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

(21) BR 112021007323-0 A2

(22) Data do Depósito: 18/10/2019

(43) Data da Publicação Nacional: 27/07/2021

(54) Título: CONSTRUTOS DE ÁCIDO NUCLEICO E MÉTODOS PARA USO

(51) Int. Cl.: C12N 15/11; C12N 15/10; C12N 15/86.

(30) Prioridade Unionista: 18/10/2018 US 62/747,393; 29/04/2019 US 62/840,343.

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

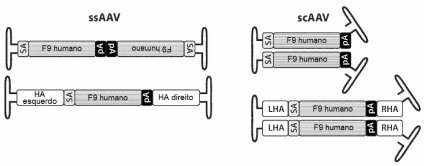
(72) Inventor(es): JONATHAN DOUGLAS FINN; HON-REN HUANG.

(86) Pedido PCT: PCT US2019057084 de 18/10/2019

(87) Publicação PCT: WO 2020/082041 de 23/04/2020

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

(57) Resumo: CONSTRUTOS DE ÁCIDO NUCLEICO E MÉTODOS PARA USO. Trata-se de construtos de ácido nucleico que permitem a inserção e/ou expressão de uma sequência de interesse, como um transgene. Composições e métodos para usar tais construtos para expressão de um polipeptídeo ou agente terapêutico, por exemplo, também são fornecidos.



“CONSTRUTOS DE ÁCIDO NUCLEICO E MÉTODOS PARA USO”

[001] Este pedido reivindica o benefício de prioridade do Pedido Provisório US nº 62/747.393, depositado em 18 de outubro de 2018 e Pedido Provisório US nº 62/840.343, depositado em 29 de abril de 2019. Os relatórios descritivos de cada um dos pedidos supracitados são incorporados ao presente documento a título de referência, em sua totalidade.

[002] A edição de genoma em abordagens de terapia genética surge da ideia que a introdução exógena do material ausente ou material genético comprometido de outro modo pode corrigir uma doença genética. A terapia genética foi reconhecida por seu potencial enorme em como os praticantes abordam e tratam doenças humanas. Em vez de depender de fármacos ou cirurgia, os pacientes com fatores genéticos subjacentes podem ser tratados alvejando-se a causa subjacente. Adicionalmente, alvejando-se a causa genética subjacente, a terapia genética pode ter o potencial de curar de modo eficaz os pacientes. Ainda assim, as aplicações clínicas de abordagens existentes ainda necessitarão de melhora em diversos aspectos.

[003] A presente divulgação fornece construtos de ácido nucleico bidirecional que permitem a inserção e expressão intensificadas de uma sequência de ácido nucleico de interesse, por exemplo, que codifica um agente terapêutico, como um polipeptídeo. Conforme descrito no presente documento, os construtos bidirecionais compreendem 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 como “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, os construtos 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, a bidirecionalidade dos construtos de ácido nucleico permite que o construto seja inserido em qualquer direção (não é limitado à inserção em uma direção) em um sítio de inserção alvo, o que permite a expressão do polipeptídeo de interesse de a) uma sequência de codificação de um segmento, ou 2) um complemento do outro (segundo) segmento, assim intensificando a eficiência de inserção e expressão, conforme exemplificado no presente documento.

BREVE DESCRIÇÃO DOS DESENHOS

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

[005] A Figura 2 mostra vetores sem braços de homologia que não são eficazes em uma linhagem celular de fígado imortalizada (Hepal-6). Um scAAV derivado de plasmídeo P00204 que compreende braços de homologia de 200 bp resultou em expressão detectável de hFIX nessa linhagem celular. O uso de vetores AAV derivados de P00 123 (scAAV sem braços de homologia) e P00147 (construto bidirecional ssAAV sem braços de homologia) não resultou em expressão detectável de hFIX.

[006] 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 de fígado conforme medido por formação de indel de -60% foram detectados em cada grupo de animais tratados com LNPs que compreendem componentes de sistema CRISPR/Cas9. A Figura 3B mostra animais que recebem os vetores ssAAV sem braços de homologia (derivados de P00147) em combinação com tratamento de LNP resultaram no nível

mais alto de expressão de hFIX em soro.

[007] 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 inserção alvejada com vetores derivados de plasmídeos P00350, P00356, P00362 (que têm braços de homologia assimétricos conforme mostrado) e P00147 (construto bidirecional, conforme mostrado na Figura 4B). A Figura 4B compara a inserção em um segundo sítio alvejado com vetores derivados de plasmídeos P00353, P00354 (que têm braços de homologia simétricos, conforme mostrado), e P00147.

[008] As Figuras 5A a 5D mostram resultados de inserção alvejada por três construtos bidirecionais através de 20 sítios alvo em hepatócitos de camundongo primários. A Figura 5A mostra o esquema de cada um dos vetores testados. A Figura 5B mostra a edição conforme medido por 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 (em um sítio alvo específico) não necessariamente resultaram em inserção ou expressão mais eficiente dos transgenes. Os construtos testados resultaram de modo eficaz em expressão de transgene nesse estudo de inserção alvejada. hSA= aceitante de splicing F9 humano; mSA= aceitante de splicing de albumina de camundongo; HiBit= etiqueta para detecção com base em luciferase; pA= sequência sinal poliA; Nluc= repórter nanoluciferase; GFP= repórter fluorescente verde.

[009] A Figura 6 mostra resultados de triagem *in vivo* de inserção alvejada com construtos bidirecionais através de 10 sítios alvo com o uso de ssAAV derivado de P00147. Conforme mostrado, níveis significativos de edição não necessariamente resultam em altos níveis de expressão de transgene.

[010] As Figuras 7A a 7D mostram resultados de triagem *in vivo* de construtos bidirecionais através de 20 sítios alvo com o uso de ssAAV derivado de P00147. A Figura 7A mostra a edição detectada para cada um dos grupos de tratamento para

cada combinação de LNP/vetor testada. A Figura 7B fornece dados de inserção alvejada correspondentes. Os resultados mostram correlação insatisfatória entre edição e inserção/expressão dos construtos bidirecionais (Figura 7B e Figura 7D), e uma correlação positiva entre resultados in vitro e in vivo (Figura 7C).

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

[012] A Figura 9a mostra a durabilidade de expressão de hFIX *in vivo.* A Figura 9b demonstra que a expressão de íntron 1 de albumina foi sustentada.

[013] As Figuras 10A a 10B mostram que a dose de AAV ou LNP variável pode modular a quantidade de expressão de hFIX de íntron 1 do gene de albumina *in vivo.*

[014] As Figuras 11A a 11C mostram resultados de triagem de construtos bidirecionais através de sítios alvo em hepatócitos de cynomolgus primários. A Figura 11A mostra níveis variados de edição conforme medido por formação de indel detectada para cada uma das amostras. A Figura 11B e a Figura 11C mostram que níveis significativos de formação de indel não foram preditivos para inserção ou expressão dos construtos bidirecionais em íntron 1 de albumina.

[015] As Figuras 12A a 12C mostram resultados de triagem de construtos bidirecionais através de sítios alvo em hepatócitos humanos primários. A Figura 12A mostra edição conforme medido por formação de indel detectada para cada uma das amostras. A Figura 12B, a Figura 12C e a Figura 12D 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.

[016] A Figura 13 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.

DESCRIÇÃO DETALHADA

[017] Será feita agora referência em detalhes a certas modalidades da invenção, cujos exemplos são ilustrados nos desenhos anexos. Embora os presentes ensinamentos sejam descritos em combinação com várias modalidades, não se pretende limitar os presentes ensinamentos a essas modalidades. Ao contrário, os presentes ensinamentos abrangem várias alternativas, modificações e equivalentes, conforme será observado por aqueles versados na técnica.

[018] 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 ser observado que, conforme usado neste relatório descritivo e nas reivindicações anexas, a forma singular “um”, “uma”, “o” e “a” incluem referências no plural, a menos que o contexto dite de outro modo. Desse modo, 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.

[019] 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.

[020] 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.

[021] 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.

I. DEFINIÇÕES

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

[023] “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, como 5-metóxiuridina, 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 a 36, Adams et al., 11ª edição, 1992). Ácidos nucleicos podem incluir um ou mais resíduos “abásicos” quando a cadeia principal não inclui base nitrogenosa 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 nucletídeos 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): 13.233-41). 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.

[024] “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, 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 guia” ou “gRNA” se referem 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, como sgRNAs ou dgRNAs, podem incluir RNAs modificados, conforme descrito no presente documento.

[025] 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 ter 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 pelo menos 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 a 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.

[026] 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”, endente-se que a sequência guia pode direcionar um RNA guia para se ligar às fitas de sentido e 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.

[027] 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 al., Cell, 163: 1 a 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 distribuição de um agente de ligação de DNA guiada por RNA *(*por exemplo, uma Cas nuclease, uma Cas9 nuclease, ou uma Cas9 nuclease de S. *pyogenes*) inclui a distribuição do polipeptídeo ou mRNA.

[028] 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, 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, 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.

[029] 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, como uridinas modificadas não contribuem para as diferenças em identidade ou complementaridade entre polinucleotídeos desde que os nucleotídeos relevantes (como timidina, uridina ou uridina modificada) tenham o mesmo complemento (por exemplo, adenosina para toda timidina, uridina ou uridina modificada; outro exemplo é citosina e 5-metilcitosina, 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, como pseudouridina, pseudouridina de N1-metila ou 5-metóxiuridina, 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.

[030] 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.

[031] “mRNA” é usado no presente documento para se referir a um polinucleotídeo que é inteira ou predominantemente RNA ou RNA modificado e 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 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. As bases de um mRNA podem ser bases modificadas, como pseudouridina, N-1-metil-pseudouridina, ou outras bases de ocorrência natural ou ocorrência não natural.

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

[033] 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.

[034] 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.

[035] 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, em um lócus genômico, como um lócus de porto seguro que inclui 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 determinadas modalidades, o gene heterólogo não ocorre naturalmente na célula de recipiente. Por exemplo, o gene heterólogo pode ser heterólogo com relação tanto a seu sítio de inserção quanto com relação a sua célula receptora.

[036] 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.

[037] 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. 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, inserção alvejada do transgene.

[038] 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 de complemento “perfeita” é 3’ GACCTGGCT 5’ (SEQ ID NO: 501), e a sequência de complemento reversa “perfeita” é gravada 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, pelo menos 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.

[039] 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, 1.000 ou mais resíduos de aminoácido.

II. CONSTRUTO DE ÁCIDO NUCLEICO BIDIRECIONAL

[040] São descritos no presente documento os construtos de ácido nucleico bidirecional que facilitam a inserção intensificada, por exemplo, intensificam a inserção produtiva e a expressão de um gene de interesse. Brevemente, vários 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 agente de interesse, por exemplo, um gene heterólogo (a sequência de codificação pode ser denominada no presente documento como “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, por exemplo, um gene heterólogo, ou um segundo transgene. O agente pode ser agente terapêutico, como um polipeptídeo, RNA funcional, mRNA ou semelhantes. O transgene pode codificar um agente, como um polipeptídeo, um RNA funcional, um mRNA ou um fator de transcrição. Em algumas modalidades, uma sequência de codificação codifica um agente terapêutico, como um polipeptídeo ou RNA funcional. 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.

[041] 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 splicing e uma segunda sequência de codificação, em que o complemento codifica um gene heterólogo na outra orientação, também ligado a um aceitante de splicing.

[042] Conforme usado no presente documento, tal construto é, por vezes, denominado um “construto doador/modelo”. Em algumas modalidades, o construto é um construto de DNA. Métodos para projetar e produzir várias modificações funcionais/estruturais em construtos doadores são conhecidos na técnica. Em algumas modalidades, o construto pode compreender qualquer uma ou mais dentre uma sequência de cauda de poliadenilação, uma sequência sinal de poliadenilação, sítio de aceitante de splicing ou marcador selecionável. Em algumas modalidades, a sequência de cauda de poliadenilação é codificada, por exemplo, como uma faixa “poli-A”, na extremidade 3’ da sequência de codificação.

[043] 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 b) 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, 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).

[044] 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).

[045] 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 de codificação 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).

[046] Em algumas modalidades, a sequência de codificação que codifica o Polipeptídeo A no primeiro segmento é menos do 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 a sequência de codificação (1) e o complemento reverso da sequência de codificação (2) possuam 100% ou menos do 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.

[047] 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.

[048] 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).

[049] 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 os mesmos, 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 adição de um aminoácido em uma terminação 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 sinal, sequência de marcação *(*por exemplo, HiBit), ou sequência funcional heteróloga (por exemplo, sequência de localização nuclear (NLS) ou autoclivagem) ligada ao polipeptídeo. Uma sequência de codificação que codifica um polipeptídeo pode compreender, opcionalmente, sequências que codificam uma ou mais sequências de peptídeo sinal de terminal amino. Cada uma dessas sequências adicionais pode ser igual ou diferente no primeiro segmento e no segundo segmento do construto.

[050] 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 ativa) 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.

[051] Em algumas modalidades, o polipeptídeo é um polipeptídeo intracelular. Em algumas modalidades, o polipeptídeo é um em que sua função é normalmente efetuada (por exemplo, funcionalmente ativa) 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.

[052] 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 do mesmo. 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 com um ou mais outros tipos de tecido. Em algumas modalidades, o polipeptídeo é uma proteína não proveniente de fígado. Em algumas modalidades, o polipeptídeo inclui, mas sem limitação, Fator IX e variantes do mesmo.

[053] 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, 1.000, 1.500, 2.000 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.

[054] 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, o construto de ácido nucleico bidirecional divulgado no presente documento é um construto doador independente 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. 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 splicing a montante de um transgene. Em determinadas modalidades, o aceitante de splicing é compatível com a sequência doadora de splicing do sítio de porto seguro da célula hospedeira, por exemplo, o doador de splicing de íntron 1 de um gene de albumina humana.

[055] Em algumas modalidades, a composição descrita no presente documento compreende um ou mais sítios de entrada de ribossomo interno (IRES). Primeiro identificado como um RNA de picornavírus de recurso, IRES tem uma função importante na iniciação de 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 não 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 sinal de albumina).

Exemplos de sequências de IRES que podem ser utilizadas incluem, sem limitação, aqueles de picornavírus (por exemplo, FMDV), vírus de pragas (CFFV), poliovírus (PV), vírus de encefalomiocardite (ECMV), vírus de doença de pés e boca (FMDV), vírus da hepatite C (HCV), vírus de febre suína clássica (CSFV), vírus da leucemia murina (MLV), vírus de imunodeficiência símia (SIV) ou vírus Cricket Paralysis (CrPV).

[056] Em algumas modalidades, o construto de ácido nucleico compreende uma sequência que codifica um peptídeo de autoclivagem, como sequência 2A ou uma sequência do tipo 2A. 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 usada 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

[057] *(*isto é, sequência sinal de albumina endógena). Como 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 na proteína A, proteína B e no peptídeo 2A. A proteína A e a proteína B podem ser os polipeptídeos iguais ou diferentes de interesse.

[058] Em algumas modalidades, um ou ambos dentre o primeiro e o segundo segmentos compreendem uma sequência de cauda de poliadenilação e/ou uma sequência 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. Em algumas modalidades, uma sequência de cauda de poliadenilação é fornecida cotranscricionalmente como um resultado de uma sequência sinal de poliadenilação que é codificada na extremidade 3’ ou próxima à mesma, do primeiro e/ou do segundo segmento. Em algumas modalidades, uma 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. Métodos para projetar uma sequência de cauda de poliadenilação e/ou sequência sinal de poliadenilação adequada são bem conhecidos na técnica. Sequências de aceitante de splicing adequadas são divulgadas e exemplificadas no presente documento, que incluem sítios de aceitante de splicing de albumina de camundongo e FIX humano. Em algumas modalidades, a sequência sinal de poliadenilação AAUAAA (SEQ ID NO: 800) é normalmente usada em sistemas de mamífero, embora variantes, 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): 1.770-82, 2011. Em algumas modalidades, uma sequência de cauda poliA é incluída.

[059] Em algumas modalidades, os construtos divulgados no presente documento podem ser DNA ou RNA, de fita simples, de 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 US nº 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 as extremidades. Consultar, por exemplo, Chang et al. (1987) *Proc. Natl. Acad. Sci. EUA* 84:4.959-4.963; 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, como, por exemplo, fosforotioatos, fosforamidatos, e resíduos de ribose

de O-metila ou desoxirribose.

[060] 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 uma sequência sinal que codifica um peptídeo sinal e operacionalmente ligada à mesma.

[061] Em algumas modalidades, o construto de ácido nucleico funciona na inserção independente de homologia de um ácido nucleico que codifica uma proteína heteróloga. 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. Por exemplo, os construtos podem ser DNA de fita simples ou dupla. Em algumas modalidades, o ácido nucleico pode ser modificado (por exemplo, com o uso de análogos de nucleosídeo), conforme descrito no presente documento.

[062] Em algumas modalidades, os construtos divulgados no presente documento compreendem um sítio de aceitante de splicing 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 splicing compreende

NAG. Em modalidades adicionais, o sítio de aceitante de splicing consiste em NAG. Em algumas modalidades, o aceitante de splicing é um aceitante de splicing de albumina, por exemplo, um aceitante de splicing de albumina usado no splicing em conjunto de éxons 1 e 2 de albumina. Em algumas modalidades, o aceitante de splicing é derivado do gene de albumina humana. Em algumas modalidades, o aceitante de splicing é derivado do gene de albumina de camundongo. Em algumas modalidades, o aceitante de splicing é um aceitante de splicing de F9 (ou “FIX”), por exemplo, o aceitante de splicing de F9 usado no splicing em conjunto de éxons 1 e 2 de F9. Em algumas modalidades, o aceitante de splicing é derivado do gene de F9 humano. Em algumas modalidades, o aceitante de splicing é derivado do gene de F9 de camundongo. Sítios de aceitante de splicing adequados adicionais úteis em eucariontes, que incluem aceitantes de splicing artificiais, são conhecidos e podem ser derivados da técnica. Consultar, por exemplo, Shapiro, et al., 1987, Nucleic Acids Res., 15, 7.155-7.174, Burset, et al, 2001, Nucleic Acids Res., 29, 255-259.

[063] Em algumas modalidades, os construtos divulgados no presente documento podem ser modificados em qualquer uma ou ambas as extremidades para incluir um ou mais recursos estruturais adequados conforme necessário, e/ou para conferir um ou mais benefícios funcionais. Por exemplo, modificações estruturais podem variar dependendo do método (ou métodos) usado para distribuir os construtos divulgados no presente documento a uma célula hospedeira - por exemplo, uso de distribuição de vetor viral ou acondicionamento em nanopartículas lipídicas para distribuição. Tais modificações incluem, sem limitação, por exemplo, estruturas de terminal, como estruturas de repetições de terminal invertidas (ITR), grampo, laços e outras, como toroide. Em algumas modalidades, os construtos divulgados no presente documento compreendem uma, duas ou três ITRs. Em algumas modalidades, os construtos divulgados no presente documento compreendem não mais que duas ITRs. Vários métodos de modificações estruturais são conhecidos na técnica.

