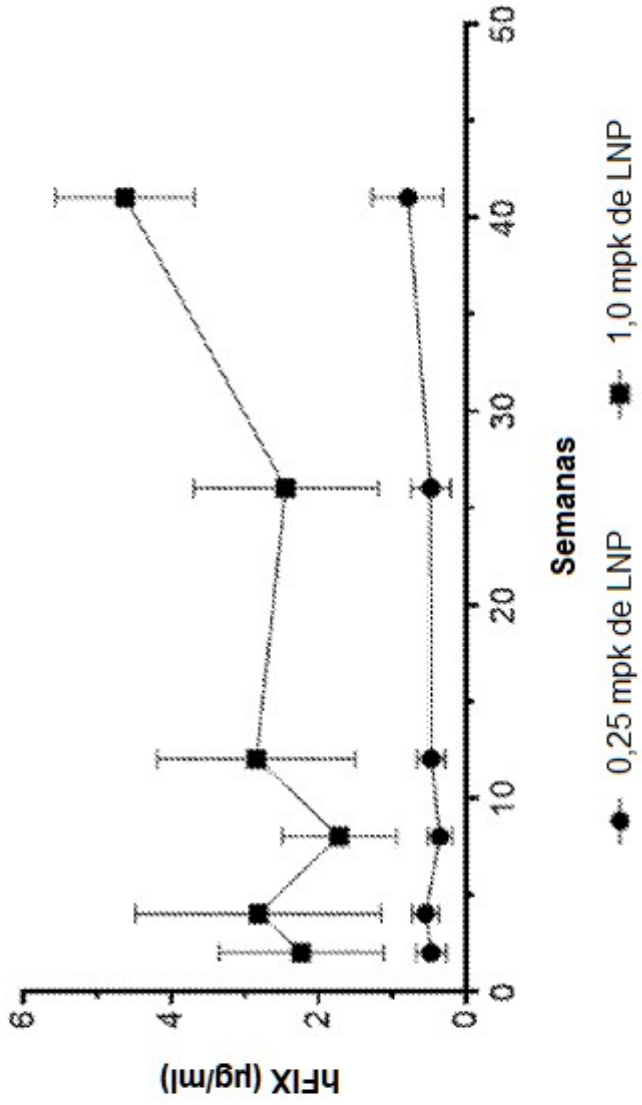

Semana 1
Semana 2
Semana 4
Semana 6
hFIX (ug/ml)
0
2
4
6
8
10
Plasma humano
Plasma deficiente de hFIX
Camundongo WT
PBS
Modelo apenas
Modelo + LNP (1x)
Modelo + LNP (2x)
Modelo + LNP (3x)
ND
ND
ND