[064] Em algumas modalidades, uma ou ambas as extremidades do construto podem ser protegidas (por exemplo, da degradação exonucleolítica) através de métodos conhecidos 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 as extremidades. Consultar, por exemplo, Chang et al. (1987) *Proc. Natl. Acad. Set. EUA* 84:4.959-4.963; Nehls et al. (1996) *Science* 272:886-889. Métodos adicionais para proteger os construtos da degradação incluem, porém, sem limitação, adição de grupo (ou grupos) amino terminal e o uso de ligações internucleotídeos modificadas, como, por exemplo, fosforotioatos, fosforamidatos, e resíduos de ribose de O-metila ou desoxirribose.

[065] Em algumas modalidades, os construtos divulgados no presente documento podem ser introduzidos em uma célula como parte de um vetor que tem sequências adicionais, como, por exemplo, origens de replicação, promotores e genes que codificam resistência antibiótica. Um construto pode omitir elementos virais. Em algumas modalidades, os construtos podem ser introduzidos como ácido nucleico nu, como ácido nucleico complexado com um agente, como um lipossoma, polímero ou poloxâmero, ou podem ser distribuídos por vetores virais (por exemplo, adenovírus, AAV, herpesvírus, retrovírus, lentivírus).

[066] 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.

[067] 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):3.335-44.

III. SISTEMA DE EDIÇÃO DE GENE

[068] Vários sistemas de edição de gene conhecidos podem ser usados na prática da presente divulgação, que incluem, por exemplo, um sistema de CRISPR/Cas; sistema de nuclease de dedo de zinco (ZFN); e sistema de nuclease efetora semelhante ao ativador de transcrição (TALEN). Geralmente, esses métodos podem envolver o uso de sistemas de clivagem modificados geneticamente para induzir uma quebra de fita dupla (DSB) ou um corte (por exemplo, uma quebra de fita simples, ou SSB) em uma sequência de DNA-alvo. Clivagem ou corte pode ocorrer através do uso de nucleases específicas, como ZFN modificada geneticamente, TALENs ou com o uso do sistema de CRISPR/Cas com um RNA guia modificado geneticamente para guiar clivagem ou corte específico de uma sequência de DNA-alvo. Adicionalmente, nucleases alvejadas têm sido desenvolvidas, e nucleases adicionais estão em desenvolvimento, por exemplo, com base no sistema Argonauta (por exemplo, de T. thermophilus, conhecido como 'TtAgo', consultar Swarts et al (2014) Nature 507(7491): 258 a 261), que também pode ter o potencial para usos em edição de genoma e terapia de gene.

[069] Em algumas modalidades, um sistema CRISPR/Cas pode ser usado para criar um sítio de inserção em um lócus desejável em um genoma hospedeiro, em que, no sítio, um construto bidirecional divulgado no presente documento pode ser inserido para expressar um ou mais polipeptídeos de interesse. Métodos para projetar RNAs guia adequados que alvejam qualquer lócus desejável de um genoma hospedeiro para inserção são bem conhecidos na técnica. Um construto bidirecional que compreende um transgene pode ser heterólogo em relação a seu sítio de inserção, por exemplo, inserção de um transgene heterólogo em um lócus "de porto seguro". Um construto bidirecional que compreende um transgene pode ser não heterólogo em relação a seu sítio de inserção, por exemplo, inserção de um transgene

de tipo selvagem em seu lócus endógeno.

[070] Um lócus de “porto seguro” é um lócus dentro do genoma no qual um ácido nucleico exógeno pode ser inserido 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. 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 expressão de um ácido nucleico exógeno (por exemplo, um gene exógeno) sem efeitos negativos significativos na célula hospedeira ou população de células, 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 um controle ou população de células. O porto seguro pode estar dentro de um gene de albumina, como um gene de albumina humano. O porto seguro pode estar dentro de uma região de íntron 1 de albumina, por exemplo, íntron 1 de albumina humano. O porto seguro pode ser um porto seguro humano, por exemplo, para um tecido hepático ou célula hospedeira de hepatócito. Exemplos não limitantes de lócus de porto seguro que são alvejados por nuclease (ou nucleases) incluem CCR5, HPRT, AAVS1, Rosa, albumina, AAVS1 (PPP1 R12C), AngptiS, ApoC3, ASGR2, FIX (F9), G6PC, Gys2, HGD, Lp(a), Pcsk9, SERPINA1, TF e TTR. Consultar, por exemplo, as Patentes US nº 7.951.925 e 8.110.379; Publicações US nº 2008/0159996; 2010/00218264; 2012/0017290; 2011/0265198; 2013/0137104; 2013/0122591; 2013/0177983;2013/0177960; e WO 2017093804. Conforme exemplificado no presente documento, em algumas modalidades, RNAs guia podem ser projetados para alvejar um lócus de albumina humana ou de camundongo (por exemplo, íntron 1). Exemplos de RNAs guia exemplificados no presente documento são mostrados

nas Tabelas 5 a 10. Será observado que quaisquer outros lócus podem ser alvejados para inserção de um construto bidirecional que compreende um transgene de acordo com os presentes métodos.

[071] Em algumas modalidades, o gene heterólogo pode ser inserido em um lócus de porto seguro e usar a sequência sinal endógena do lócus de porto seguro, por exemplo, a sequência sinal de albumina codificada pelo 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 sinal de éxon de albumina humana 1.

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

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

[074] Em algumas modalidades, o gene pode compreender sua própria sequência sinal e IRES, pode ser inserido no lócus de porto seguro, e não usa a sequência sinal endógena do lócus de porto seguro. Por exemplo, uma sequência de codificação que compreende sua sequência sinal nativa e uma sequência de IRES pode ser inserida em íntron de albumina humana 1 de modo que não se funda à sequência sinal de albumina humana codificada pelo éxon 1. Nessas modalidades, a

proteína é traduzida do sítio de IRES e não é quimérica *(*por exemplo*,* peptídeo 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.

[075] Em algumas modalidades, o gene pode ser inserido no lócus de porto seguro e pode compreender um IRES e não usa qualquer sequência sinal. Por exemplo, uma sequência de codificação que compreende uma sequência de IRES e nenhuma sequência sinal nativa pode ser inserida em íntron de albumina humana 1 de modo que não se funda à sequência sinal de albumina humana codificada pelo éxon 1. Em algumas modalidades, as proteínas são traduzidas do sítio de IRES sem qualquer sequência sinal. Em algumas modalidades, a proteína não é secretada e/ou transportada de modo extracelular.

[076] Também será observado que um RNA guia para uma Cas nuclease, como uma Cas9 nuclease que pode ser usada nos presentes métodos pode incluir qualquer uma das diversas variações e modificações conhecidas (por exemplo, modificações químicas), incluindo a presença de um ou mais componentes de ocorrência não natural e/ou de ocorrência natural ou configurações que são usadas em vez de ou além dos resíduos A, G, C e U canônicos. Por exemplo, cada uma das sequências guia exemplificadas no presente documento (Tabelas 5-10) podem compreender adicionalmente nucleotídeos adicionais para formar um crRNA, RNA guia e/ou sgRNA, por exemplo, de um sistema SpyCas9 CRISPR/Cas. Por exemplo, cada uma das sequências guia exemplificadas no presente documento (Tabelas 5-10) pode compreender adicionalmente nucleotídeos adicionais para formar um crRNA ou sgRNA com a sequência de nucleotídeo exemplificativo seguinte à sequência guia em sua extremidade 3': GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 200) na orientação 5' para 3'. No caso de um sgRNA, as sequências guia, como as sequências guia listadas nas Tabelas 5-10 podem compreender adicionalmente nucleotídeos

adicionais para formar um sgRNA, por exemplo, com a sequência de nucleotídeo exemplificativa seguinte (uma sequência guia SpyCas9) seguinte à extremidade 3' da sequência guia: GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 201) ou GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 202) na orientação 5' para 3'.

[077] O RNA guia pode compreender opcionalmente um trRNA. Em cada modalidade de composição e método descrita no presente documento, um 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. 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 em qualquer uma dentre as Tabelas 5-10, 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.

[078] 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.

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

[080] Em algumas modalidades, a sequência de RNA guia compreende uma sequência que é complementar a uma sequência dentro de uma região genômica selecionada dentre as tabelas no presente documento 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 dentre as Tabelas 5-10. 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 dentre as Tabelas 5-10.

[081] 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).

[082] Métodos para usar vários agentes de ligação de DNA guiados por RNA, por exemplo, uma nuclease, como uma Cas nuclease, por exemplo, Cas9, também são bem conhecidos na técnica. Embora o uso de um ácido nucleico bidirecional com um sistema de CRISPR/Cas seja exemplificado no presente documento, será recomendado que as variações adequadas para o sistema também possam ser

usadas. Será reconhecido que, dependendo do contexto, o agente de ligação de DNA guiado por RNA pode ser fornecido como um ácido nucleico (por exemplo, DNA ou mRNA) ou como uma proteína. Em algumas modalidades, o presente método pode ser praticado em uma célula hospedeira que já compreende e/ou expressa um agente de ligação de DNA guiado por RNA.

[083] Em algumas modalidades, o agente de ligação de DNA guiado por RNA, 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, 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 de CRISPR de tipo II de S. *pyogenes, S. aureus,* e outros procariotas (consultar, por exemplo, a lista no próximo parágrafo), e versões variantes ou mutantes (por exemplo, modificadas geneticamente, de ocorrência não natural, de ocorrência natural, ou outra variante) dos mesmos. Consultar, por exemplo, os documentos nº US2016/0312198 Al; US 2016/0312199 A1.

[084] 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.*

[085] 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*.

[086] 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.

[087] 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.

[088] 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, como Fok1. Em algumas modalidades, uma Cas nuclease pode ser uma nuclease modificada.

[089] 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.

[090] 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, o documento de Patente US nº 8.889.356 para a discussão de Cas nickases e alterações de domínio catalítico exemplificativo. Em algumas modalidades, uma Cas nickase, como uma Cas9 nickase tem um domínio RuvC ou HNH inativado.

[091] 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.

[092] 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 (CPF1FRATN)).

[093] 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.

[094] 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).

[095] 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 a 10 NLS (ou NLSs). Em algumas modalidades, o agente de ligação de DNA guiado por RNA pode ser fundido com 1 a 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, 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, 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.

[096] Conforme observado acima, o agente de ligação de DNA guiado por RNA pode ser um ácido nucleico que codifica polipeptídeos de ligação de DNA guiado por RNA. Em algumas modalidades, um agente de ligação de DNA guiado por RNA

compreende um mRNA que compreende um mRNA que compreende um quadro de leitura aberto (ORF) que codifica um agente de ligação de DNA guiado por RNA, 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, como uma Cas nuclease, é fornecido, usado ou administrado. 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 de sítio específico. 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.

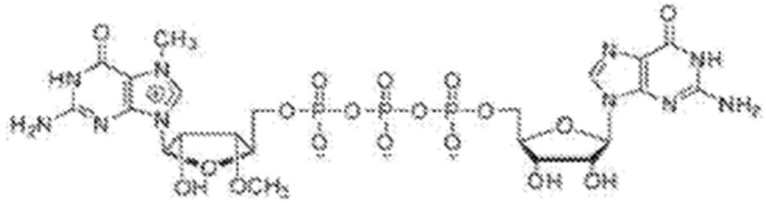
[097] ORFs de Cas9, que incluem ORFs de Cas9 modificados, são fornecidos no presente documento e são conhecidos na técnica. Como um exemplo, o ORF de Cas9 pode ser de códon otimizado, de modo que a sequência de codificação inclua um ou mais códons alternativos para um ou mais aminoácidos. Um “códon alternativo”, 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.

[098] 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-metóxiuridina, 5-iodouridina ou uma combinação das mesmas. Em algumas modalidades, a uridina modificada é 5-metóxiuridina. 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-metóxiuridina. Em algumas modalidades, a uridina modificada é uma combinação de N1-metil-pseudouridina e 5-metóxiuridina. 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-metóxiuridina.

[099] Em algumas modalidades, um mRNA divulgado no presente documento compreende um cap de 5', 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

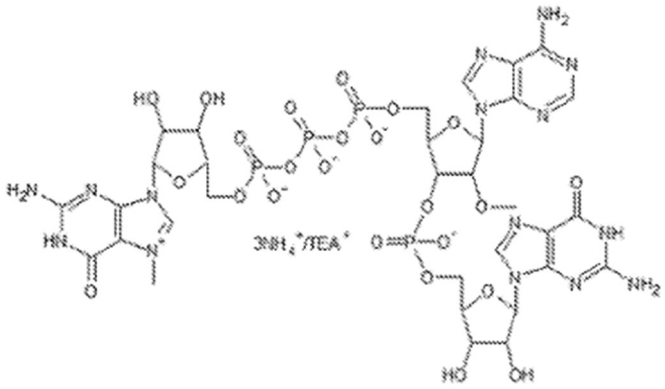
compreendem um 2’-metóxi e uma 2’-hidroxila, respectivamente. Em Cap2, as riboses do primeiro e do segundo nucleotídeos próximos ao cap do mRNA compreendem ambas uma 2’-metóxi. Consultar, por exemplo, Katibah et al. (2014) *Proc Natl Acad Sci USA* 111(33): 12.025-30; Abbas et al. (2017) *Proc Natl Acad Sci USA* 114(11):E2106-E2115. A maioria dos mRNAs eucarióticos superiores endógenos, que incluem mRNAs de mamíferos, como mRNAs humanos, compreendem Cap1 ou Cap2. Cap0 e outras estruturas de cap que se diferem de Cap 1 e Cap2 podem ser imunogênicos em mamíferos, como seres humanos, devido ao reconhecimento como “não próprio” pelos componentes do sistema imunológico inato, como IFIT-1 e IFIT-5, que pode resultar em níveis de citocina elevados que incluem interferon tipo I. Componentes do sistema imunológico inato, 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.

[0100] 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: 1.486-1.495. A estrutura de ARCA é mostrada abaixo.



[0101] CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies Cat. nº N-7113) ou CleanCap™ GG (m7G(5')ppp(5')(2'OMeG)pG; TriLink

Biotechnologies Cat. nº 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.



[0102] 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. Set. EUA* 87, 4.023-4.027; Mao, X. e Shuman, S. (1994) J. *Biol. Chem.* 269, 24.472-24.479.

[0103] 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.

IV. MÉTODOS DE DISTRIBUIÇÃO

[0104] Os construtos de ácido nucleico divulgados no presente documento podem ser distribuídos para uma célula hospedeira ou indivíduo, in vivo ou ex vivo,

com o uso de vários métodos conhecidos e adequados disponíveis na técnica. Os construtos de ácido nucleico podem ser distribuídos junto com componentes de um sistema de edição de gene adequado (por exemplo, agente de ligação de DNA guiado por RNA, como uma Cas nuclease com seu RNA guia correspondente) conforme descrito no presente documento.

[0105] Os métodos de distribuição de gene com base viral e não viral convencional podem ser usados para introduzir os construtos divulgados no presente documento e componentes do sistema de edição de gene em células (por exemplo, células de mamíferos) e tecidos alvo. Conforme fornecido adicionalmente no presente documento, os sistemas de distribuição de vetor não viral incluem ácidos nucleicos, como vetores não virais, vetores de plasmídeo e, por exemplo, ácido nucleico complexado com um veículo de distribuição, como um lipossoma, nanopartícula lipídica (LNP) ou poloxâmero. Sistemas de distribuição de vetor viral incluem vírus de DNA e RNA.

[0106] Métodos e composições para distribuição 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 distribuição de ácidos nucleicos.

[0107] Sistemas de distribuição 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 US nº 6.008.336). A lipofecção é descrita, por exemplo, nos documentos de Patente US nº 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, como complexos imunolipídicos, é bem conhecida na técnica, e conforme descrito no presente documento.

[0108] Vários sistemas de distribuição (por exemplo, vetores, lipossomos, LNPs) que contêm os construtos bidirecionais e/ou componentes de edição de gene (por exemplo, RNA guia e Cas) também podem ser administrados a um organismo para distribuição a 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.

[0109] Em determinadas modalidades, a presente divulgação fornece vetores que compreendem os construtos de ácido nucleico bidirecional divulgados no presente documento para distribuição a uma célula hospedeira. Em determinadas modalidades, os componentes do sistema de edição de gene (por exemplo, agente de ligação de DNA guiado por RNA e RNA guia) também são distribuídos a uma célula hospedeira como parte de um vetor. Em determinadas modalidades, os vetores virais também podem ser usados para distribuir quaisquer um ou mais dentre um construto de ácido nucleico bidirecional, RNA guia e/ou agente de ligação de DNA guiado por RNA a uma célula hospedeira.

[0110] Em algumas modalidades, são fornecidas no presente documento composições e métodos para distribuir o construto de ácido nucleico bidirecional divulgado no presente documento a uma célula hospedeira ou indivíduo, em que o construto é parte de um sistema de vetor conforme descrito no presente documento. Em algumas modalidades, o sistema de vetor compreende componentes adicionais, como componentes de um sistema de edição de gene (por exemplo, RNA guia e/ou um agente de ligação de DNA guiado por RNA).

[0111] Em algumas modalidades, uma composição de vetor que compreende o construto de ácido nucleico bidirecional divulgado no presente documento é fornecida. Em algumas modalidades, a composição compreende adicionalmente componentes de um sistema de edição de gene (por exemplo, RNA guia e/ou um agente de ligação de DNA guiado por RNA).

[0112] Em algumas modalidades, o vetor pode ser circular. Em outras modalidades, o vetor pode ser linear. Em algumas modalidades, o vetor pode ser distribuído por meio de uma nanopartícula lipídica, lipossoma, nanopartícula não lipídica ou capsídeo viral. Vetores exemplificativos sem limitação incluem plasmídeos, fagomídeos, cosmídeos, cromossomos artificiais, minicromossomos, transposões, vetores virais e vetores de expressão.

[0113] Em algumas modalidades, o sistema de vetor pode ter capacidade para acionar expressão de um ou mais componentes de nuclease em uma célula. Em algumas modalidades, o construto bidirecional, opcionalmente como parte de um sistema de vetor, pode compreender um promotor que tem capacidade para acionar a expressão de uma sequência de codificação em uma célula. Em algumas modalidades, a célula pode ser uma célula eucariótica, 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. Em algumas

modalidades, o vetor não compreende um promotor que aciona a expressão de uma ou mais sequências de codificação em uma célula (por exemplo, a expressão da sequência de codificação, uma vez inserida em um lócus endógeno alvo, é acionada por um promotor endógeno).

[0114] 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 (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 (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.

[0115] O uso de sistemas com base viral de RNA ou DNA para a distribuição de ácidos nucleicos toma a vantagem de processos altamente evoluídos para alvejar um vírus a células específicas no corpo e tráfego da carga útil viral para o núcleo. Os vetores virais podem ser administrados diretamente a um indivíduo (in vivo) ou podem ser usados para tratar células in vitro. Em algumas modalidades, as células modificadas in vitro são administradas a um indivíduo (por exemplo, como uma manipulação ex vivo de células derivadas do indivíduo ou de uma fonte doadora). 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. A integração no genoma hospedeiro é possível com, por exemplo, o retrovírus, lentivírus e métodos de transferência de gene de adenovírus associado, resultando frequentemente em expressão a longo prazo do transgene inserido. Adicionalmente, altas eficiências de transdução foram observadas em muitos tipos de célula diferentes e tecidos alvo.

[0116] O tropismo de um retrovírus pode ser alterado incorporando-se proteínas de envelope estranhas, expandindo a população alvo potencial de células alvo. Os vetores lentivirais são vetores retrovirais que podem transduzir ou infectar células não divisoras e produzir títulos virais tipicamente altos. Seleção de um sistema de transferência de gene retroviral depende do tecido alvo. Os vetores retrovirais são compreendidos de repetições terminais longas de atuação cis com capacidade de

empacotamento por até 6-10 kb de sequência estranha. As LTRs de atuação cis mínimas são diferentes para replicação e empacotamento dos vetores, que são, então, usadas para integrar o construto bidirecional que compreende um transgene na célula alvo para fornecer expressão de transgene permanente. Os vetores retrovirais amplamente usados incluem aqueles com base em vírus de leucemia murina (MuLV), vírus de leucemia de gibão (GaLV), vírus de imunodeficiência símia (SIV), vírus de imunodeficiência humana (HIV) e combinações dos mesmos (consultar, por exemplo, Buchscher et al., *J. Virol.* 66:2.731-2.739 (1992); Johann et al., *J. Virol.* 66:1.635-1.640 (1992); Sommerfelt et al., *Virol.* 176:58-59 (1990); Wilson et al., *J. Virol.* 63:2.374-2.378 (1989); Miller et al., *J. Virol.* 65:2.220-2.224 (1991); PCT/US94/05700).

[0117] Em algumas modalidades, os sistemas com base adenoviral podem ser usados. Os vetores com base adenoviral têm capacidade para eficiência de transdução muito alta em muitos tipos de célula e não necessitam de divisão de célula. Com tais vetores, alto título e altos níveis de expressão foram obtidos. Esse vetor pode ser produzido em grandes quantidades em um sistema relativamente simples. Os vetores adenovirais recombinantes deficientes de replicação podem ser produzidos em alto título e infectam prontamente vários tipos de célula diferentes. A maioria dos vetores adenovirais são modificados geneticamente de modo que um transgene substitua os genes Ad Ela, Elb e/ou E3; subsequentemetne, o vetor defeituoso de replicação é propagado em células 293 humanas que fornecem função de gene deletado em trans. Os vetores Ad podem transduzir múltiplos tipos de tecidos in vivo, incluindo células diferenciadas não divisíveis como aquelas encontradas no fígado, rim e músculo. Os vetores Ad convencionais têm uma grande capacidade de transporte. Um exemplo do uso de um vetor Ad em um teste clínico envolveu a terapia de polinucleotídeo para imunização antitumoral com injeção intramuscular (Sterman et al., *Hum. Gene Ther.* 7:1.083-9 (1998)). Exemplos adicionais do uso de vetores de

adenovírus para transferência de gene em testes clínicos incluem Rosenecker et al., *Infection* 24:1 5-10 (1996); Sterman et al., *Hum. Gene Ther.* 9:7 1.083-1.089 (1998); Welsh et *a\.,Hum. Gene Ther.* 2:205-18 (1995); Alvarez et *a\.. Hum. Gene Ther.* 5:597-613 (1997); Topf et al., *Gene Ther.* 5:507-513 01998); Sterman et. *a&..,Hum. Gene Ther.* 7:1.083-1.089 (1998).

[0118] Em algumas modalidades, vetores de adenovírus associado (AAV) são usados para distribuir construtos de ácido nucleico bidirecional fornecidos no presente documento. Os vetores AAV são bem conhecidos e foram usados para transduzir células com ácidos nucleicos alvo, por exemplo, na produção in vitro de ácidos nucleicos e peptídeos, e para procedimentos de terapia genética in vivo e ex vivo (consultar, por exemplo, West et al., *Virology* 160:38-47 (1987); Patente US nº 4.797.368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1.351 (1994). A construção de vetores AAV recombinantes é descrita em várias publicações, incluindo a Patente US nº 5.173.414; Tratschin et al., *Mol. Cell. Biol.* 5:3.251-3.260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2.072-2.081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); e Samulski et al., *J. Virol.* 63:03.822-3.828 (1989). Em algumas modalidades, o vetor viral pode ser um vetor de AAV. Em algumas modalidades, o vetor AAV é, por exemplo, AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37,

[0119] AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10 ou AAVLK03 assim coo qualquer sorotipo AAV inovador pode ser usado de acordo com a presente invenção. Os vetores de adenovírus associado recombinantes de vetor AAV são uma alternativa promissora para sistemas de distribuição de ácido nucleico, por exemplo, aqueles com base no adenovírus associado tipo 2 de parvovírus não patogênico e defeituoso.

[0120] Conforme usado no presente documento, “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, AV 10, AAV11, AAV12, 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, 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).

[0121] Em outras modalidades, o vetor viral pode se um vetor de lentivírus. 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 30kb 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 distribuir 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, como uma proteína Cas (por exemplo, Cas9), enquanto um segundo vetor de AAV pode conter uma ou mais sequências guia.

[0122] As células de empacotamento são usadas para formar partículas de vírus que têm capacidade para infectar uma célula hospedeira. Tais células incluem 293 células, que podem empacotar adenovírus e AAV, e células ψ2 ou células PA317, que empacotam retrovírus. Os vetores virais usados em terapia genética são usualmente gerados por uma linhagem celular produtora que empacota um vetor de ácido nucleico em uma partícula viral. Os vetores contêm tipicamente as sequências virais mínimas necessárias para empacotar, outras sequências virais substituídas por sequências que codificam a proteína a ser expressa. As funções virais ausentes são fornecidas em trans pela linhagem celular de empacotamento. Por exemplo, os

vetores AAV usados em terapia genética possuem apenas tipicamente sequências de repetição terminal invertida do (ITR) do genoma AAV que são necessárias para o empacotamento. O DNA viral é empacotado em uma linhagem celular, que contém um plasmídeo auxiliar que codifica outros genes AAV, a saber, rep e cap, mas que não têm sequências de ITR. A linhagem celular também pode ser infectada com adenovírus como um auxiliar. O vírus auxiliar promove replicação do vetor AAV e expressão de genes AAV do plasmídeo auxiliar.

[0123] Os vetores de terapia de gene podem ser distribuídos in vivo por administração a um paciente individual, tipicamente por administração sistêmica (por exemplo, infusão intravenosa, intraperitoneal, intramuscular, subdérmica ou intracraniana) ou aplicação tópica, conforme descrito abaixo. Alternativamente, os vetores podem ser distribuídos a células ex vivo, como células explantadas de um paciente individual (por exemplo, linfócitos, aspirados de medula óssea, biópsia de tecido) ou células-tronco hematopoiéticas doadoras universais, seguidas por reimplantação das células em um paciente, usualmente após a seleção para células que incorporaram o vetor.

[0124] Em algumas modalidades, além dos construtos de ácido nucleico bidirecional divulgados no presente documento, o sistema de vetor pode compreender adicionalmente ácidos nucleicos que codificam uma nuclease. Em algumas modalidades, além dos construtos de ácido nucleico bidirecional divulgados no presente documento, 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 Cas proteína, 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 bidirecionais 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 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 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 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, 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 naturalmente encontrados juntos com o crRNA, trRNA, ou crRNA e trRNA. Qualquer um dos vetores descritos no presente documento pode ser distribuído por lipossoma, uma nanopartícula, um exossoma, uma microvesícula, e/ou nanopartículas de lipídio (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 distribuídos 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 distribuídos por LNP. Qualquer uma das LNPs e formulações de LNP descritas no presente documento são adequadas para distribuição dos guias.

[0125] As nanopartículas lipídicas (LNPs) são um meio bem conhecido para distribuição de nucleotídeo e carga de proteína, e podem ser usadas para a distribuição dos construtos de ácido nucleico bidirecional divulgados no presente documento. Em algumas modalidades, LNPs podem ser usadas para distribuir componentes de um sistema de edição de gene. Em algumas modalidades, as LNPs distribuem ácido nucleico (por exemplo, DNA ou RNA), proteína (por exemplo, agente de ligação de DNA guiado por RNA), ou ácido nucleico junto com proteína.

[0126] Em algumas modalidades, é fornecido no presente documento um método para distribuir o construto de ácido nucleico bidirecional divulgado no presente documento a uma célula hospedeira ou indivíduo, em que o construto é distribuído por meio de uma LNP. Em algumas modalidades, é fornecido no presente documento um método para distribuir o construto de ácido nucleico bidirecional divulgado no presente documento a uma célula hospedeira ou indivíduo, em que um ou mais componentes de um sistema de edição de gene, como um sistema de CRISPR/Cas nuclease, são distribuídos por meio de uma LNP. Em algumas modalidades, as LNPs compreendem um construto bidirecional e/ou um ou mais componentes de um sistema de edição de gene (por exemplo, RNA guia e/ou agente de ligação de DNA guiado por RNA ou um mRNA que codifica agente de ligação de DNA guiado por RNA).

[0127] Em algumas modalidades, é fornecida no presente documento uma composição que compreende o construto de ácido nucleico bidirecional divulgado no presente documento e uma LNP. Em algumas modalidades, a composição compreende adicionalmente componentes de um sistema de edição de gene (por exemplo, RNA guia e/ou um agente de ligação de DNA guiado por RNA, como Cas9

ou um sistema de vetor com capacidade para codificar o mesmo). Em algumas modalidades, uma composição que compreende o construto de ácido nucleico bidirecional divulgado no presente documento e uma LNP que compreende um RNA guia e/ou um mRNA que codifica um agente de ligação de DNA guiado por RNA, como Cas9, é fornecida no presente documento.

[0128] 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.

[0129] Eletroporação é um meio bem conhecido para distribuição de carga, e qualquer metodologia de eletroporação pode ser usada para distribuição do construto bidirecional divulgado no presente documento. Em algumas modalidades, a eletroporação pode ser usada para distribuir o construto bidirecional divulgado no presente documento, opcionalmente com um RNA guia e/ou 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) distribuído pelos mesmos ou por diferentes meios.

[0130] Em algumas modalidades, a presente divulgação inclui um método para distribuir o construto bidirecional divulgado no presente documento a uma célula in vitro, em que o construto bidirecional é distribuído por meio de uma LNP. Em

algumas modalidades, o construto bidirecional é distribuído por um meio não LNP, como por meio de um sistema AAV, e um RNA guia e/ou 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) é distribuído por uma LNP.

[0131] Em algumas modalidades, o construto bidirecional descrito no presente documento, sozinho ou parte de um vetor, é formulado em ou administrado por meio de uma nanopartícula lipídica; consultar, por exemplo, WO/2017/173054, cujo conteúdo está aqui incorporado a título de referência, em sua totalidade.

[0132] Qualquer um dos vetores descritos no presente documento pode ser distribuído por LNP. Qualquer um dos LNPs e formulações de LNP descritos no presente documento são adequados para a distribuição dos gRNAs, uma Cas nuclease ou um mRNA que codifica uma Cas nuclease, combinações dos mesmos, e/ou o construto bidirecional divulgado no presente documento. Em algumas modalidades, uma composição de LNP é abrangida que compreende: um componente de RNA e um componente de lipídio, em que o componente de lipídio compreende um lipídio de amina, como um lipídio ionizável biodegradável; e em que o componente de RNA compreende um RNA guia e/ou um mRNA que codifica uma Cas nuclease.

[0133] Em alguns casos, o componente de lipídio compreende lipídio ionizável biodegradável, colesterol, DSPC e PEG-DMG.

[0134] Será evidente que componentes do sistema de edição de gene (por exemplo, RNA guia e/ou agente de ligação de DNA guiado por RNA) e construtos bidirecionais podem ser distribuídos com o uso dos mesmos ou diferentes sistemas. Por exemplo, o RNA guia, agente de ligação de DNA guiado por RNA sequência e construto bidirecional podem ser carreados pelo mesmo vetor (por exemplo, vetor AAV) ou ser formulados em uma ou mais composições de LNP. Alternativamente, o agente de ligação de DNA guiado por RNA (como uma proteína ou mRNA) e/ou gRNA pode ser carregado por ou associado a uma LNP, enquanto os construtos

bidirecionais podem ser carregados por um vetor, ou vice-versa. Adicionalmente, os diferentes sistemas de distribuição podem ser administrados pelas mesmas ou diferentes rotas.

[0135] Os sistemas de distribuição diferentes podem ser distribuídos *in vitro* ou *in vivo* simultaneamente ou em qualquer ordem sequencial. Em algumas modalidades, o construto bidirecional, RNA guia e agente de ligação de DNA guiado por RNA pode ser distribuído 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 bidirecional pode ser distribuído *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) distribuir o RNA guia e/ou Cas nuclease, como um vetor e/ou agente de ligação de DNA guiado por RNA a uma LNP individualmente ou em conjunto como uma ribonucleoproteína (RNP). Em algumas modalidades, o construto doador pode ser distribuído 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 distribuído em intervalos de uma semana, por exemplo, na semana 1, semana 2 e semana 3, etc. Como um exemplo adicional, o RNA guia/ou 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 distribuídos *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) distribuir o construto bidirecional como um vetor e/ou associado a uma LNP. Em algumas modalidades, o RNA guia de albumina pode ser distribuído 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 distribuído em intervalos de uma semana, por exemplo, na semana 1, semana 2 e semana 3, etc. Em algumas modalidades, a Cas nuclease pode ser distribuída em múltiplas administrações, por exemplo, pode ser distribuída 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.

V. MÉTODOS DE USO

[0136] A presente divulgação fornece métodos para usar o construto de ácido nucleico bidirecional descrito no presente documento em várias aplicações. Em algumas modalidades, os métodos para usar o construto de ácido nucleico bidirecional descrito no presente documento em várias aplicações incluem o uso de um sistema de edição de gene, como o sistema CRISPR/Cas, conforme descrito no presente documento.

[0137] Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para modificar um lócus alvo (por exemplo, inserir um gene trans em um sítio alvo em um lócus) que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, um RNA guia e um agente de ligação de DNA guiado por RNA conforme descrito no presente documento (por exemplo, uma Cas nuclease como Cas9). Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para modificar um lócus alvo que compreende clivar uma sequência alvo em uma célula hospedeira e inserir um construto de ácido nucleico bidirecional descrito no presente documento, que utiliza opcionalmente um RNA guia e um agente de ligação de DNA guiado por RNA conforme descrito no presente documento (por exemplo, uma Cas nuclease, como Cas9) para a etapa de clivagem.

[0138] Em algumas modalidades, é fornecido no presente documento um

método in vitro ou in vivo para introduzir um construto em uma célula hospedeira que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, um RNA guia, e um agente de ligação de DNA guiado por RNA conforme descrito no presente documento (por exemplo, uma Cas nuclease, como Cas9). Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para introduzir um construto em uma célula hospedeira que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema ZFN, TALEN ou CRISPR/Cas9.

[0139] Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para aumentar a expressão de um polipeptídeo em uma célula hospedeira que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, um RNA guia, e um agente de ligação de DNA guiado por RNA conforme descrito no presente documento (por exemplo, uma Cas nuclease, como Cas9). Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para aumentar a expressão de um polipeptídeo em uma célula hospedeira que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema ZFN, TALEN ou CRISPR/Cas9. O polipeptídeo pode ser extracelular.

[0140] O construto bidirecional pode ser administrado por meio de um vetor, como um vetor de ácido nucleico. O RNA guia e o agente de ligação de DNA guiado por RNA podem ser administrados individualmente, ou em qualquer combinação, por exemplo, por meio de uma LNP que compreende um RNA guia e um mRNA que codifica o agente de ligação de DNA guiado por RNA. A administração e a distribuição a uma célula hospedeira podem ser efetuadas por qualquer um dentre os métodos de

distribuição descritos no presente documento.

[0141] Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para expressar um polipeptídeo codificado por um transgene em um lócus alvo que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, um RNA guia, e um agente de ligação de DNA guiado por RNA conforme descrito no presente documento (por exemplo, uma Cas nuclease, como Cas9). Em algumas modalidades, é fornecido no presente documento um método in vitro ou in vivo para expressar um polipeptídeo codificado por um transgene em um lócus alvo que compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema ZFN, TALEN ou CRISPR/Cas9. Em algumas modalidades, um método para produzir uma célula hospedeira para expressar um polipeptídeo compreende administrar ou distribuir a uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema ZFN, TALEN ou CRISPR/Cas9.

[0142] O construto bidirecional, RNA guia e agente de ligação de DNA guiado por RNA, por exemplo, pode ser administrado individualmente, ou em qualquer combinação, conforme descrito no presente documento. Em algumas modalidades, o construto bidirecional, RNA guia e o agente de ligação de DNA guiado por RNA podem ser distribuídos simultaneamente ou sequencialmente, por exemplo, em um vetor, dois vetores, vetores individuais, uma LNP, duas LNPs, LNPs individuais ou uma combinação dos mesmos. A administração e a distribuição a uma célula hospedeira podem ser efetuadas por qualquer um dentre os métodos de distribuição descritos no presente documento.

[0143] Além disso, em algumas modalidades, os métodos envolvem a inserção no lócus de albumina, como íntron de albumina 1, por exemplo, com o uso

de um RNA guia que compreende uma sequência selecionada dentre qualquer uma das Tabelas 5, 6, 7, 8, 9 e 10. Em determinadas modalidades 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.