FIG. 10

FIG. 11A


hFIX (µg/ml)
0
2
4
6
0
10
20
30
40
50
Semanas
0,25 mpk de LNP
1,0 mpk de LNP

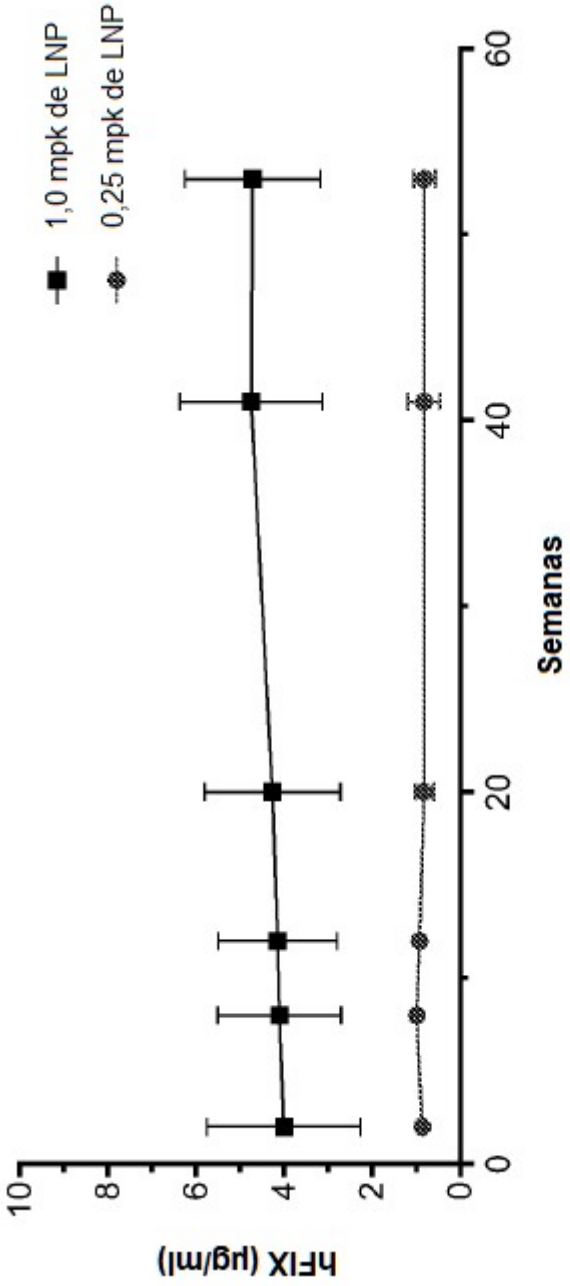

1,0 mpk de LNP
0,25 mpk de LNP
hFIX (µg/ml)
0
2
4
6
8
10
0
20
40
60
Semanas

Figura 11B

FIG. 12A

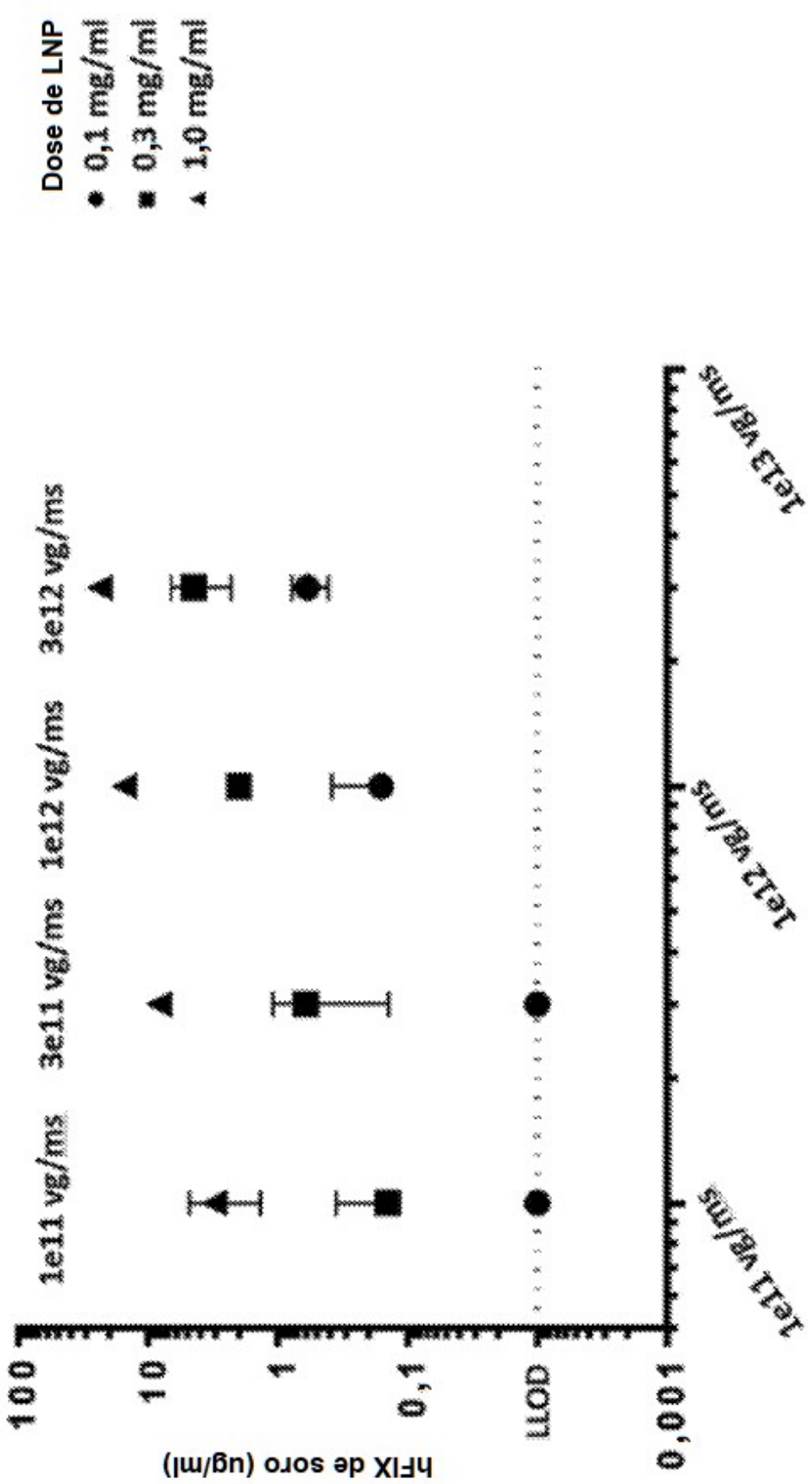
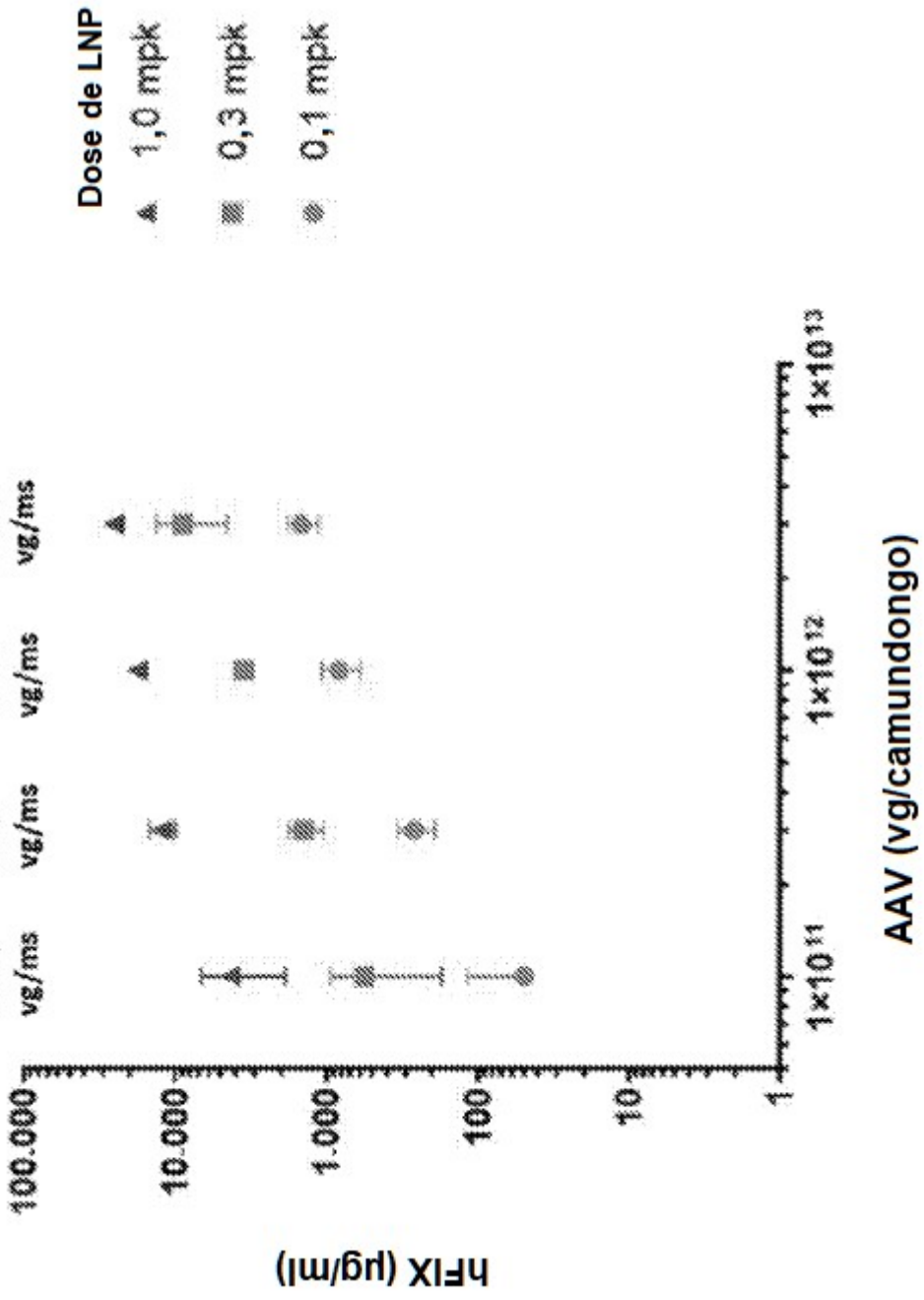