[0144] 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, como um gene de albumina humana, que compreende administrar ou distribuir 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, como um íntron 1 de albumina humana, que compreende administrar ou distribuir a uma célula hospedeira ou população de células hospedeiras qualquer um ou mais dentre os gRNAs, construto doador (por exemplo, construto bidirecional que compreende um ácido nucleico que codifica um polipeptídeo heterólogo) e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease ou ácido nucleico que codifica uma 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 porto seguro humano, como tecido hepático ou célula hospedeira de hepatócito, que compreende administrar ou distribuir 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 um polipeptídeo heterólogo), e agentes de ligação de DNA guiados por RNA (por exemplo, Cas nuclease ou ácido nucleico que codifica uma Cas nuclease) descritos no presente documento.

[0145] Inserção e/ou expressão de um transgene pode ser em seu lócus cognato, (por exemplo, inserção de um transgene de tipo selvagem no lócus endógeno) ou em um lócus não cognato (por exemplo, lócus de porto seguro, como albumina) conforme descrito no presente documento.

[0146] Em algumas modalidades, a célula hospedeira é um tipo de célula sem divisão. 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, retêm 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, como um hepatócito de camundongo, cyno ou ser humano. Em algumas modalidades, a célula hospedeira é um miócito, como um monócito de camundongo, cyno 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 transgene codificado pelo construto bidirecional divulgado no presente documento. Em algumas modalidades, é fornecida no presente documento uma célula hospedeira feita por um método divulgado no presente documento. Em determinadas modalidades, a célula hospedeira é produzida administrando ou distribuindo-se uma célula hospedeira um construto de ácido nucleico bidirecional descrito no presente documento, e um sistema de edição de gene, como um sistema ZFN, TALEN ou CRISPR/Cas9.

[0147] Um método para expressar um polipeptídeo do construto bidirecional descrito no presente documento também é fornecido. De modo similar, uma célula hospedeira que compreende o construto bidirecional descrito no presente documento pode expressar um polipeptídeo codificado pelo construto. Em algumas modalidades, o polipeptídeo é um polipeptídeo secretado. Em algumas modalidades, o polipeptídeo é um cuja função é normalmente efetuada *(*por exemplo*,* funcionalmente ativa) como um polipeptídeo secretado. Um “polipeptídeo secretado”, conforme usado no presente documento, se refere a uma proteína que é secretada pela célula. Em algumas modalidades, o polipeptídeo é um polipeptídeo intracelular. Em algumas modalidades, o polipeptídeo é um em que sua função é normalmente efetuada (por exemplo, funcionalmente ativa) 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. Em algumas modalidades, o polipeptídeo é um polipeptídeo mutante (por exemplo, um mutante hiperativo de um polipeptídeo de tipo selvagem). Em algumas modalidades, o polipeptídeo é uma proteína de fígado. Em algumas modalidades, o polipeptídeo é uma proteína não proveniente de fígado. Em algumas modalidades, o polipeptídeo inclui, mas sem limitação, Fator IX e variantes do mesmo. Em algumas modalidades, o fígado polipeptídeo é, por exemplo, um polipeptídeo para abordar um distúrbio do fígado, como, sem limitação, tirosinemia, doença de Wilson, doença de Tay-Sachs, hiperbilirubinema (Crigler-Najjar), porfiria intermitente aguda, citrulinemia tipo 1, colestase intra-hepática familiar progressiva ou doença da urina em xarope de ácer.

[0148] 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 polipeptídeo heterólogo 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 do polipeptídeo é alcançado por pelo menos 7 dias, pelo menos 14 dias, ou pelo menos 28 dias. Em modalidades adicionais, o método compreende manter a atividade de polipeptídeo heteróloga e/ou proteína leves após uma dose única de construto bidirecional por pelo menos 1, 2, 4 ou 6 meses, ou pelo menos 1, 2, 3, 4, ou 5 anos.

[0149] 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 por um controle de célula hospedeira

que não foi administrado ao construto que compreende o transgene. 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). Em algumas modalidades, a expressão do polipeptídeo pela célula hospedeira (in vitro ou in vivo) é aumentada a pelo menos cerca de 1 μg/ml, 2 μg/ml, 3 μg/ml, 4 μg/ml, 5 μg/ml, 6 μg/ml, 7 μg/ml, 8 μg/ml, 9 μg/ml, 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, 1.000 μg/ml, 1.100 μg/ml, 1.200 μg/ml, 1.300 μg/ml, 1.400 μg/ml, 1.500 μg/ml, 1.600 μg/ml, 1.700 μg/ml, 1.800 μg/ml, 1.900 μg/ml, 2.000 μg/ml ou mais, conforme determinado, por exemplo, na célula, plasma e/ou soro de um indivíduo.

[0150] Em algumas modalidades, é fornecido no presente documento um método para tratar um distúrbio associado ao fígado de acordo com os métodos descritos no presente documento. Conforme usado no presente documento, um “distúrbio associado ao fígado” se refere a distúrbios que causam danos ao tecido de fígado diretamente, distúrbios que resultam dos danos ao tecido de fígado, e/ou distúrbios de órgãos tecido não provenientes do fígado que resultaram de um defeito no fígado.

[0151] Em algumas modalidades, o construto bidirecional, RNA guia e agente de ligação de DNA guiado por RNA são administrados individualmente ou em qualquer combinação de modo local ou sistêmico, por exemplo, de modo intravenoso. Em algumas modalidades, o construto bidirecional, RNA guia, e agente de ligação de DNA

guiado por RNA são administrados individualmente ou em qualquer combinação na circulação hepática.

[0152] Em algumas modalidades, o hospedeiro ou indivíduo é um mamífero. Em algumas modalidades, o hospedeiro ou indivíduo é um ser humano. Em algumas modalidades, o hospedeiro ou indivíduo é um primata. Em algumas modalidades, o hospedeiro ou indivíduo é um roedor (por exemplo, camundongo, rato), vaca, porco, macaco, carneiro, cão, gato, peixe ou aves.

[0153] 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

[0154] 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

[0155] Um construto de inserção bidirecional flanqueado por suas ITRs foi

sintetizado e clonado em pUC57-Kan por um fornecedor comercial. O construto resultante (P00147) 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

[0156] A tripla transfecção em células HEK293 foi usada para acondicionar genomas com construtos de interesse para produção de AAV8 e AAVDJ 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): 1.259-71). 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 11. AAV isolado foi dialisado em tampão de armazenamento (PBS com 0,001% de Pluronic F68). O título 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

[0157] 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 XbaI para completar a digestão seguida por inativação térmica de XbaI a 65 °C. O plasmídeo linearizado foi purificado de enzimas e sais tampões. 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 Bioanalyzer (Agilent).

[0158] mRNAs de Cas9 abaixo compreendem ORF de Cas9, SEQ ID NO: 703 ou a 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 DISTRIBUIÇÃO DE MRNA E GRNA DE CAS9

[0159] mRNA e gRNA de Cas9 foram distribuídos 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.

[0160] 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.

[0161] 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.

[0162] 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.

[0163] 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 WO2016010840 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 4ºC 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 DISTRIBUIÇÃO *IN VITRO* DE MRNA DE CAS9, GRNA E CONSTRUTOS DE INSERÇÃO

CÉLULAS HEPA1-6

[0164] Células Hepa1-6 foram colocadas em placa em densidade de 10.000 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 37ºC 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 distribuição 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

[0165] Hepatócitos de camundongo primários (PMH), hepatócitos de cyno 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 pelotizadas 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 33.000 células/poço para PHH e 50.000 células/poço para PCH e 15.000 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 ºC 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 37ºC.

[0166] Para experimentos que utilizam distribuição 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

[0167] 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 s com o uso de diluição 1:10 de amostras (50 µl de reagente + 40 µl de água + 10 µl de meio célula).

[0168] Para experimentos que envolvem a detecção da etiqueta 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 s com o uso de diluição 1:10 de amostras (50 µl de reagente + 40 µl de água + 10 µl de meio célula).

DISTRIBUIÇÃO IN VIVO DISTRIBUIÇÃO DE LNP E/OU AAV

[0169] 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)

[0170] 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.

[0171] Para estudos *in vivo*, o sangue foi coletado e o soro foi isolado, conforme indicado. Os níveis séricos 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

[0172] 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.

[0173] 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.

[0174] 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*

[0175] 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.

[0176] 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 COM E SEM BRAÇOS DE HOMOLOGIA

[0177] 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 distribuição 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 expressão de transgene (por exemplo, níveis de Fator IX humano), conforme descrito no Exemplo 1.

[0178] 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), apenas o vetor (scAAV derivado de plasmídeo P00204) que compreende 200 bp de braços de homologia resultou na expressão detectável de hFIX. 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 em nenhuma 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 COM E SEM BRAÇOS DE HOMOLOGIA

[0179] 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 transgene (por exemplo, hFIX), respectivamente.

[0180] 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 ssAAV sem braços de homologia (vetor de ssAAV 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: % DE INDEL

Modelo	Indel Médio (%)	Desvio Padrão de Indel (%)
scAAV cego (P00123)	66,72	4,09
ssAAV cego (P00147)	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

Modelo	Fator IX Médio (ug/ml)	Desvio Padrão de Fator IX (ug/ml)
scAAV cego (P00123)	0,75	0,28
ssAAV cego (P00147)	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 COM E SEM BRAÇOS DE HOMOLOGIA

[0181] O experimento descrito nesse Exemplo examinou o efeito de incorporação de braços de homologia em vetores de ssAAV *in vivo*.

[0182] 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 transgene (por exemplo, hFIX).

[0183] Conforme mostrado na Figura 4A e na Tabela 14, o uso dos vetores de ssAAV com braços de homologia assimétricos (braços de 300/600 bp, braços de 300/2000 bp e braços de 300/1500 bp para vetores derivados de plasmídeos P00350, P00356 e P00362, respectivamente) para inserção no sítio 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.

[0184] De maneira similar, o uso dos vetores de ssAAV com braços de homologia simétricos (braços de 500 bp e braços de 800 bp para vetores derivados de plasmídeos P00353 e P00354, respectivamente) para inserção no sítio alvejado por G666 resultou em níveis inferiores, porém, detectáveis, em comparação com o uso do vetor bidirecional sem braços de homologia (derivados de P00147) (consultar a Figura 4B e a Tabela 15).

TABELA 14: NÍVEIS DE FIX SÉRICO

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 DE FIX SÉRICO

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 SÍTIOS-ALVO EM HEPATÓCITOS DE CAMUNDONGO PRIMÁRIOS

[0185] Ao ter demonstrado que construtos bidirecionais que carecem de braços de homologia superaram vetores com outras configurações, o experimento descrito nesse Exemplo examinou os efeitos da alteração dos módulos do construto bidirecional, aqui, o ORF e os aceitantes de splicing, e alteração dos gRNAs para alvejar a inserção mediada por CRISPR/Cas9.

[0186] 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).

[0187] Os materiais de distribuição de pacote de ssAAV e lipídio testados nesse Exemplo foram preparados e distribuídos 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 AAV que compreendem os ORFs de hFIX continham um peptídeo HiBit fundido em suas extremidades 3’, e o vetor AAV que compreende apenas genes repórteres compreenderam um ORF de NanoLuc (além de GFP). Esquemas de cada um dos vetores testados são fornecidos na Figura 5A. Os gRNAs testados são mostrados nas Figuras 5B e 5C, com o uso de um número encurtado para aqueles

listados na Tabela 4 (por exemplo, quando os zeros iniciais são omitidos, por exemplo, quando “G551” corresponde a “G000551” na Tabela 4).

[0188] 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 e na Tabela 17. Conforme mostrado na Figura 5D, um nível significativo de formação de indel não resulta necessariamente em expressão mais eficiente dos transgenes. Ao usar modelos derivados de P00411 e P00418, os valores de R^2 foram 0,54 e 0,37, respectivamente, quando os guias com menos que 10% de edição não estão incluídos. O aceitante de splicing de albumina de camundongo e aceitante de splicing de FIX humano resultaram, cada um, na expressão de transgene eficaz. De modo interessante, apesar de ORFs diferentes e aceitantes de splicing, 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 (consultar Figura 5C).

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
G011722	4,20	0,72	4,27	1,20	4,20	0,26
G011723	5,63	1,27	6,07	0,15	5,93	0,15
G011724	6,10	1,28	8,50	2,69	7,13	1,27
G011725	1,93	0,29	2,60	0,79	2,53	0,65
G011726	10,73	1,46	11,70	0,50	12,43	1,33
G011727	14,20	1,56	14,80	2,36	16,20	2,69
G011728	10,55	1,20	13,65	0,92	15,50	1,56
G011729	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: NÍVEIS DE LUCIFERASE

ID Guia	P00411		P00418		P00415	
	Média de Luciferase (RLU)	Desvio Padrão de Luciferase (RLU)	Média de Luciferase (RLU)	Desvio Padrão de Luciferase (RLU)	Média de Luciferase (RLU)	Desvio Padrão de Luciferase (RLU)
G000551	58.000,00	4.331,28	41.800,00	2.165,64	78.633,33	20.274,70
G000552	95.700,00	10.573,08	80.866,67	27.911,35	20.5333,33	30.664,86
G000553	205.333,33	52.993,71	177.333,33	32.929,22	471.666,67	134.001,00
G000554	125.333,33	55.949,38	91.933,33	19.194,10	232.666,67	67.002,49
G000555	59.933,33	11.566,04	77.733,33	11.061,80	155.666,67	15.947,83
G000666	88.500,00	28.735,87	93.266,67	30.861,19	313.000,00	15.394,80
G000667	75.333,33	22.653,11	68.966,67	27.222,11	153.000,00	30.805,84
G000668	164.000,00	56.320,51	133.400,00	65.111,29	429.000,00	120.751,80
G000669	28.933,33	11.636,29	22.033,33	2.413,16	46.466,67	6.543,19
G000670	162.666,67	32.959,57	200.000,00	33.867,39	424.666,67	36.473,73
G011722	16.766,67	3.384,28	8.583,33	4.103,10	24.000,00	8.915,16
G011723	22.733,33	7.252,82	17.133,33	4.905,44	26.100,00	8.109,87
G011724	17.300,00	2.400,00	28.033,33	9.091,94	30.933,33	3.365,02
G011725	8.253,33	1.163,20	8.890,00	1.429,27	20.366,67	13.955,05
G011726	12.223,33	3.742,54	11.610,00	2.490,44	14.950,00	8.176,03
G011727	35.600,00	8.128,35	36.300,00	12.301,22	86.700,00	5.023,94
G011728	14.900,00	5.011,99	22.466,67	7.130,45	38.166,67	13.829,08
G011729	10.460,00	2.543,95	11.223,33	2.220,28	26.966,67	16.085,50
G011730	14.833,33	2.307,24	21.700,00	8.681,59	41.233,33	25.687,03
G011731	16.433,33	3.274,65	22.566,67	2.205,30	20.756,67	13.096,20
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

[0189] O ssAAV e as LNPs testados nesse Exemplo foram preparados e distribuídos 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 transgene (por exemplo, hFIX), respectivamente.

[0190] Em um experimento inicial, 10 diferentes formulações de LNP que contêm 10 diferentes gRNAs que alvejam íntron 1 de albumina foram distribuídas aos camundongos juntamente com ssAAV derivado de P00147. O AAV e a LNP foram distribuídos em 3e11 vg/ms e 4 mg/kg (com relação ao teor de carga de RNA total), respectivamente. Os gRNAs testados nesse experimento são mostrados na Figura 6 e na Tabela 18. Conforme mostrado na Figura 6 e conforme observado *in vitro*, um nível significativo de formação de indel não foi preditivo para inserção ou expressão dos transgenes.

[0191] Em um experimento separado, o painel completo de 20 gRNAs que alvejam os 20 diferentes sítios-alvo testados *in vitro* no Exemplo 5 foram testados *in vivo*. Para essa finalidade, 20 formulações de LNP que contêm os 20 gRNAs que alvejam íntron 1 de albumina foram distribuídas aos camundongos juntamente com ssAAV derivado de P00147. O AAV e a LNP foram distribuídos em 3e11 vg/ms e 1 mg/kg (com relação ao teor de carga de RNA total), respectivamente. Os gRNAs testados neste experimento são mostrados nas Figuras 7A e 7B e Tabelas 19 e 20, com o uso de um número encurtado para aqueles listados na Tabela 4.

[0192] Conforme mostrado, na Figura 7A, 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 transgene (hFIX), conforme visto na plotagem fornecida na Figura 7D. 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*.

[0193] 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 e Tabela 14, 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 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 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.

TABELA 18: NÍVEIS DE FATOR IX E % DE INDEL

Guia	Indel Médio (%)	Desvio Padrão de Indel (%)	Média de Luciferase (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

TABELA 19: % DE EDIÇÃO DE 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
G011722	10,34	1,68
G011723	4,02	0,84
G011724	2,46	0,64
G011725	8,26	1,24
G011726	6,90	1,01
G011727	13,33	6,43
G011728	35,78	9,34
G011729	4,62	1,46
G011730	12,68	3,14
G011731	26,70	1,86

TABELA 20: NÍVEIS DE FIX

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)
G000551	10,88	2,74	10,25	2,51	9,39	3,48
G000555	13,34	2,09	12,00	2,75	12,43	2,57
G000553	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
G000555	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

EXEMPLO 7- DURABILIDADE DE EXPRESSÃO DE HFIX *IN VIVO*

[0194] A durabilidade de expressão de hFIX ao longo do tempo em animais

tratados foi avaliada nesse Exemplo. Para essa finalidade, hFIX foi medido no soro de animais tratados depois da dose, como parte de um estudo de durabilidade de um ano.

[0195] O ssAAV e as LNPs testados nesse Exemplo foram preparados e distribuídos 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 distribuído em 3e11 vg/ms e a LNP foi distribuída em 0,25 ou 1,0 mg/kg (com relação ao teor de carga de RNA total) (n = 5 para cada grupo).

[0196] Conforme mostrado na Figura 9A e 9B e Tabelas 21 e 22, a expressão de hFIX foi sustentada em cada ponto no tempo avaliado para ambos os grupos dentro de 41 semanas ou 52 semanas, respectivamente. Acredita-se que uma queda nos níveis observada nas 8 semanas na Figura 9A 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 HFIX

Semana	Dose			
	0,25 mpk LNP		1 mpk 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 LNP		1 mpk 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 8 - EFEITOS DE DOSES VARIADAS DE AAV E LNP PARA MODULAR EXPRESSÃO DE HFIX *IN VIVO*

[0197] Nesse Exemplo, os efeitos de variação da dose tanto de AAV quanto de LNP para modular a expressão de hFIX foram avaliados em camundongos C57B1/6.

[0198] O ssAAV e as LNPs testados nesse Exemplo foram preparados e distribuídos aos camundongos, conforme descrito no Exemplo 1. A formulação de LNP conteve G000553 e o ssAAV foi derivado de P00147. O AAV foi distribuído em 1e11, 3e11, 1e12 ou 3e12 vg/ms e a LNP foi distribuída 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. Soros foram coletados em dois pontos no tempo para análise de expressão de hFIX.

[0199] Conforme mostrado na Figura 10A (1 semana), Figura 10B (2 semanas) e Tabela 23, a variação da dose de AAV ou LNP pode modular a quantidade de expressão de hFIX *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 9- EXAME *IN VITRO* DE CONSTRUTOS BIDIRECIONAIS ATRAVÉS DE SÍTIOS-ALVO EM HEPATÓCITOS CYNOMOLGUS PRIMÁRIOS E

HUMANOS PRIMÁRIOS

[0200] 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 de cynomolgus (“cyno”) e humana em hepatócitos de cynomolgus primários (PCH) e humanos primários (PHH), respectivamente.

[0201] Os materiais de distribuição de pacote de ssAAV e lipídio testados nesse Exemplo foram preparados e distribuídos 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 P00415), plotado nas Figuras 11B e 12B 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 nas Figuras 11B e 12B. 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 7.

[0202] Conforme mostrado na Figura 11A para PCH e Figura 12A 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 11A e na Tabela 1 devido à falha de determinados pares iniciadores usados para o sequenciamento à base de amplicon). Os dados de edição mostrados nas Figuras 11A e 12A graficamente, são reproduzidos numericamente na Tabela 1 e Tabela 2 abaixo. No entanto, conforme mostrado nas Figuras 11B, 11C e Figuras 12B e 12C, 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 11C é de 0,13, e o valor de R^2

da Figura 12D é de 0,22.

TABELA 1: DADOS DE EDIÇÃO DE ÍNTRON 1 E EXPRESSÃO DE TRANSGENE DE ALBUMINA PARA SGRNAS DISTRIBUÍDOS AOS HEPATÓCITOS CYNOMOLGUS PRIMÁRIOS

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médio	Desvio Padrão de RLU
G009867	25,05	0,21	10.650,67	1.455,97
G009866	18,7	3,96	75.556,67	12.182,98
G009876	14,85	4,88	27.463,33	10.833,53
G009875	12,85	2,33	51.660,00	6.362,36
G009874	28,25	6,01	270.433,30	133.734,10
G009873	42,65	5,59	178.600,00	87.607,25
G009865	59,15	0,21	301.666,70	18.610,03
G009872	48,15	3,46	320.233,30	63.517,43
G009871	46,5	5,23	211.966,70	65.852,44
G009864	33,2	8,34	210.033,30	61.201,33
G009863	54,8	12,45	69.853,33	15.216,92
G009862	44,6	7,21	508.666,70	119.876,30
G009861	28,65	0,21	178.666,70	15.821,93
G009860	33,2	7,07	571.333,30	52.728,87
G009859	0,05	0,07	258.333,30	79.052,73
G009858	14,65	1,77	402.333,30	25.579,94
G009857	23	0,99	312.333,30	73.036,52
G009856	14,8	0,99	95.900,00	21.128,42
G009851	1,5	0,42	105.766,70	27.048,91
G009868	12,15	2,47	43.033,33	9.141,85
G009850	63,45	13,93	228.200,00	101.542,10
G009849	57,55	8,27	225.400,00	46.001,30
G009848	33	5,37	156.333,30	20.647,84
G009847	66,75	7	100.866,70	22.159,72
G009846	61,85	5,02	31.766,67	10.107,59
G009845	54,4	7,5	43.020,00	11.582,23
G009844	47,15	2,05	110.466,70	32.031,44

TABELA 2: DADOS DE EDIÇÃO DE ÍNTRON 1 E EXPRESSÃO DE TRANSGENE DE ALBUMINA PARA SGRNAS DISTRIBUÍDOS AOS HEPATÓCITOS HUMANOS PRIMÁRIOS

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médio	Desvio Padrão de RLU
G009844	19,07	2,07	268.333,30	80.432,17
G009851	0,43	0,35	18.033,33	2.145,54
G009852	47,20	3,96	18.400,00	2.251,67
G009857	0,10	0,14	71.100,00	14.609,24
G009858	8,63	9,16	32.000,00	18.366,55
G009859	3,07	3,50	59.500,00	16.014,99
G009860	18,80	4,90	190.333,30	54.307,76
G009861	10,27	2,51	62.233,33	9.865,26
G009866	13,60	13,55	96.200,00	46.573,81
G009867	12,97	3,04	3.916,67	1.682,03
G009868	0,63	0,32	10.176,67	2.037,80
G009874	49,13	0,60	318.000,00	114.118,40
G012747	3,83	0,23	51.000,00	6.161,17
G012748	1,30	0,35	17.433,33	2.709,86
G012749	9,77	1,50	75.066,67	11.809,04
G012750	42,73	4,58	5.346,67	2.977,35
G012751	7,77	1,16	32.066,67	18.537,62
G012752	32,93	2,27	402.000,00	83.144,45
G012753	21,20	2,95	71.800,00	32.055,73
G012754	0,60	0,10	16.933,33	4.254,80
G012755	1,10	0,10	13.833,33	3.685,56
G012756	2,17	0,40	35.600,00	6.055,58
G012757	1,07	0,25	13.993,33	6.745,08
G012758	0,90	0,10	34.900,00	15.308,82
G012759	2,60	0,35	30.566,67	15.287,36
G012760	39,10	6,58	6.596,67	2.133,13
G012761	36,17	2,43	467.666,70	210.965,20
G012762	8,50	0,57	217.000,00	13.000,00

ID GUIA	% de Edição Média	Desvio Padrão de % de Edição	RLU Médio	Desvio Padrão de RLU
G012763	47,07	3,07	142.333,30	37.581,02
G012764	44,57	5,83	1.423.333,00	261.023,60
G012765	19,90	1,68	179.666,70	57.011,69
G012766	8,50	0,28	243.333,30	17.473,79

[0203] 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).

[0204] Os materiais de ssAAV e LNP foram preparados e distribuídos 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. Conforme acima, 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 12C e mostrado na Tabela 23 como unidades de luciferase relativa (“RLU”). Por exemplo, os vetores AAV continham o ORF de NanoLuc

[0205] (além de GFP). Esquemas dos vetores testados são fornecidos nas Figuras 11B e 12B. Os gRNAs testados são mostrados na Figura 12C com o uso de um número encurtado para aqueles listados na Tabela 1 e Tabela 7.

TABELA 23: DADOS DE EXPRESSÃO DE TRANSGENE DE ÍNTRON 1 DE ALBUMINA PARA SGRNAS DISTRIBUÍDOS AOS HEPATÓCITOS HUMANOS PRIMÁRIOS

Guia	Média de Luciferase (RLU)	Desvio Padrão de Luciferase (RLU)
G009844	3.700.000	509.117
G009852	281.000	69.296
G009857	1.550.000	127.279

Guia	Média de Luciferase (RLU)	Desvio Padrão de Luciferase (RLU)
G009858	551.000	108.894
G009859	1.425.000	77.782
G009860	2.240.000	183.848
G009861	663.500	238.295
G009866	274.000	11.314
G009867	44.700	566
G009874	2.865.000	431.335
G012747	651.000	59.397
G012749	867.000	93.338
G012752	4.130.000	268.701
G012753	1.145.000	162.635
G012757	579.000	257.387
G012760	129.000	36.770
G012761	4.045.000	728.320
G012762	2.220.000	127.279
G012763	1.155.000	205.061
G012764	11.900.000	1.555.635
G012765	1.935.000	134.350
G012766	2.050.000	169.706
LNP	8.430	212

EXEMPLO 10 - TESTE *IN VIVO* DE EXPRESSÃO DE FATOR IX A PARTIR DE UM LÓCUS DE PORTO SEGURO ALTERNATIVO

[0206] Nesse Exemplo, a inserção de ssAAV que compreende um construto de hFIX bidirecional em um lócus de porto seguro alternativo foi avaliada. Para testar a inserção em um lócus de porto seguro alternativo, AAV foi preparado conforme descrito acima. Camundongos foram administrados com AAVs em uma dose de 3e11 vg/camundongo imediatamente seguidos pela administração de LNPs formuladas com mRNAs de Cas9 e RNAs guia em uma dose de 0,3 mg/kg. Animais foram sacrificados 4 semanas depois da dose, e amostras de fígado e sangue foram

coletadas. Edição nas amostras hepáticas foi determinada através de NGS. Níveis de hFIX humano no soro foram determinados através de ELISA. Os dados de NGS e ELISA mostraram inserção e expressão eficaz de hFIX dentro do lócus de porto seguro alternativo.

EXEMPLO 11- TESTE *IN VIVO* DA INSERÇÃO DE GENE DE FATOR IX HUMANO EM PRIMATAS NÃO HUMANOS

[0207] Nesse exemplo, um estudo de 8 semanas foi realizado para avaliar a inserção de gene de Fator IX humano e a expressão de proteína de hFIX em macacos cynomolgus 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.

[0208] Macacos cynomolgus 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 3. Depois do tratamento de AAV, animais receberam tampão ou LNP, conforme descrito na Tabela 3, através de bolo lento ou infusão.

[0209] Duas semanas depois da dose, espécimes hepáticos foram coletados através de biópsia percutânea guiada por ultrassom única. Cada espécime de biópsia foi instantaneamente congelado em nitrogênio líquido e armazenado a -86 a -60 °C. A edição de análise dos espécimes de fígado foi realizada por Sequenciamento de NGS conforme anteriormente descrito.

[0210] 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.

[0211] 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 cynomolgus foram incubados a seguir em poços individuais. O anticorpo de detecção (anticorpo policlonal de Fator 9 anti-humano de ovelha, Abeam, nº de cat ab128048) 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.

[0212] 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 4 e Figura 13. Níveis plasmáticos de hFIX alcançaram níveis anteriormente descritos como terapeuticamente eficazes (George, et al., NEJM 377(23), 2.215-2.227, 2017).

[0213] Conforme medido, níveis de proteína de hFIX de circulação foram prolongados através do estudo de oito semanas (consultar a Figura 13 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), 1.671-1.675, 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 a 40% de atividade de Fator IX de tipo selvagem (faixa abrange 12 a 67% de atividade de Fator IX de tipo selvagem).

TABELA 3: EDIÇÃO EM FÍGADO

ID 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 4: EXPRESSÃO DE HFIX

ID 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

ID 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)
6001	ND	ND	ND	ND
6002	ND	ND	ND	ND
6003	ND	ND	ND	ND

EXEMPLO 12- TESTE *IN VIVO* DA INSERÇÃO DE FATOR IX EM PRIMATAS NÃO HUMANOS

[0214] Nesse exemplo, um estudo foi realizado para avaliar a inserção de gene de Fator IX e expressão de proteína de hFIX em macacos cynomolgus 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.

[0215] 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 11. Níveis de proteína de FIX humano >3 vezes superiores àqueles alcançados no experimento do Exemplo 13 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 a 5 ug/ml; Amiral et al, Clin. Chem., 30(9), 1.512 a 1.516, 1984) foram alcançados com o uso de G009860 nas NHPs pelo menos nos pontos no tempo do dia 14 e 28. Dados iniciais indicaram níveis de proteína de FIX humano de circulação de ~3 a 4 µg/ml no dia 14 após uma dose única, com níveis prolongados através dos primeiros 28 dias (~3-5 µg/ml) do estudo.

[0216] 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.

[0217] 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/ml. 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 24.

TABELA 24: NÍVEIS DE PROTEÍNA DE FATOR IX HUMANO SÉRICO

	Concentração Média Calculada (ug/ml)		
	3001	3002	3003
Dia 7	7,85	5,63	11,20
Dia 14	8,65	11,06	14,70
Dia 28	9,14	14,12	10,85
Dia 42	9,03	33,12*	13,22
Dia 56	10,24	16,72	33,84*

EXEMPLO 13 - ANÁLISE FORA DO ALVO DE GUIAS HUMANOS DE ALBUMINA

[0218] 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 26. O ensaio identificou potenciais sítios fora do alvo para os sgRNAs testados.

TABELA 25: ANÁLISE FORA DO ALVO

ID de gRNA	Alvo	Sequência Guia (SEQ ID NO:)	Contagem de Sítio Fora do Alvo
G012753	Albumina	GACUGAAACUUCACAGAAUA (SEQ ID NO: 20)	62
G012761	Albumina	AGUGCAAUGGAUAGGUCUUU (SEQ ID NO: 28)	75
G012752	Albumina	UGACUGAAACUUCACAGAAU (SEQ ID NO: 19)	223
G012764	Albumina	CCUCACUCUUGUCUGGGCAA (SEQ ID NO: 31)	3.985
G012763	Albumina	UGGGCAAGGGAAGAAAAAAA (SEQ ID NO: 30)	5.443

ID de gRNA	Alvo	Sequência Guia (SEQ ID NO:)	Contagem de Sítio Fora do Alvo
G009857	Albumina	AUUUAUGAGAUCAACAGCAC (SEQ ID NO: 5)	131
G009859	Albumina	UUAAAUAAAGCAUAGUGCAA (SEQ ID NO: 7)	91
G009860	Albumina	UAAAGCAUAGUGCAAUGGAU (SEQ ID NO: 8)	133
G012762	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: 1.129)	304
G000645	VEGFA	GACCCCCUCCACCCCGCCUC (SEQ ID NO: 1.130)	1.641

[0219] Em ensaios de detecção fora do alvo conhecidos, 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 sobrerrepresenta, 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 esses métodos podem ser validados com o uso de sequenciamento alvejado dos potenciais sítios fora do alvo identificados.

EXEMPLO 14 - CONSTRUÇÃO DE CONSTRUTOS PARA A EXPRESSÃO DE PROTEÍNAS SECRETORAS OU NÃO SECRETORAS

[0220] Construtos, como construtos bidirecionais, podem ser projetados de modo que expressem proteínas secretoras ou não secretoras. Para a produção de uma proteína secretora, um construto pode compreender uma sequência sinal que auxilia na translocação do polipeptídeo para o lúmen ER. Alternativamente, um construto pode utilizar a sequência sinal endógena da célula hospedeira (por exemplo, a sequência sinal de albumina endógena quando o transgene é integrado a um lócus de albumina da célula hospedeira).

[0221] Em contraste, construtos para a expressão de proteínas não secretoras podem ser projetados de modo que não compreendam uma sequência sinal e de modo que não utilizem a sequência 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. As sequências de IRES, como IRES de EMCV, permitem a iniciação de tradução de qualquer posição em um mRNA imediatamente a jusante de onde o IRES está localizado. Isso permitiria a expressão de uma proteína que não tem a sequência sinal endógena da célula hospedeira de um sítio de inserção que contém uma sequência sinal a montante (por exemplo, a sequência sinal encontrada no Éxon 1 de lócus de albumina não seria incluída na proteína expressa). Na ausência de uma sequência sinal, a proteína não seria secretada. Exemplos de sequências de IRES que podem ser usados em um construto incluem aqueles de picornavírus (por exemplo, FMDV), vírus de pragas (CFFV), poliovírus (PV), vírus de encefalomiocardite (ECMV), vírus de doença de pés e boca (FMDV), vírus da hepatite C (HCV), vírus de febre suína clássica (CSFV), vírus da leucemia murina (MLV), vírus de imunodeficiência símia (SIV) ou vírus Cricket Paralysis (CrPV).

[0222] 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, como sequências do tipo 2A ou 2A, servem como sinais de salto de ribossomo para produzir múltiplas proteínas individuais de uma transcrição de mRNA única. Conforme mostrado no Plasmídeo ID P00415 de 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 sinal endógena da célula hospedeira (por exemplo, a sequência 2A localizada a montante da proteína de interesse resultaria em clivagem entre a sequência sinal de albumina endógena e a proteína de interesse). Os peptídeos 2A representativos que poderiam ser utilizados são mostrados na Tabela 12. Adicionalmente, (GSG) resíduos podem ser adicionados à extremidade 5' do peptídeo para melhorar a eficiência de clivagem conforme mostrado na Tabela 12.

Tabela 26: Peptídeos de autoclivagem para uso em construtos

Peptídeo	Sequência de Aminoácidos
T2A (SEQ ID NO: 1.131)	EGRGSLLTCGDVEENPGP
P2A (SEQ ID NO: 1.132)	ATNFSLLKQAGDVEENPGP
E2A(SEQ ID NO: 1.133)	QCTNYALLKLAGDVESNPGP
F2A (SEQ ID NO: 1.134)	VKQTLNFDLLKLAGDVESNPGP
T2A com resíduos GSG (SEQ ID NO: 1.135)	GSGEGRGSLLTCGDVEENPGP
P2A com resíduos GSG (SEQ ID NO: 1.136)	GSGATNFSLLKQAGDVEENPGP
E2A com resíduos GSG (SEQ ID NO: 1.137)	GSGQCTNYALLKLAGDVESNPGP
F2A com resíduos GSG (SEQ ID NO: 1.138)	GSGVKQTLNFDLLKLAGDVESNPGP

TABELA 5: SEQUÊNCIAS DE RNA GUIA HUMANO

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	chr4:73404715-73404735	9
G009866	UACUAAAACUUUAUUUUACU	chr4:73404452-73404472	10
G009867	AAAGUUGAACAAUAGAAAAA	chr4:73404418-73404438	11
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	ACCUUUUUUUUUUUUUUACCU	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

ID Guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G012765	ACCUCACUCUUGUCUGGGCA	chr4:73405108-73405128	32
G012766	UGAGCAACCUCACUCUUGUC	chr4:73405114-73405134	33

TABELA 6: SEQUÊNCIAS DE RNA GUIA 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	chr5:904615 69-904615 89	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
G011722	UGCUUGUAUUUUUCUAGUAA	chr5:90461039-90461059	108
G011723	GUAAAUAUCUACUAAGACAA	chr5:90461425-90461445	109
G011724	UUUUUCUAGUAAUGGAAGCC	chr5:90461047-90461067	110
G011725	UUAUAUUAUUGAUAUAUUUU	chr5:90461174-90461194	111

ID Guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
G011726	GCACAGAUAUAAACACUUAA	chr5:90461480-90461500	112
G011727	CACAGAUAUAAACACUUAAC	chr5:90461481-90461501	113
G011728	GGUUUUAAAAAUAAUAAUGU	chr5:90461502-90461522	114
G011729	UCAGAUUUUCCUGUAACGAU	chr5:90461572-90461592	115
G011730	CAGAUUUUCCUGUAACGAUC	chr5:90461573-90461593	116
G011731	CAAUGGUAAAUAAGAAAUAA	chr5:90461408-90461428	117
G013018	GGAAAAUCUGAAGGUGGCAA	chr5:90461563-90461583	118
G013019	GGCGAUCUCACUCUUGUCUG	chr5:90461717-90461737	119