Dose de LNP
0,1 mg/ml
0,3 mg/ml
1,0 mg/ml
1e11 vg/ms
3e11 vg/ms
1e12 vg/ms
3e12 vg/ms
100
10
1
0,1
LLOD
0,01
hFIX de soro (ug/ml)

FIG. 12B


Dose de LNP
1,0 mpk
0,3 mpk
0,1 mpk
1e11 vg/ms
3e11 vg/ms
1e12 vg/ms
3e12 vg/ms
100.000
10.000
1.000
100
10
1
hFIX (µg/ml)
1×10^11
1×10^12
1×10^13
AAV (vg/camundongo)

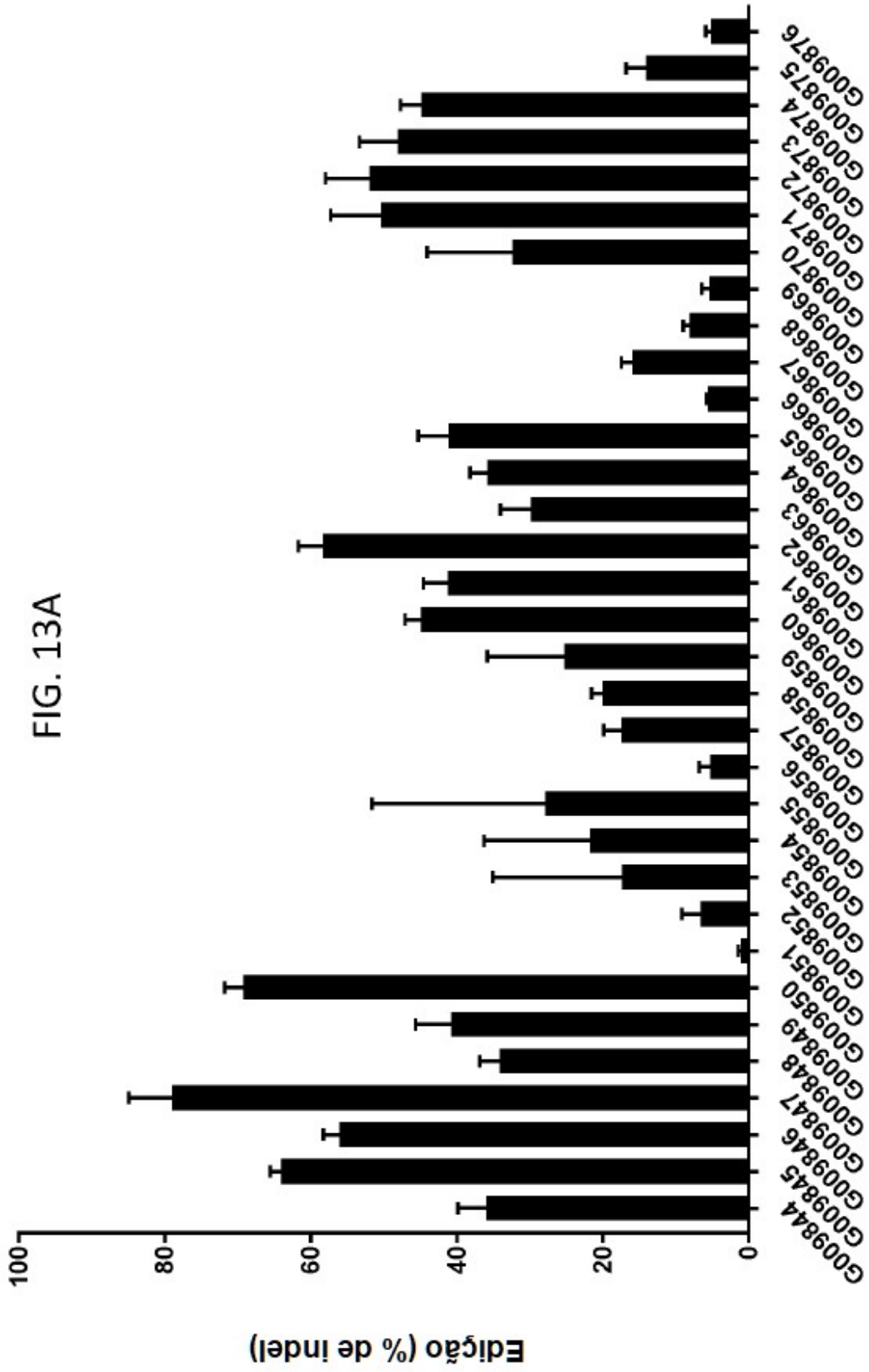

Edição (% de indel)
100
80
60
40
20
0
G009844
G009845
G009846
G009847
G009848
G009849
G009850
G009851
G009852
G009853
G009854
G009855
G009856
G009857
G009858
G009859
G009860
G009861
G009862
G009863
G009864
G009865
G009866
G009867
G009868
G009869
G009870
G009871
G009872
G009873
G009874
G009875
G009876

FIG. 13A

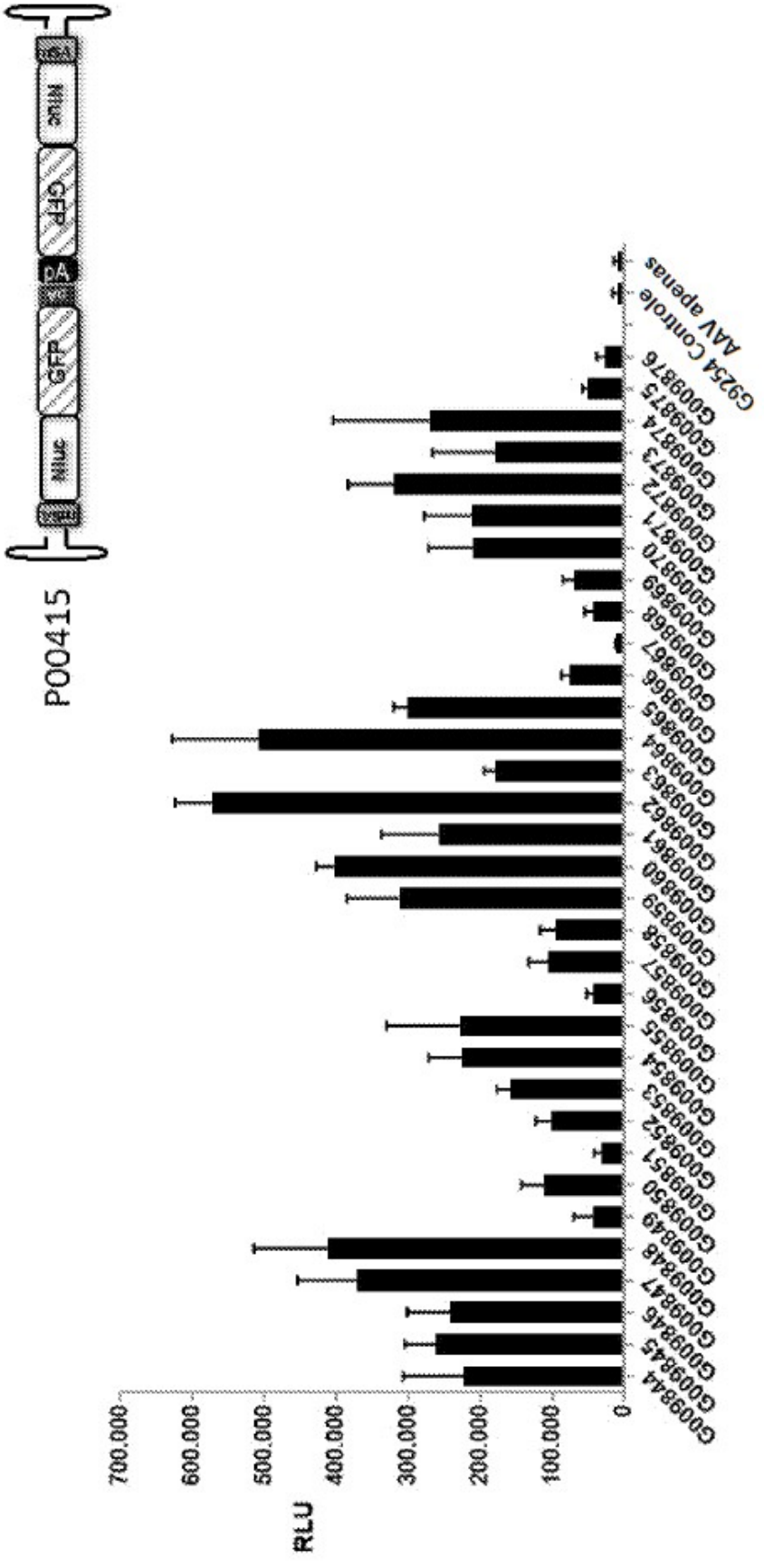

P00415
hRluc
GFP
pA
GFP
mRuc
RLU
700.000
600.000
500.000
400.000
300.000
200.000
100.000
0
G009844
G009845
G009846
G009847
G009848
G009849
G009850
G009851
G009852
G009853
G009854
G009855
G009856
G009857
G009858
G009859
G009860
G009861
G009862
G009863
G009864
G009865
G009866
G009867
G009868
G009869
G009870
G009871
G009872
G009873
G009874
G009875
G009876
G254 Controle
AAV apenas
ID guia

FIG. 13B

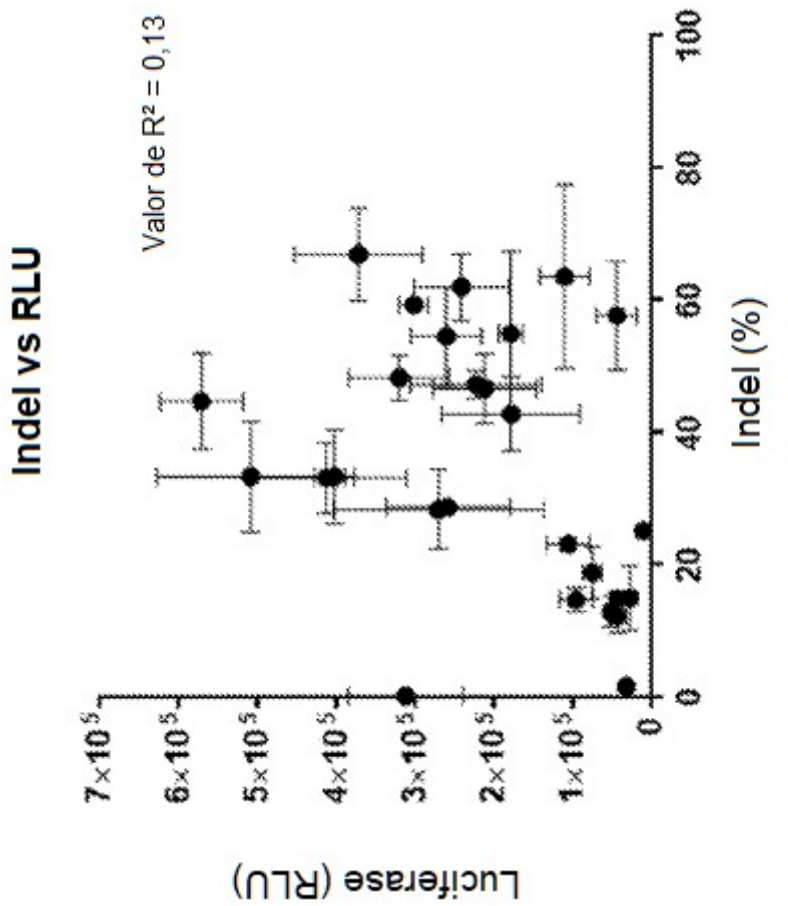

Indel vs RLU
Valor de R² = 0,13
Luciferase (RLU)
7×10⁵
6×10⁵
5×10⁵
4×10⁵
3×10⁵
2×10⁵
1×10⁵
0
0
20
40
60
80
100
Indel (%)