TABELA 7: SEQUÊNCIAS DE RNA GUIA DE CYNO

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

ID Guia	Sequência Guia	Coordenadas Genômicas	SEQ ID NO:
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
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 8. SGRNA DE ALBUMINA HUMANA E PADRÕES DE MODIFICAÇÃO

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG	34	mG*mA*mG*CAACCUCACUCUUG UCUGUUUU AGAmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGCUAGUCC GUU AU C AmAmCmU mUmGmAmA mAmAmAmGmU mGmGmCm	66

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	UGCUUU U		AmCmCmGm A mGmUm CmGmGmU mGmCmU * mU*mU * mU	
G009851	AUGCAUUUGUUUCAAAAU AUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	35	mA*mU*mG*CAUUUGUUUC AAAA UAUG UUUU AGAmGmCmU mAmGmAmA mAmUm AmGmCAAGUUAAAAUAAG GCUA GUCCG UU AU C AmAmCmU mU mGmAmAm AmAmAm GmU mGmGmCm AmCmCmGmAmG mUmCm GmGmU mGmCmU * mU * mU * mU	67
G009852	UGCAUUUGUUUCAAAAUA UUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	36	mU*mG*mC*AUUUGUUUCA AAAU AUUGU UUU AGAmGmCmU mAmGmAmAm AmUmAm GmCAAGUUAAAAUAAGGC UAGU CCGUUA U CAmAmCmUmUmGmAmAm AmA mAmGmUmGmGmCmAmCm CmGm AmGmU mCmGmGmUmGmCmU * m U*mU*mU	68
G009857	AUUUAUGAGAUCAACAGC ACGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	37	mA*mU*mU*UAUGAGAUCA ACAG CACGU UUU AGAmGmCmU mAmGmAmAm AmUmAm GmCAAGUUAAAAUAAGGC UAGU CCGUUA U CAmAmCmUmUmGmAmAm AmA mAmGm U mGmGmCmAmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU * m U	69
G009858	GAUCAACAGCACAGGUUU UGGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA	38	mG*mA*mU*CAACAGCACA GGUU UUGGU UUU AGAmGmCmU mAmGmAmAm AmUmAm GmCAAGUUAAAAUAAGGC UAGU CCGUUA U	70

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U		CAmAmCmUmUmGmAmAm AmA mAmGm U mGmGmCm AmCmCmGm AmGmU mCmGm GmU mGmCmU * mU * mU * mU	
G009859	UUAAAUAAAGCAUAGUGC AAGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	39	mU*mU*mA*AAUAAAGCAUA GU GCAAGUUUUAGAmGmCmU mAm GmAmAmAmU m AmGmC A AGUU A AAAUAAGGCUAGUCCGUU AUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU mCmGmGmU mGmCmU * mU * mU *mU	71
G009860	UAAAGCAUAGUGCAAUGG AUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	40	mU*mA*mA*AGCAUAGUGC AAUG GAU GUUUU AGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	72
G009861	UAGUGCAAUGGAUAGGUC UUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	41	mU*mA*mG*UGCAAUGGAU AGG UCUUGUUUUAGAmGmCmU mAm GmAmAmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUU AUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU mCmGmGmU mGmCmU * mU * mU *mU	73
G009866	UACUAAAACUUUAUUUUA CUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA	42	mU*mA*mC*UAAAACUUUAU UUU ACU GUUUU AGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU	74

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U		mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	
G009867	AAAGUUGAACAAUAGAAA AAGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	43	mA*mA*mA*GUUGAACAAUA GA AAAAGUUUU AGAmGmCmUmAm GmAmAmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUU AUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU mCmGmGmU mGmCmU * mU * mU *mU	75
G009868	AAUGCAUAAUCUAAGUCA AAGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	44	mA*mA*mU*GCAUAAUCUAA GUC A A AGUUUU AGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	76
G009874	UAAUAAAAUUCAAACAUC CUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A AAAGUGGCACCGAGUCGG UGCUUU U	45	mU*mA*mA*UAAAAUUCAAA CAU C CU GUUUU AGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	77
G012747	GCAUCUUUAAAGAAUUAU UUGUUU U AGAGCUAGAAAUAGCAAG UUAAAA U AAGGCUAGUCCGUUAUCA ACUUGA A	46	mG*mC*mA*UCUUUAAAGA AUUA UUUGUUUU AG AmGmCmUmAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU	78

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	AAAGUGGCACCGAGUCGGUGCUUU U		mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	
G012748	UUUGGCAUUUAUUUCUAAAAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	47	mU*mU*mU*GGCAUUUAUUUCUA A A AGUUUUAGAmGmCmU mAmG mAmAmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	79
G012749	UGUAUUUGUGAAGUCUUACAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	48	mU*mG*mU*AUUUGUGAAGUCU U ACAGUUUUAGAmGmCmUmAm GmAmAmAmU mAmGmCA AGUU A AAAUAAGGCUAGUCCGUUAUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU mCmGmGmU mGmCmU * mU * mU *mU	80
G012750	UCCUAGGUAAAAAAAAAAAAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	49	mU*mC*mC*UAGGUAAAAAAAAA A A AGUUUU AG AmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	81
G012751	UAAUUUUCUUUUGCGCACUAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGG	50	mU*mA*mA*UUUUCUUUUGCGCA CU AGUUUUAGAmGmCmUmAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU	82

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	UGCUUU U		mGmCmU * mU * mU * mU	
G012752	UGACUGAAACUUCACAGAAUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	51	mU*mG*mA*CUGAAACUUCACAG A AU GUUUU AG AmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	83
G012753	GACUGAAACUUCACAGAAUAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	52	mG*mA*mC*UGAAACUUCACAGA AU AGUUUU AGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	84
G012754	UUCAUUUUAGUCUGUCUUCUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	53	mU*mU*mC*AUUUUAGUCUGUCU UCUGUUUUAGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmU mGmCmU * mU * mU * mU	85
G012755	AUUAUCUAAGUUUGAAUAUAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU	54	mA*mU*mU*AUCUAAGUUUGAA U AU AGUUUU AGAmGmCmU mAmGmAmAmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUUAUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU	86

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	U		mCmGmGmU mGmCmU * mU * mU *mU	
G012756	AAUUUUUAAAAUAGUAUUCUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	55	mA*mA*mU*UUUUAAAAUAGUA UUCUGUUUUAGAmGmCmU mAm GmAmAmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUU AUCA mAmCmU mU mGmAmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mU mCmGmGmU mGmCmU * mU * mU *mU	87
G012757	UGAAUUAUUCUUCUGUUUAAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	56	mU*mG*mA*AUUAUUCUUC UGUU U A AGUUUU AGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	88
G012758	AUCAUCCUGAGUUUUUCUGUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	57	mA*mU*mC*AUCCUGAGUU UUUC UGUGUUUUAGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	89
G012759	UUACUAAAACUUUAUUUUACGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU	58	mU*mU*mA*CUAAAACUUUA UUU U ACGUUUU AGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUA UCAm AmCmU mU mGmAmAmAmAmAmG mUmGmGmCmAmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU * mU	90

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	U			
G012760	ACCUUUUUUUUUUUUUACCUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	59	mA*mC*mC*UUUUUUUUUUUUUA C CU GUUUUAGAmGmCmU mAmG mAmAmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	91
G012761	AGUGCAAUGGAUAGGUCUUUGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	60	mA*mG*mU*GCAAUGGAUAGGUC UUUGUUUU AGAmGmCmUmAmG mAmAmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	92
G012762	UGAUUCCUACAGAAAAACUCGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	61	mU*mG*mA*UUCCUACAGAAAAA CUCGUUUUAGAmGmCmU mAmG mAm AmAmU mAmGmC A AGUU A A AAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmG mUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	93
G012763	UGGGCAAGGGAAGAAAAAAAGUUU U AGAGCUAGAAAUAGCAAGUUAAAA U AAGGCUAGUCCGUUAUCAACUUGA A AAAGUGGCACCGAGUCGGUGCUUU U	62	mU*mG*mG*GCAAGGGAAGAAA AAAAGUUUUAGAmGmCmUmAm GmAmAmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUUAUCA mAmCmU mUmGmAmAmAmAm GmU mGmGmCm AmCmCmGmAmG mUmCmGmGmU mGmCmU * mU * mU *mU	94

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G012764	CCUCACUCUUGUCUGGGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	63	mC*mC*mU*CACUCUUGUCUGGGCAAGUUUUAGAmGmCmUmAmG m Am AmAmUmAmGmC A AGUU A AAAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	95
G012765	ACCUCACUCUUGUCUGGGCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	64	mA*mC*mC*UCACUCUUGUCUGG GCAGUUUUAGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	96
G012766	UGAGCAACCUCACUCUUGUCGUUU UAGAGCUAGAAAUAGCAAGUUAAAA UAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGCUUU U	65	mU*mG*mA*GCAACCUCACUCUU GU CGUUUUAGAmGmCmU mAmG m Am AmAmU mAmGmC A AGUU A AAAUAAGGCUAGUCCGUUAUCAm AmCmU mUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGm U mCmGmGmUmGmCmU * mU * mU * mU	97

TABELA 9. SGRNA GUIA DE ALBUMINA DE CAMUNDONGO E PADRÃO DE MODIFICAÇÃO

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCCUU GUUUUAGAGCUAGAAAUAGC	120	mA*mU*mU*UGCAUCUGAGAACCCU U GUUUU AGAmGmCmU m AmGmAmAm AmUmAmGmCAAGUUAAAAUA	142

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU		AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	
G000552	AUCGGGAACUGGCAUCUUCA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	121	mA*mU*mC*GGGAACUGGCAUCUUC AGUUUU AGAmGmCmUm AmGmAmAmAmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	143
G000553	GUUACAGGAAAAUCUGAAGG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	122	mG*mU*mU*ACAGGAAAAUCUGAAG GGUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	144
G000554	GAUCGGGAACUGGCAUCUUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	123	mG*mA*mU*CGGGAACUGGCAUCUU CGUUUU AGAmGmCmUm AmGmAmAm AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC	145
			mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	
G000555	UGCAUCUGAGAACCCUUAGG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	124	mU*mG*mC*AUCUGAGAACCCUUAG GGUUUU AGAmGmCmUm AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU	146

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	GGCACCGAGUCGGUGC UUUU		mCmGmGmU mGmCm U*mU*mU*mU	
G00066 6	CACUCUUGUCUGUGGAA ACA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	125	mC*mA*mC*UCUUGUCUGUGG AAAC AGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	147
G00066 7	AUCGUUACAGGAAAAUC UGA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	126	mA*mU*mC*GUUACAGGAAAAU CUG AGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	148
G00066 8	GCAUCUUCAGGGAGUAG CUU GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	127	mG*mC*mA*UCUUCAGGGAGU AGCU U GUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	149
G00066 9	CAAUCUUUAAAUAUGUU GUG GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	128	mC*mA*mA*UCUUUAAAUAUGU UGU GGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	150
G00067 0	UCACUCUUGUCUGUGGA AAC GUUUUAGAGCUAGAAAU AGC	129	mU*mC*mA*CUCUUGUCUGUG GAAA CGUUUU AGAmGmCmU m AmGmAmAm AmUmAmGmCAAGUUAAAAUA	151

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU		AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	
G011722	UGCUUGUAUUUUUCUAGUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	130	mU*mG*mC*UUGUAUUUUUCU AGUA AGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	152
G011723	GUAAAUAUCUACUAAGACAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	131	mG*mU*mA*AAUAUCUACUAAG ACA AGUUUU AGAmGmCmU m AmGmAmAm AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	153
G011724	UUUUUCUAGUAAUGGAAGCC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	132	mU*mU*mU*UUCUAGUAAUGG AAGC CGUUUU AGAmGmCmU m AmGmAmAm AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	154
G011725	UUAUAUUAUUGAUAUAUUUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	133	mU*mU*mA*UAUUAUUGAUAUA UUU U GUUUU AGAmGmCmU m AmGmAmAm AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCmU*mU*mU*mU	155

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G011726	GCACAGAUAUAAACACUUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	134	mG*mC*mA*CAGAUAUAAACACUUA AGUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	156
G011727	CACAGAUAUAAACACUUAAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	135	mC*mA*mC*AGAUAUAAACACUUAA CGUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	157
G011728	GGUUUUAAAAAUAAUAAUGU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	136	mG*mG*mU*UUUAAAAAUAAUAAUG U GUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	158
G011729	UCAGAUUUUCCUGUAACGAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGCUUUU	137	mU*mC*mA*GAUUUUCCUGUAACGA U GUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	159
G011730	CAGAUUUUCCUGUAACGAUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAA	138	mC*mA*mG*AUUUUCCUGUAACGAU CGUUUU AGAmGmCmU mAmGmAmAm AmUmAmGmCAAGUUAAAAUAAGGC U AGU C CGUU AU C AmAmCmU mU mGmAmAmAmAmGmUmGmGmC	160

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	AGU GGCACCGAGUCGGUGC UUUU		mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	
G011731	CAAUGGUAAAUAAGAAA UAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	139	mC*mA*mA*UGGUAAAUAAGAA AUA AGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	161
G013018	GGAAAAUCUGAAGGUGG CAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	140	mG*mG*mA*AAAUCUGAAGGU GGCA AGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	162
G013019	GGCGAUCUCACUCUUGU CUG GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGC UUUU	141	mG*mG*mC*GAUCUCACUCUU GUCU GGUUUU AGAmGmCmU m AmGmAm Am AmUmAmGmCAAGUUAAAAUA AGGC U AGU C CGUU AU C AmAmCmU mU mGm AmAmAmAmAmGmUmGmGmC mAmC mCmGmAmGmU mCmGmGmU mGmCm U*mU*mU*mU	163

TABELA 10: SGRNA CYNO E PADRÕES DE MODIFICAÇÃO

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009844	GAGCAACCUCACUCU U GUCU GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC	197	mG*mA*mG*CAACCUCACUCUU GUCUGU UUUAG AmGmCmU mAmGmAmAmAmUmAmGmCA AGUUA AAAUAAGGCUAGUCCGUUAUC	230

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC uuuu		AmAmCm UmUm GmAmAmAmAmAmGmU mGmGmCmAmCm CmGm AmGmU mCmGmGmU mGmCmU * mU * mU*m U	
G009845	AGCAACCUCACUCUUG UCUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU	198	mA*mG*mC*AACCUCACUCUUGUCUGGU UUUAG AmGmCmU mAmGmAmAmAmUmAmGmCA AGUUA AAAUAAGGCUAGUCCGUUAUC AmAmCm UmUm GmAmAmAmAmAmGmU mGmGmCmAmCm CmGm AmGmU mCmGmGmU mGmCmU * mU * mU*m U	231
G009846	ACCUCACUCUUGUCUG GGGA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU	199	mA*mC*mC*UCACUCUUGUCUGGGGAGU UUU AGAmGmCmU mAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAGUCCGUUAUCAm AmCm U mU mGmAmAmAmAmAmGmU mGmGmCm AmCm CmGm AmGmU mCmGmGmU mGmCmU * mU * mU*mU	232
G009847	CCUCACUCUUGUCUG GGAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	200	mC*mC*mU*CACUCUUGUCUGGGGAAGU UUUA GAmGmCmU mAmGmAmAmAmU mAmGmCAAGU UAAAAUAAGGCUAGUCCGUUAUCAmAm	233
G009848	CUCACUCUUGUCUGGG GAAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	201	mC*mU*mC*ACUCUUGUCUGGGGAAGGU UUU AGAmGmCmU mAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAGUCCGUUAUCAm AmCm U mU mGmAmAmAmAmAmGmU mGmGmCm	234

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
G009849	GGGGAAGGGGAGAAAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	202	mG*mG*mG*GAAGGGGAGAAAAAAAAG UUUUAG AmGmCmU mAmGmAmAmAmU mAmGmCA AGUUAA AAUAAGGCUAGUCCGUUAUCA mAmCmU mUmGm AmAmAmAmAmGmUmGmGmC	235
G009850	GGGAAGGGGAGAAAAAAAAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	203	mG*mG*mG*AAGGGGAGAAAAAAAAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm	236
G009851	AUGCAUUUGUUUCAAAAUAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	204	mA*mU*mG*CAUUUGUUUCAAAAUAUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm	237
G009852	UGCAUUUGUUUCAAAAUAUU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	205	mU*mG*mC*AUUUGUUUCAAAAUAUUGUUUUAG AmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAmAmAmGmU mGmGmCmAmCmCmGm AmGm UmCmGmGmUmGmCmU*mU*mU*mU	238
G009853	UGAUUCCUACAGAAAAGUC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC	206	mU*mG*mA*UUCCUACAGAAAAGUCGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA	239

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU		mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU *mU	
G009854	UACAGAAAAAGUCAGGAUAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	207	mU*mA*mC*AGAAAAAGUCAGGAUAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU *mU	240
G009855	UUUCUUCUGCCUUUAA ACAG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	208	mU*mU*mU*CUUCUGCCUUUAAACAGGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU *mU	241
G009856	UUAUAGUUUUAUAUUCAAAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	209	mU*mU*mA*UAGUUUUAUAUUCAAACGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU *mU	242
G009857	AUUUAUGAGAUCAACAGCAC GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU	210	mA*mU*mU*UAUGAGAUCAACAGCACGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm	243

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
			U mCmGmGmU mGmCmU * mU * mU *mU	
	GGCACCGAGUCGGUGC UUUU			
G009858	GAUCAACAGCACAGGU UUUG GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	211	mG*mA*mU*CAACAGCACAGGUUUUGGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	244
G009859	UUAAAUAAAGCAUAGUGCAA GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	212	mU*mU*mA*AAUAAAGCAUAGUGCAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	245
G009860	UAAAGCAUAGUGCAAUGGAU GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU	213	mU*mA*mA*AGCAUAGUGCAAUGGAUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	246
G009861	UAGUGCAAUGGAUAGGUCUU	214	mU*mA*mG*UGCAAUGGAUAGGUCUUGUU	247

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU		UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	
G009862	AGUGCAAUGGAUAGG UCUUA GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	215	mA*mG*mU*GCAAUGGAUAGGU CUUAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	248
G009863	UUACUUUGCACUUUC C UUAG GUUUUAGAGCUAGAA AUAGC	216	mU*mU*mA*CUUUGCACUUUCC UUAGGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA	249
	AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU		AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	
G009864	UACUUUGCACUUUCC U UAGU GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG	217	mU*mA*mC*UUUGCACUUUCCU UAGUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU	250

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	C UUUU		mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	
G009865	UCUGACCUUUUAUUU UACCU GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	218	mU*mC*mU*GACCUUUUAUUUU ACCUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	251
G009866	UACUAAAACUUUAUU UUACU GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	219	mU*mA*mC*UAAAACUUUAUUU UACUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	252
G009867	AAAGUUGAACAAUAG AAAAA GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	220	mA*mA*mA*GUUGAACAAUAGA AAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	253
G009868	AAUGCAUAAUCUAAG UCAAA GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C	221	mA*mA*mU*GCAUAAUCUAAGU CAAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU *	254

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	UUUU		mU *mU	
G009869	AUUAUCCUGACUUUU UCUGU GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	222	mA*mU*mU*AUCCUGACUUUUU CUGUGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	255
G009870	UGAAUUAUUCCUCUG UUUAA GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	223	mU*mG*mA*AUUAUUCCUCUGU UUAAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGm U mCmGmGmU mGmCmU * mU * mU *mU	256
G009871	UAAUUUUCUUUUGCC C ACUA GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC CGUUAUCAACUUGAA AAAGU GGCACCGAGUCGGUG C UUUU	224	mU*mA*mA*UUUUCUUUUGCCC ACUAGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm Um GmAmAmAmAmGmU mGmGmCmAmCmCm Gm AmGmU mCmGmGmU mGmCmU * mU * mU * mU	257
G009872	AAAAGGUCAGAAUUG UUUAG GUUUUAGAGCUAGAA AUAGC AAGUUAAAAUAAGGC UAGUC	225	mA*mA*mA*AGGUCAGAAUUGU UUAGGUU UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA	258

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU		mAmCmUmUmGmAmAmAm Am AmGmUmGmGmCm AmCmCmGmAmGmU mCmGmGmU mGmCmU * mU * mU *mU	
G009873	AACAUCCUAGGUAAAAUAAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	226	mA*mA*mC*AUCCUAGGUAAAAUAAAGUUUUAG AmGmCmU mAmGmAmAmAmUmAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm Am AmGmUmGmGmCm AmCmCmGmAmGmU mCmGmGmU mGmCmU * mU * mU *mU	259
G009874	UAAUAAAAUUCAAACAUCCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC	227	mU*mA*mA*UAAAAUUCAAACAUCCUGUUUUAG AmGmCmU mAmGmAmAmAmUmAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm Am AmGmUmGmGmCm AmCmCmGmAmGmU mCmGmGmU mGmCmU * mU * mU *mU	260
	CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU		AmAmAm Am AmGmUmGmGmCm AmCmCmGmAmGmU mCmGmGmU mGmCmU * mU * mU *mU	
G009875	UUGUCAUGUAUUUCUAAAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	228	mU*mU*mG*UCAUGUAUUUCUAAAAUGUUUUAG AmGmCmU mAmGmAmAmAmUmAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGmAmAmAm Am AmGmUmGmGmCm AmCmCmGmAmGmU mCmGmGmU mGmCmU * mU * mU *mU	261
G009876	UUUGUCAUGUAUUUCUAAAA	229	mU*mU*mU*GUCAUGUAUUUCUAAAAGUU	262

ID Guia	Sequência Completa	SEQ ID NO:	Sequência Completa Modificada	SEQ ID NO:
	GUUUUAGAGCUAGAAAUAGC AAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC UUUU		UUAG AmGmCmU mAmGmAmAmAmU mAmGmCAAG UUAA AAUAAGGCUAGUCCGUUAUCA mAmCmUm UmGm AmAmAm Am AmGmU mGmGmCm AmCmCmGm AmGmU mCmGmGmU mGmCmU * mU * mU *mU	

TABELA 11: COMPONENTES DE VETOR E SEQUÊNCIAS

ID de Plasmídeo	ITR 5'	Aceitante de Splicing (1ª orientação)	Transgene (1ª orientação)	Poli-A (1ª orientação)	Poli-A (2ª orientação)	Transgene (2ª orientação)	Aceitante de Splicing (2ª orientação)	ITR 3'
P00147	(SEQ ID NO: 263)	Aceitante de Splicing 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 Splicing de Albumina de Camundongo (SEQ ID NO: 269)	(SEQ ID NO: 270)
P00411	(SEQ ID NO: 263)	Aceitante de Splicing 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 Splicing de Fator IX Humano (SEQ ID NO: 274)	(SEQ ID NO: 270)
P00415	(SEQ ID NO: 263)	Aceitante de Splicing de Albumina de Camundongo (SEQ ID	Nluc-P2A- GFP (SEQ ID	SEQ ID NO: 266	SEQ ID NO: 267	Nluc-P2A- GFP (SEQ ID	Aceitante de Splicing de Albumina de Camundongo (SEQ ID	(SEQ ID NO: 270)

ID de Plasmídeo	ITR 5’	Aceitante de Splicing (1ª orientação)	Transgene (1ª orientação)	Poli-A (1ª orientação)	Poli-A (2ª orientação)	Transgene (2ª orientação)	Aceitante de Splicing (2ª orientação)	ITR 3’
		NO: 264)	NO: 275)			NO: 276)	NO: 269)	
P00418	(SEQ ID NO: 263)	Aceitante de Splicing 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 Splicing de Albumina de Camundongo (SEQ ID NO: 269)	(SEQ ID NO: 270)

Íntron 1 de albumina humana: (SEQ ID NO: 1)

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

Sequência de ITR 5’ (SEQ ID NO: 263):

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGA
CCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGC

GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT

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

TAGGTCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTC
TTCATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAG

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
CTAGGGGGTATCCCC

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

AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCA
ATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAG
CATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGT
CCAAACTCATCAATGTATCTTATCATGTCTG

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 Splicing de Albumina de Camundongo (2ª orientação) (SEQ ID NO:269):

CTGTGGAAACAGGGAGAGAAAAACCACACAACATATTTAAAGATTGATG
AAGACAACTAACTGTAATATGCTGCTTTTTGTTCTTCTCTTCACTGACCTA

Sequência de ITR 3’ (2ª orientação) (SEQ ID NO: 270):

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCT
CGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCG
CCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

GATTATTTGGATTAAAAACAAAGACTTTCTTAAGAGATGTAAAATTTTCAT
GATGTTTTCTTTTTTGCTAAAACTAAAGAATTATTCTTTTACATTTCAG

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
AGACTGTTCAAGAAGATCAGCTAA

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
TTATTCGCGTTTTCGTGGTCCAGAAA

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

CTGAAATGTAAAAGAATAATTCTTTAGTTTTAGCAAAAAAGAAAACATCAT
GAAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTTAATCCAAATAATC

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 a 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 a 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
TGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Sequência completa de P00415 (de ITR a 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
TCAGGATCTTGTTGGCGTTCTCGTGGTCCAGGAA

Sequência completa de P00418 (de ITR a 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 a 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 a 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 a 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 a ITR):(SEQ ID NO: 284)TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACC
AAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGA
GCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTTAGC
CTCTGGCAAAATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTC
TGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTT
CATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAAT
TTTCCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTATTATTGTCA
TTATTTGCATCTGAGAACCCTTAGGTGGTTATATTATTGATATATTTTTGGTATCT
TTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAAA
AAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATACAGTTGAATTTATTAC
GGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAAATTAAAAGTGTGT
GTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTATTTCTTATTTACC
ATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTTGAGTTTGAAT
GCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTGAAAAAATAT
AACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCTGTAACG
ATCGGGAACTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAACAAA
AAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTT
TTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGA

ATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAGGGAACC
TTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTT
TTGAAAACACTGAAAGAACAACTGAATTTTGGAAGCAGTATGTTGATGGAGATCA
GTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTC
CTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTA
ACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATA
ACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGT
CCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTTC
TAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTACT
GAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACT
TCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAG
GTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAA
AATGGATTGTAACTGCTGCCCACTGTGTTGAAACTGGTGTTAAAATTACAGTTGT
CGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATGT
GATTCGAATTATTCCTCACCACAACTACAATGCAGCTATTAATAAGTACAACCAT
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 HAF9 de 300/600 bp (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 HAF9 de 300/2000 bp (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 HAF9 de 300/1500 bp (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. Construto de ácido nucleico bidirecional **CARACTERIZADO** pelo fato de que compreende:

a) um primeiro segmento que compreende uma sequência de codificação para um agente, como um polipeptídeo; e

b) um segundo segmento que compreende um complemento reverso de uma sequência de codificação do agente,

em que o construto não compreende um promotor que aciona a expressão do agente, como um polipeptídeo.

2. Construto de ácido nucleico bidirecional **CARACTERIZADO** pelo fato de que compreende:

a) um primeiro segmento que compreende uma sequência de codificação para um primeiro agente, como um primeiro polipeptídeo; e

b) um segundo segmento que compreende um complemento reverso de uma sequência de codificação de um segundo agente, como um segundo polipeptídeo,

em que o construto não compreende um promotor que aciona a expressão do agente (ou agentes).

3. Construto de ácido nucleico bidirecional, de acordo com a reivindicação 1 ou 2, **CARACTERIZADO** pelo fato de que o segundo segmento é 3' do primeiro segmento.

4. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 3, **CARACTERIZADO** pelo fato de que a sequência de codificação do complemento reverso no segundo segmento adota um uso de códon diferente daquele da primeira sequência de codificação do primeiro segmento.

5. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 4, **CARACTERIZADO** pelo fato de que o segundo segmento compreende uma sequência de nucleotídeo 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.

6. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 5, **CARACTERIZADO** pelo fato de que a sequência de codificação do segundo segmento codifica o polipeptídeo com o uso de um ou mais códons alternativos para um ou mais aminoácidos codificados pela sequência de codificação no primeiro segmento.

7. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 6, **CARACTERIZADO** pelo fato de que o segundo segmento compreende um complemento reverso da sequência de codificação do primeiro segmento, ou um fragmento do mesmo.

8. Construto de ácido nucleico bidirecional, de acordo com a reivindicação 7, **CARACTERIZADO** pelo fato de que o complemento reverso é:

a. não substancialmente complementar à sequência de codificação do primeiro segmento;

b. não substancialmente complementar a um fragmento da sequência de codificação do primeiro segmento;

c. altamente complementar à sequência de codificação do primeiro segmento;

d. altamente complementar a um fragmento da sequência de codificação do primeiro segmento;

e. pelo menos 60% idêntico ao complemento reverso da sequência de codificação do primeiro segmento;

f. pelo menos 70% idêntico ao complemento reverso da sequência de codificação do primeiro segmento;

f. pelo menos 90% idêntico ao complemento reverso da sequência de

codificação do primeiro segmento;

g. 50 a 80% idêntico ao complemento reverso da sequência de codificação do primeiro segmento; e/ou

h. pelo menos 60 a 100% idêntico ao complemento reverso da sequência de codificação do primeiro segmento.

9. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 8, sendo que o construto é **CARACTERIZADO** pelo fato de que não compreende um braço de homologia.

10. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 9, **CARACTERIZADO** pelo fato de que o primeiro segmento é ligado ao segundo segmento por um ligante.

11. Construto de ácido nucleico bidirecional, de acordo com a reivindicação 10, **CARACTERIZADO** pelo fato de que o ligante tem 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 500, 1.000, 1.500, 2.000 nucleotídeos em comprimento.

12. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 11, **CARACTERIZADO** pelo fato de que cada um do primeiro e segundo segmentos compreende uma sequência de poliadenilação, como uma sequência sinal de poliadenilação ou uma sequência de cauda de poliadenilação.

13. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 12, sendo que o construto é **CARACTERIZADO** pelo fato de que compreende um sítio aceitante de splicing.

14. Construto de ácido nucleico bidirecional, de acordo com a reivindicação 13, sendo que o construto é **CARACTERIZADO** pelo fato de que compreende um primeiro sítio aceitante de splicing 5’ do primeiro segmento e um segundo sítio aceitante de splicing 3’ do segundo segmento.

15. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 14, sendo que o construto é **CARACTERIZADO** pelo fato de que é

de fita dupla, opcionalmente DNA de fita dupla.

16. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 15, sendo que o construto é **CARACTERIZADO** pelo fato de que é de fita simples, opcionalmente DNA de fita simples.

17. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 16, **CARACTERIZADO** pelo fato de que a sequência que codifica o polipeptídeo é otimizada por códon.

18. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 17, sendo que o construto é **CARACTERIZADO** pelo fato de que compreende uma ou mais das seguintes estruturas terminais: grampo, alças, repetições terminais invertidas (ITR) ou toroide.

19. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 18, sendo que o construto é **CARACTERIZADO** pelo fato de que compreende uma, duas ou três repetições terminais invertidas (ITR).

20. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 19, sendo que o construto é **CARACTERIZADO** pelo fato de que compreende no máximo duas ITRs.

21. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 20, **CARACTERIZADO** pelo fato de que o polipeptídeo é um polipeptídeo secretado.

22. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 20, **CARACTERIZADO** pelo fato de que o polipeptídeo y é um polipeptídeo intracelular.

23. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 2 a 22, **CARACTERIZADO** pelo fato de que o primeiro polipeptídeo e o segundo polipeptídeo são diferentes.

24. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das

reivindicações 1 a 23, **CARACTERIZADO** pelo fato de que o polipeptídeo é um polipeptídeo variante.

25. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 24, **CARACTERIZADO** pelo fato de que o polipeptídeo é uma proteína de fígado.

26. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 24, **CARACTERIZADO** pelo fato de que o polipeptídeo é uma proteína não proveniente de fígado.

27. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 26, sendo que o construto é **CARACTERIZADO** pelo fato de que é um construto independente de homologia.

28. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 27, **CARACTERIZADO** pelo fato de que o polipeptídeo, quando expresso, compreende um peptídeo sinal heterólogo.

29. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 28, **CARACTERIZADO** pelo fato de que o ácido nucleico codifica um peptídeo sinal heterólogo.

30. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 27, **CARACTERIZADO** pelo fato de que o ácido nucleico não codifica um peptídeo sinal.

31. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 27, **CARACTERIZADO** pelo fato de que o polipeptídeo, quando expresso, compreende seu próprio peptídeo sinal.

32. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 31, **CARACTERIZADO** pelo fato de que o ácido nucleico codifica um peptídeo heterólogo.

33. Construto de ácido nucleico bidirecional, de acordo com a reivindicação

32, **CARACTERIZADO** pelo fato de que o peptídeo heterólogo é 2A.

34. Vetor **CARACTERIZADO** pelo fato de que compreende o construto, conforme definido em qualquer uma das reivindicações 1 a 33.

35. Vetor, de acordo com a reivindicação 34, sendo que o vetor é **CARACTERIZADO** pelo fato de que é um vetor viral adenoassociado (AAV).

36. Vetor, de acordo com a reivindicação 34 ou 35, **CARACTERIZADO** pelo fato de que o AAV compreende um genoma de fita simples (ssAAV) ou um genoma autocomplementar (scAAV).

37. Vetor, de acordo com a reivindicação 34 ou 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.64RI, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV12, rh10 e híbridos dos mesmos.

38. Vetor viral **CARACTERIZADO** pelo fato de que compreende um construto de ácido nucleico autocomplementar (ou de fita dupla) que compreende uma sequência de nucleotídeo que codifica um polipeptídeo, em que o vetor não compreende um promotor que aciona a expressão do polipeptídeo.

39. Vetor, de acordo com a reivindicação 38, sendo que o vetor é **CARACTERIZADO** pelo fato de que não compreende um braço de homologia.

40. Nanopartícula lipídica **CARACTERIZADA** pelo fato de que compreende o construto, conforme definido em qualquer uma das reivindicações 1 a 33.

41. Célula hospedeira **CARACTERIZADA** pelo fato de que compreende o construto, conforme definido em qualquer uma das reivindicações 1 a 33.

42. Célula hospedeira, de acordo com a reivindicação 41, sendo que a célula hospedeira é **CARACTERIZADA** pelo fato de que é uma célula de fígado, uma célula muscular ou uma célula neuronal.

43. Célula hospedeira, de acordo com a reivindicação 41, sendo que a célula

hospedeira é **CARACTERIZADA** pelo fato de que é um tipo de célula sem divisão.

44. Célula hospedeira, de acordo com qualquer uma das reivindicações 41 a 43, sendo que a célula hospedeira é **CARACTERIZADA** pelo fato de que expressa o polipeptídeo codificado pelo construto.

45. Célula hospedeira, de acordo com qualquer uma das reivindicações 41 a 44, sendo que a célula hospedeira é **CARACTERIZADA** pelo fato de que é um hepatócito.

46. Método para modificar um lócus alvo **CARACTERIZADO** pelo fato de que compreende dotar uma célula de um construto, conforme definido em qualquer uma das reivindicações 1 a 33, um vetor, conforme definido em qualquer uma das reivindicações 34 a 39, ou uma LNP, conforme definido na reivindicação 40.

47. Método para introduzir um construto em uma célula **CARACTERIZADO** pelo fato de que compreende dotar a célula de um construto, conforme definido em qualquer uma das reivindicações 1 a 33, um vetor, conforme definido em qualquer uma das reivindicações 34 a 39, ou uma LNP, conforme definido na reivindicação 40.

48. Método para expressar um polipeptídeo em uma célula **CARACTERIZADO** pelo fato de que compreende dotar uma célula de um construto, conforme definido em qualquer uma das reivindicações 1 a 33, um vetor, conforme definido em qualquer uma das reivindicações 34 a 39, ou uma LNP, conforme definido na reivindicação 40.

49. Método para aumentar a expressão de um polipeptídeo em uma célula **CARACTERIZADO** pelo fato de que compreende dotar a célula de um construto, conforme definido em qualquer uma das reivindicações 1 a 33, um vetor, conforme definido em qualquer uma das reivindicações 34 a 39, ou uma LNP, conforme definido na reivindicação 40.

50. Método, de acordo com qualquer uma das reivindicações 46 a 49, **CARACTERIZADO** pelo fato de que o construto, vetor ou LNP é fornecido in vivo.

51. Método, de acordo com qualquer uma das reivindicações 46 a 49, **CARACTERIZADO** pelo fato de que o construto, vetor ou LNP é fornecido in vitro.

52. Método ex vivo para modificar um lócus alvo **CARACTERIZADO** pelo fato de que compreende dotar uma célula de um construto, conforme definido em qualquer uma das reivindicações 1 a 33, um vetor, conforme definido em qualquer uma das reivindicações 34 a 39, ou uma LNP, conforme definido na reivindicação 40.

53. Método, de acordo com qualquer uma das reivindicações 46 a 52, **CARACTERIZADO** pelo fato de que a célula é um tipo de célula sem divisão.

54. Método, de acordo com qualquer uma das reivindicações 46 a 53, **CARACTERIZADO** pelo fato de que compreende adicionalmente dotar a célula de uma nuclease, em que a nuclease tem capacidade de introduzir um ssDNA sítio-específico ou quebra de dsDNA.