FIG. 13C

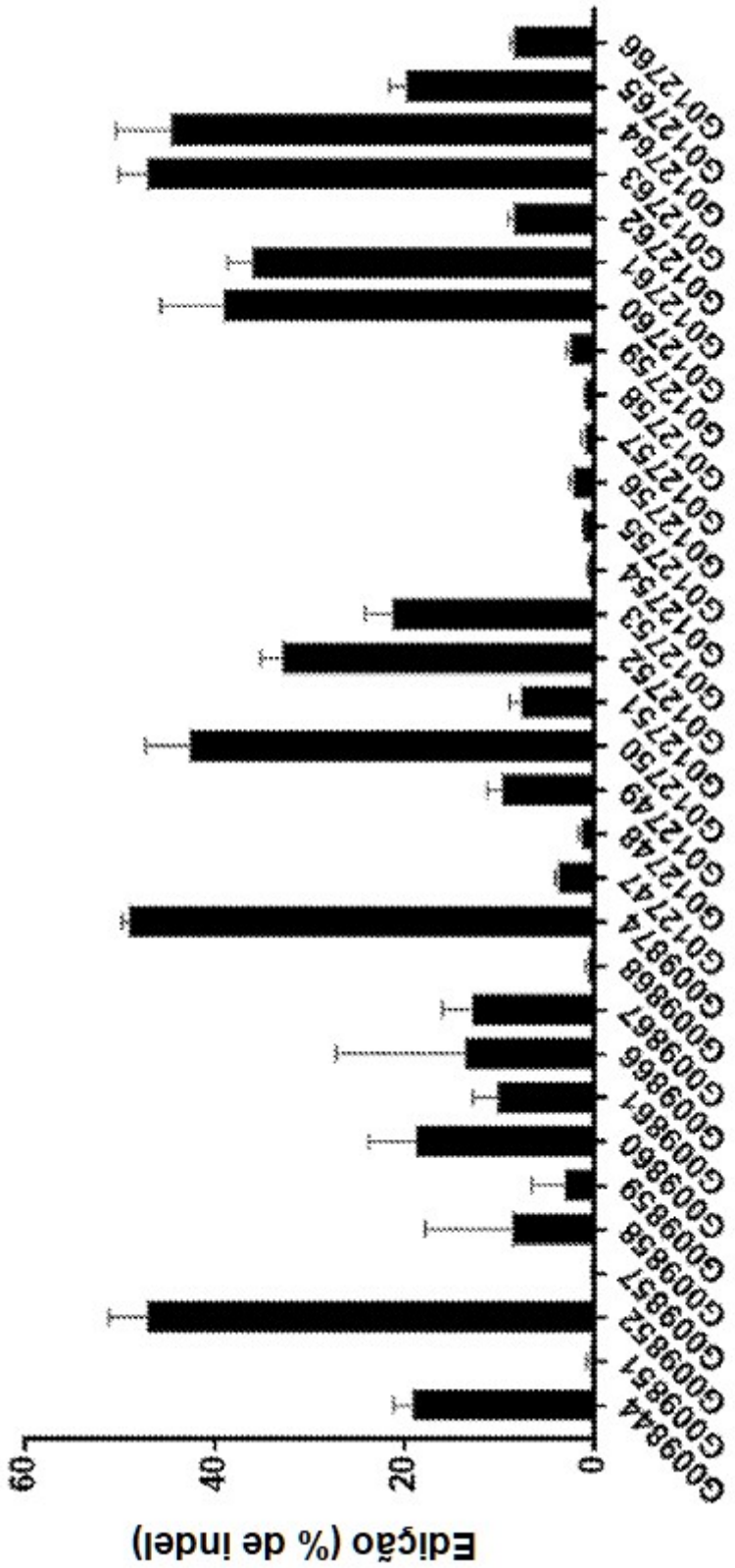

Edição (% de indel)
0
20
40
60
G009844
G009851
G009852
G009857
G009858
G009859
G009860
G009861
G009866
G009867
G009868
G012874
G012747
G012748
G012749
G012750
G012751
G012752
G012753
G012754
G012755
G012756
G012757
G012758
G012759
G012760
G012761
G012762
G012763
G012764
G012765
G012766

FIG. 14A

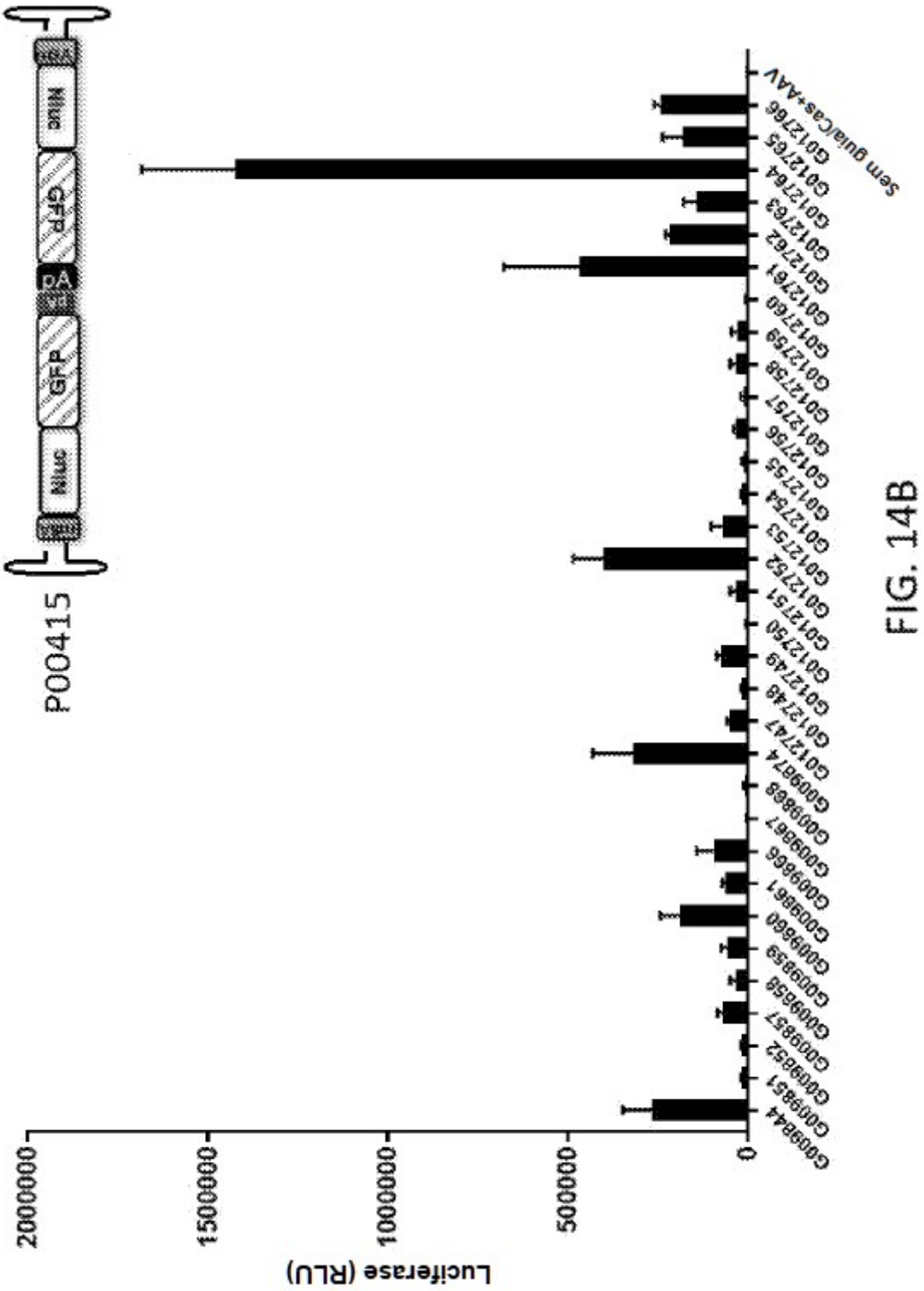

P00415
Nluc
GFP
pA
GFP
Nluc
Luciferase (RLU)
2000000
1500000
1000000
500000
0
Sem guia/Cas+AAV
G012766
G012765
G012764
G012763
G012762
G012761
G012760
G012759
G012758
G012757
G012756
G012755
G012754
G012753
G012752
G012751
G012750
G012749
G012748
G012747
G009874
G009868
G009867
G009866
G009861
G009860
G009859
G009858
G009857
G009852
G009851
G009844

FIG. 14B

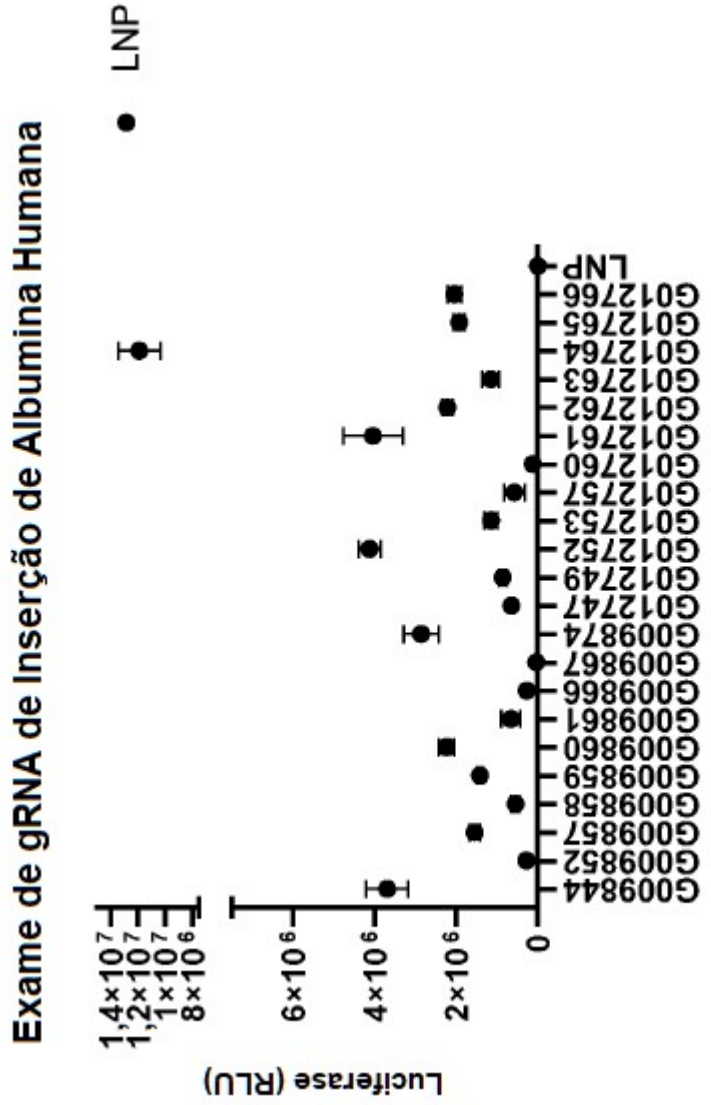

Exame de gRNA de Inserção de Albumina Humana
LNP
Luciferase (RLU)
1,4×10^7
1,2×10^7
1×10^7
8×10^6
6×10^6
4×10^6
2×10^6
0
G009844
G009852
G009857
G009858
G009859
G009860
G009861
G009866
G009867
G009874
G012747
G012749
G012752
G012753
G012757
G012760
G012761
G012762
G012763
G012764
G012765
G012766
LNP

Figura 14C

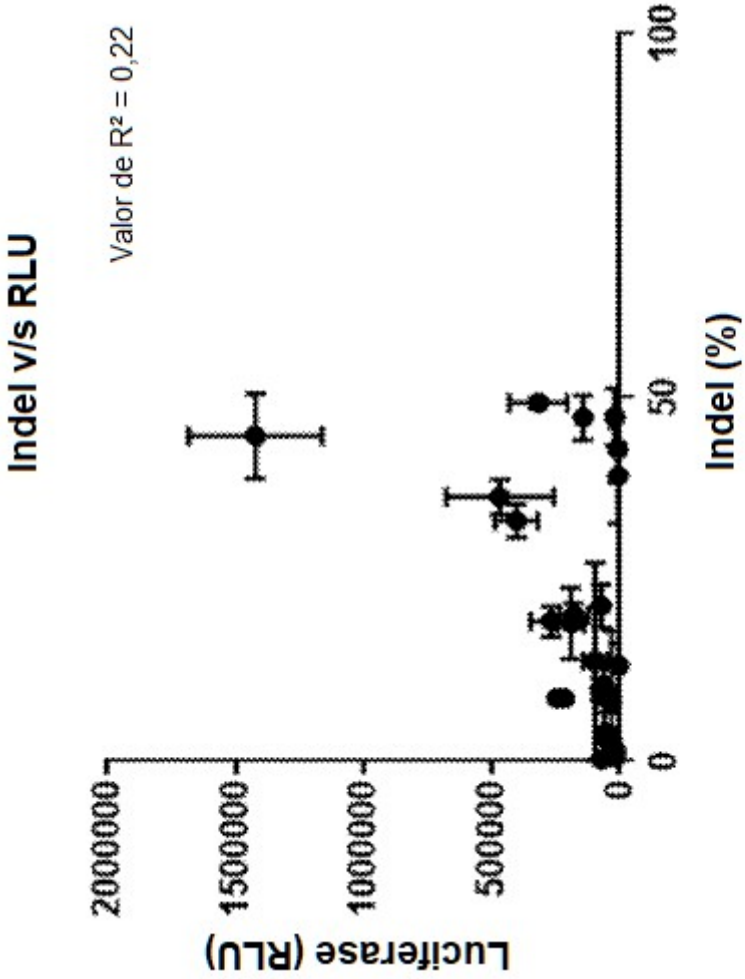
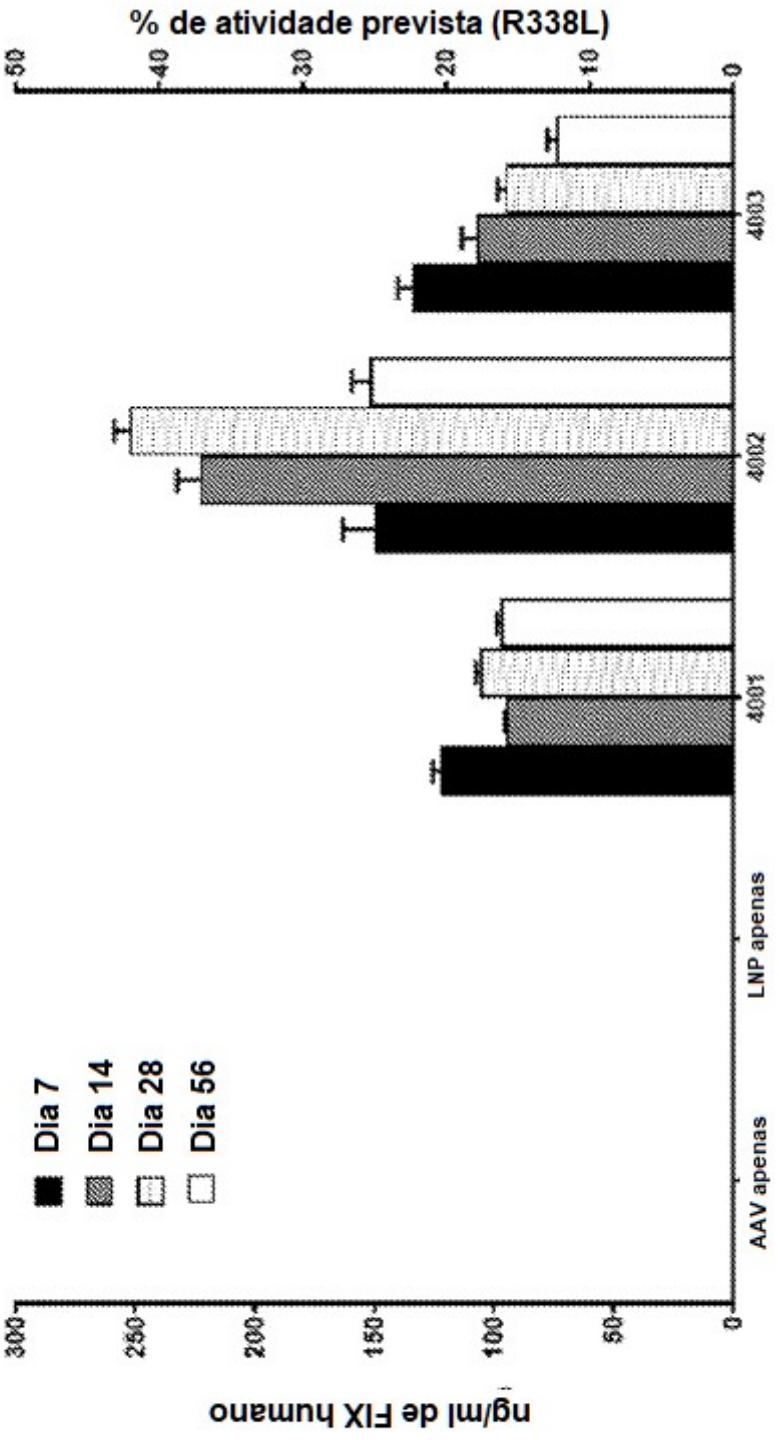

Indel v/s RLU
Valor de R² = 0,22
Luciferase (RLU)
2000000
1500000
1000000
500000
0
0
50
100
Indel (%)

FIG 14D

FIG. 15


% de atividade prevista (R338L)
50
40
30
20
10
0
4003
4002
4001
LNP apenas
AAV apenas
Dia 7
Dia 14
Dia 28
Dia 56
300
250
200
150
100
50
0
ng/ml de FIX humano

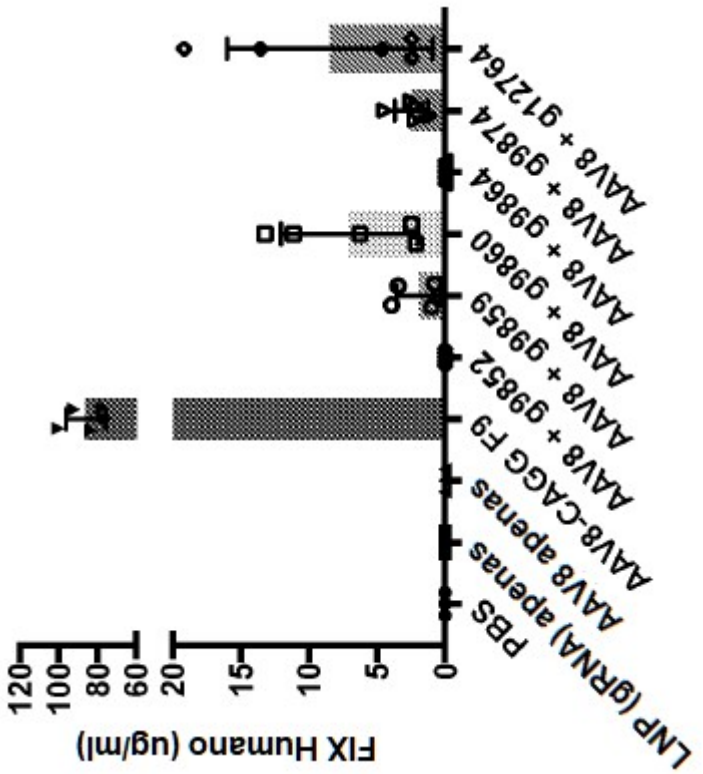

FIX Humano (ug/ml)
120
100
80
60
20
15
10
5
0
PBS
AAV8 apenas
LNP (gRNA) apenas
AAV8-CAGG F9
AAV8 + 99852
AAV8 + 99859
AAV8 + 99860
AAV8 + 99864
AAV8 + 99874
AAV8 + 91276

FIG. 16A

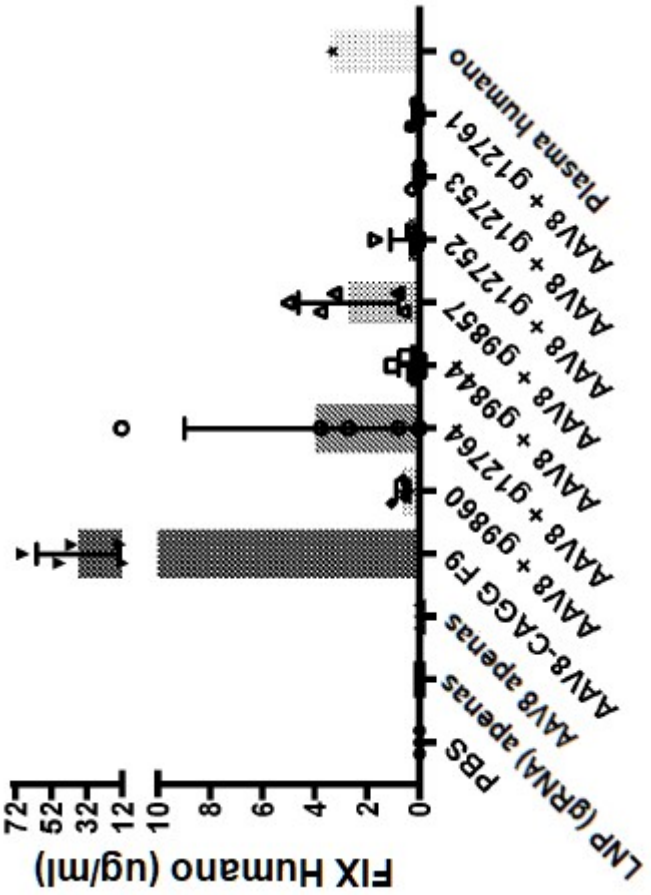

FIX Humano (ug/ml)
72
52
32
12
10
8
6
4
2
0
PBS
AAV8 apenas
LNP (gRNA) apenas
AAV8-CAGG F9
AAV8 + 99860
AAV8 + 91276 4
AAV8 + 99844
AAV8 + 99857
AAV8 + 91276 2
AAV8 + 91276 3
AAV8 + 91276 1
Plasma humano

FIG. 16B

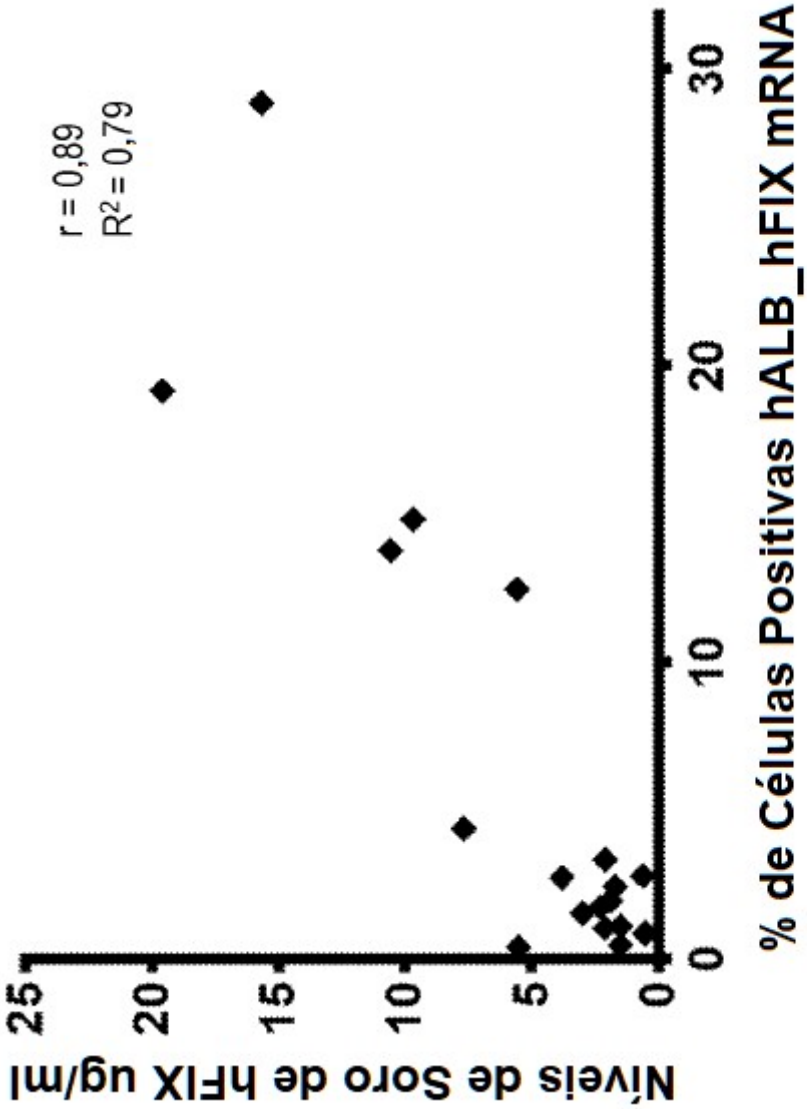

r = 0,89
R² = 0,79
Níveis de Soro de hFIX ug/ml
25
20
15
10
5
0
0
10
20
30
% de Células Positivas hALB_hFIX mRNA

FIG. 17

RESUMO

“COMPOSIÇÕES E MÉTODOS PARA EXPRESSÃO DE TRANSGENE A PARTIR DE UM LÓCUS DE ALBUMINA”

Trata-se de métodos para editar, por exemplo, introduzir um transgene heterólogo, no gene de albumina humano (por exemplo, no íntron 1).

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

Código de Controle

Campo 1



Campo 2



Outras Informações:

- Nome do Arquivo: 213.393-8 Listagem de sequências.txt
- Data de Geração do Código: 15/06/2021
- Hora de Geração do Código: 17:50:24
- Código de Controle:
  - Campo 1: BA9223427ACF028E
  - Campo 2: 05732BA5AF51CE5A