55. Método, de acordo com qualquer uma das reivindicações 46 a 54, **CARACTERIZADO** pelo fato de que compreende dotar a célula de um agente de ligação de DNA guiado por RNA e um RNA guia (gRNA).

56. Método, de acordo com a reivindicação 55, **CARACTERIZADO** pelo fato de que o gRNA é um gRNA único (sgRNA).

57. Método, de acordo com a reivindicação 55 ou 56, **CARACTERIZADO** pelo fato de que o construto, agente de ligação de DNA guiado por RNA e o gRNA são fornecidos à célula simultaneamente.

58. Método, de acordo com a reivindicação 55 ou 56, **CARACTERIZADO** pelo fato de que o construto, agente de ligação de DNA guiado por RNA e o gRNA são fornecidos à célula sequencialmente, em qualquer ordem.

59. Método, de acordo com a reivindicação 58, **CARACTERIZADO** pelo fato de que o construto é fornecido à célula antes de fornecer o gRNA e o agente de ligação de DNA guiado por RNA.

60. Método, de acordo com a reivindicação 58, **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 fornecido à célula antes de fornecer o construto.

61. Método, de acordo com qualquer uma das reivindicações 54 a 60, **CARACTERIZADO** pelo fato de que o construto é fornecido em um vetor.

62. Método, de acordo com qualquer uma das reivindicações 54 a 60, **CARACTERIZADO** pelo fato de que o construto é fornecido em uma nanopartícula de lipídio.

63. Método, de acordo com qualquer uma das reivindicações 54 a 62, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é fornecido como um ácido nucleico que codifica um agente de ligação de DNA guiado por RNA.

64. Método, de acordo com qualquer uma das reivindicações 54 a 62, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é fornecido como uma enzima Cas ou um mRNA que codifica um agente de ligação de DNA guiado por RNA.

65. Método, de acordo com qualquer uma das reivindicações 54 a 62, **CARACTERIZADO** pelo fato de que o agente de ligação de DNA guiado por RNA é uma Cas nuclease, como uma Cas9 nuclease.

66. Método, de acordo com a reivindicação 65, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma Cas nuclease de classe 2, ou uma variante da mesma.

67. Método, de acordo com qualquer uma das reivindicações 54 a 66, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma Cas9 de *S. pyogenes*, ou uma variante da mesma.

68. Método, de acordo com qualquer uma das reivindicações 54 a 67, **CARACTERIZADO** pelo fato de que a Cas nuclease é uma nickase.

69. Método, de acordo com qualquer uma das reivindicações 54 a 68,

CARACTERIZADO pelo fato de que o lócus é um lócus de albumina.

70. Método, de acordo com qualquer uma das reivindicações 54 a 69, **CARACTERIZADO** pelo fato de que a célula é um tipo de célula sem divisão.

71. Método, de acordo com qualquer uma das reivindicações 54 a 70, **CARACTERIZADO** pelo fato de que a célula é uma célula de fígado, uma célula neuronal ou uma célula muscular.

72. Célula **CARACTERIZADA** pelo fato de que é produzida pelo método, conforme definido em qualquer uma das reivindicações 46 a 71.

73. Célula, de acordo com a reivindicação 72, sendo que a célula é **CARACTERIZADA** pelo fato de que é uma célula hospedeira.

74. Célula hospedeira, de acordo com a reivindicação 73, sendo que a célula hospedeira é **CARACTERIZADA** pelo fato de que é uma célula de fígado, uma célula muscular ou uma célula neuronal.

75. Célula, de acordo com qualquer uma das reivindicações 72 a 74, **CARACTERIZADA** pelo fato de que a célula hospedeira é um tipo de célula sem divisão.

76. Célula, de acordo com qualquer uma das reivindicações 72 a 75, **CARACTERIZADA** pelo fato de que a célula hospedeira expressa o polipeptídeo codificado pelo construto.

77. Célula, de acordo com qualquer uma das reivindicações 72 a 76, **CARACTERIZADA** pelo fato de que a célula hospedeira é um hepatócito.

78. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 33, **CARACTERIZADO** pelo fato de que o agente é um agente terapêutico.

79. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 33, **CARACTERIZADO** pelo fato de que o agente é qualquer um ou mais dentre um RNA funcional, um mRNA ou um polipeptídeo.

80. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 33, **CARACTERIZADO** pelo fato de que o agente é um polipeptídeo.

81. Construto de ácido nucleico bidirecional, de acordo com qualquer uma das reivindicações 1 a 33, **CARACTERIZADO** pelo fato de que o agente é um polipeptídeo terapêutico.

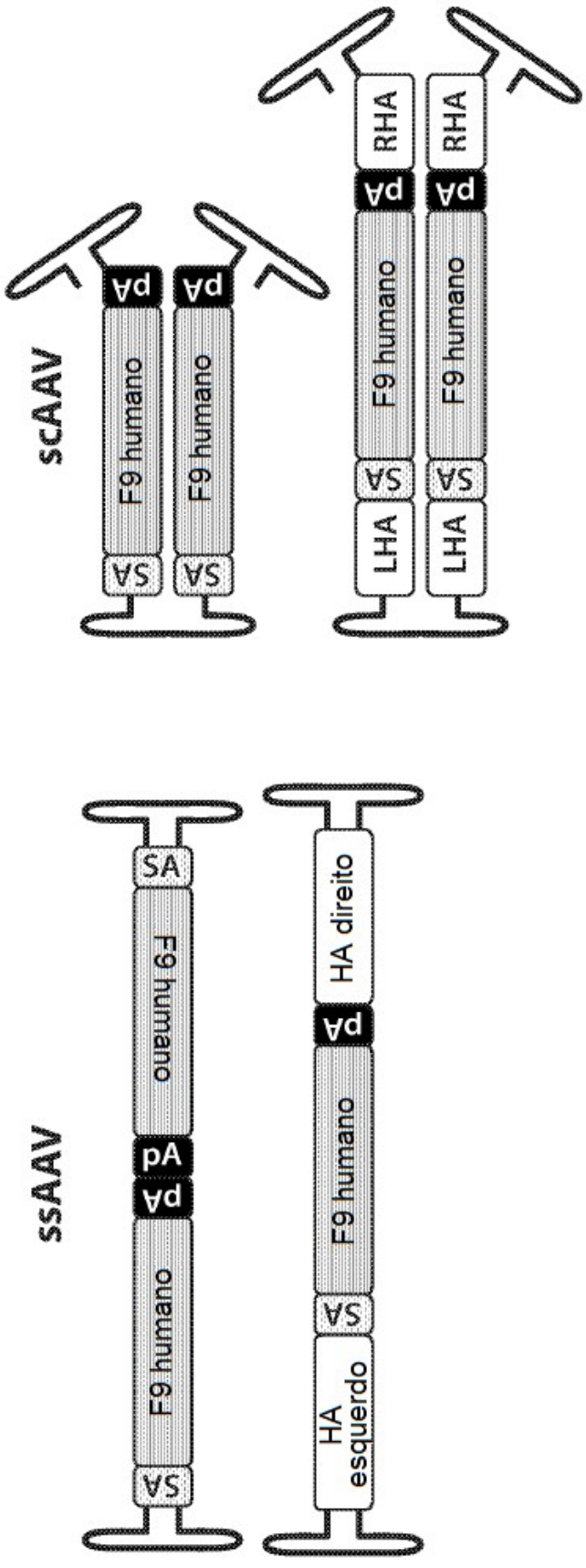

scAAV
SA
F9 humano
pA
F9 humano
SA
pA
LHA
SA
F9 humano
pA
RHA
LHA
SA
F9 humano
pA
RHA
ssAAV
SA
F9 humano
pA
F9 humano
SA
HA esquerdo
SA
F9 humano
pA
HA direito

FIG. 1

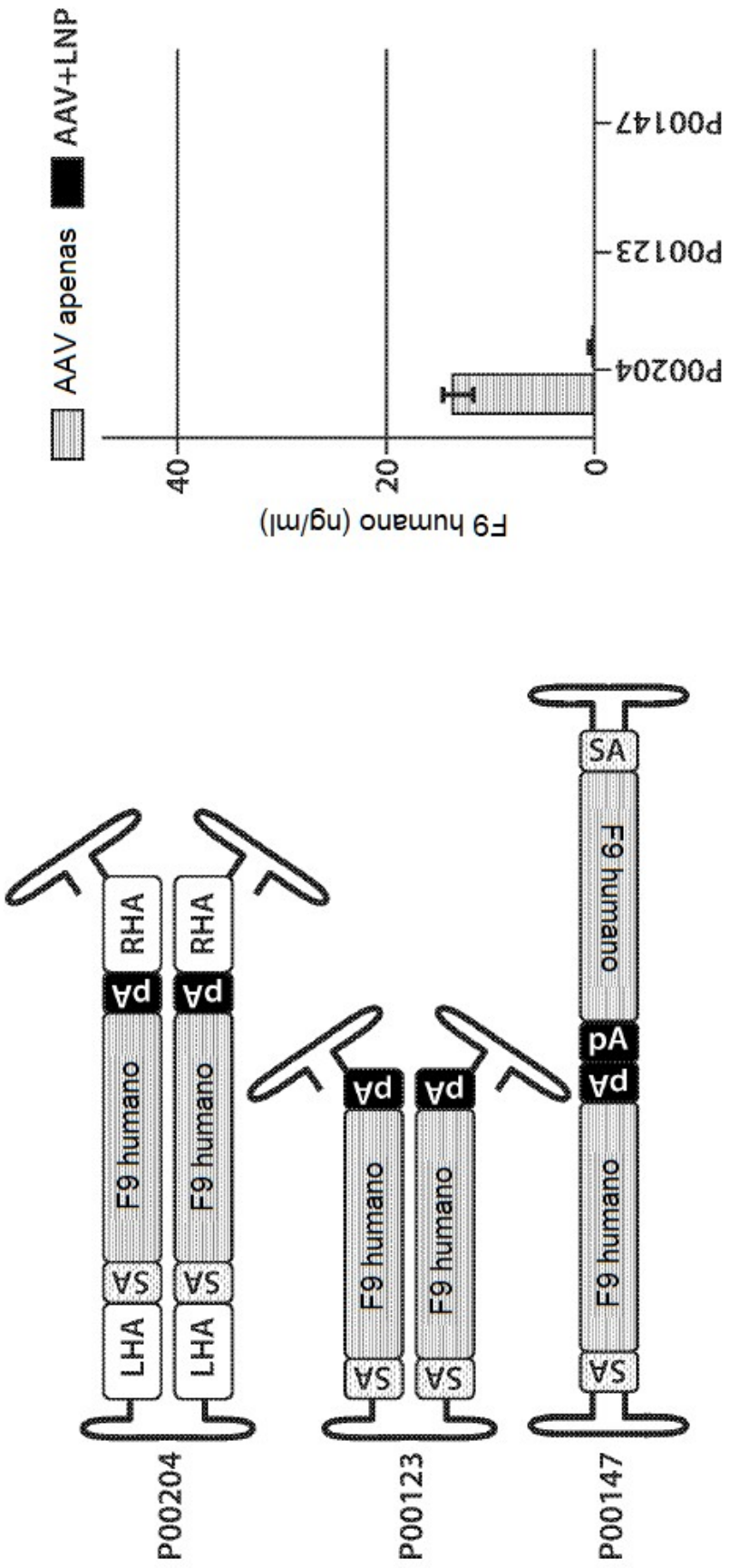

AAV apenas
AAV+LNP
F9 humano (ng/ml)
0
20
40
P00204
P00123
P00147
LHA
SA
F9 humano
pA
RHA
Vd

FIG. 2

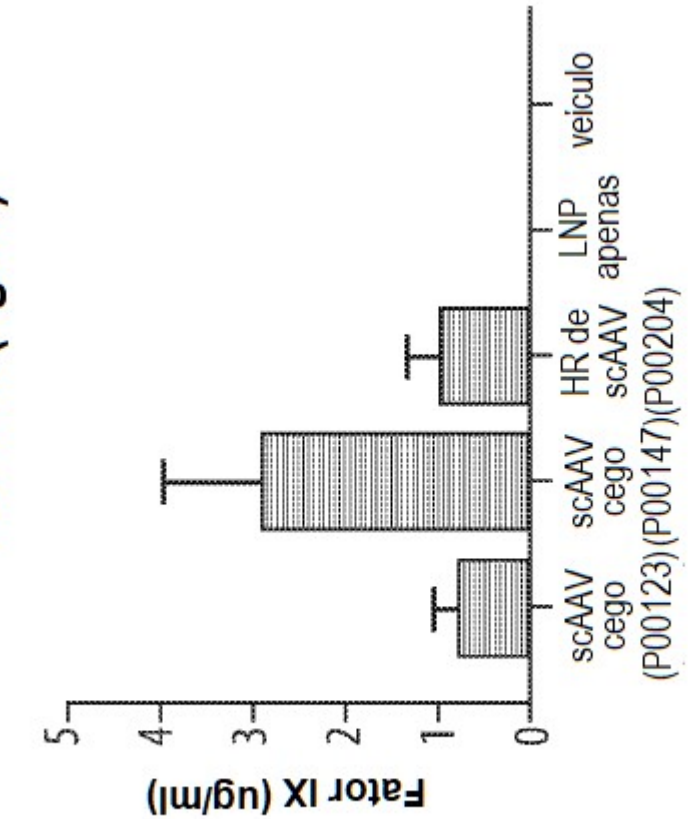

Fator IX (ug/ml)
Fator IX (ug/ml)
0
1
2
3
4
5
scAAV cego (P00123)
scAAV cego (P00147)
HR de scAAV (P00204)
LNP apenas
veículo

FIG. 3B

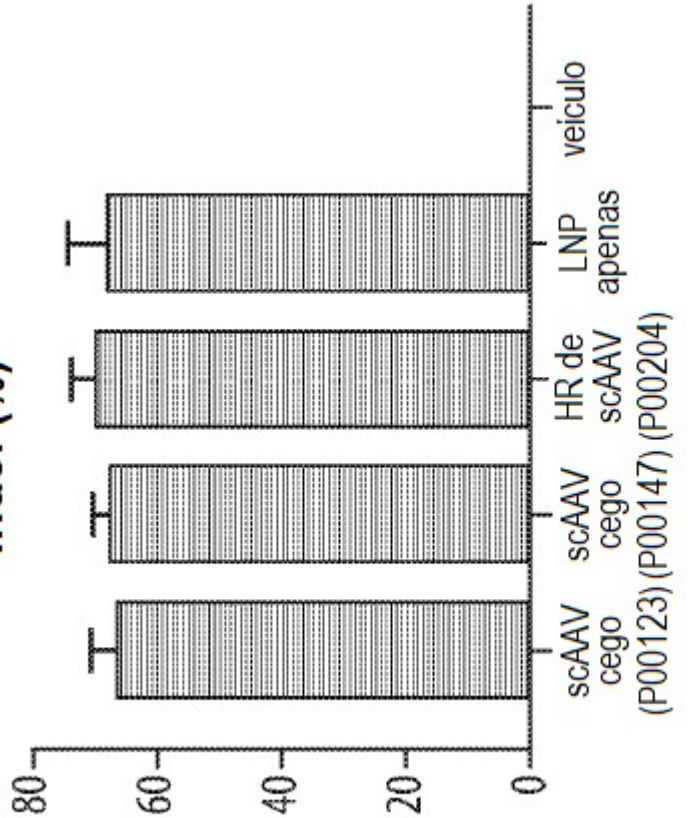

Indel (%)
Indel (%)
0
20
40
60
80
scAAV cego (P00123)
scAAV cego (P00147)
HR de scAAV (P00204)
LNP apenas
veículo

FIG. 3A

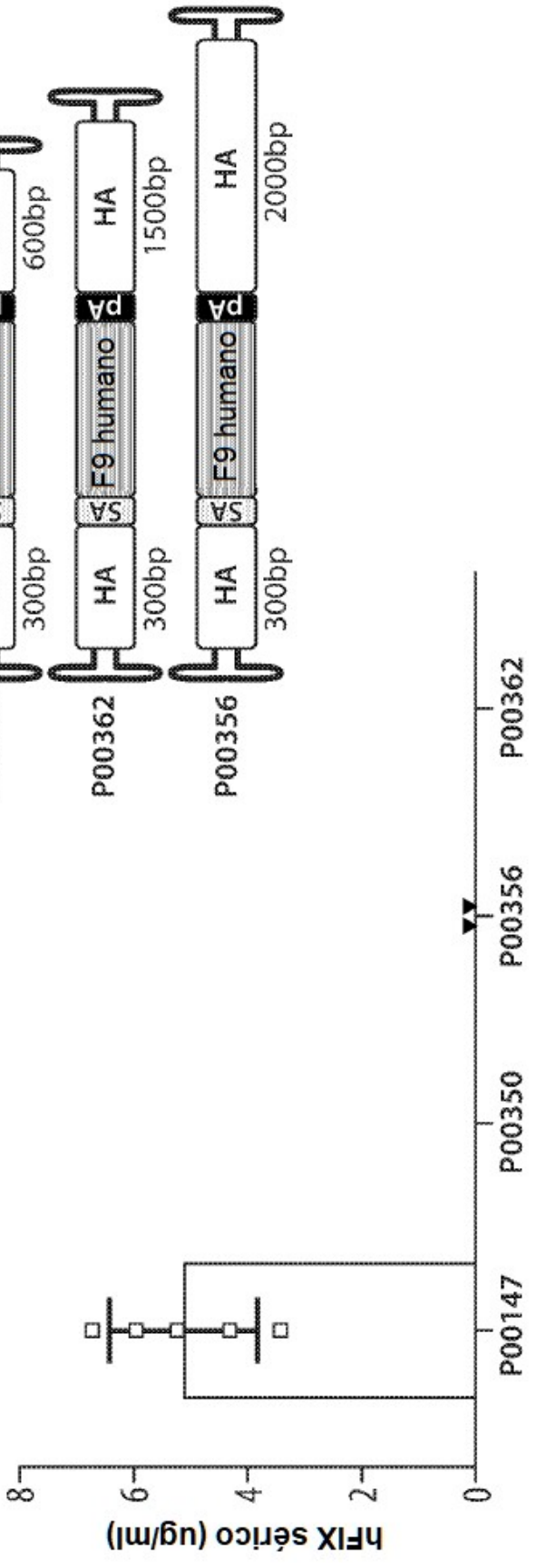

hFIX sérico (ug/ml)
0
2
4
6
8
P00147
P00350
P00356
P00362
P00350
HA
SA
F9 humano
pA
HA
300bp
600bp
P00362
HA
SA
F9 humano
pA
HA
300bp
1500bp
P00356
HA
SA
F9 humano
pA
HA
300bp
2000bp

FIG. 4A

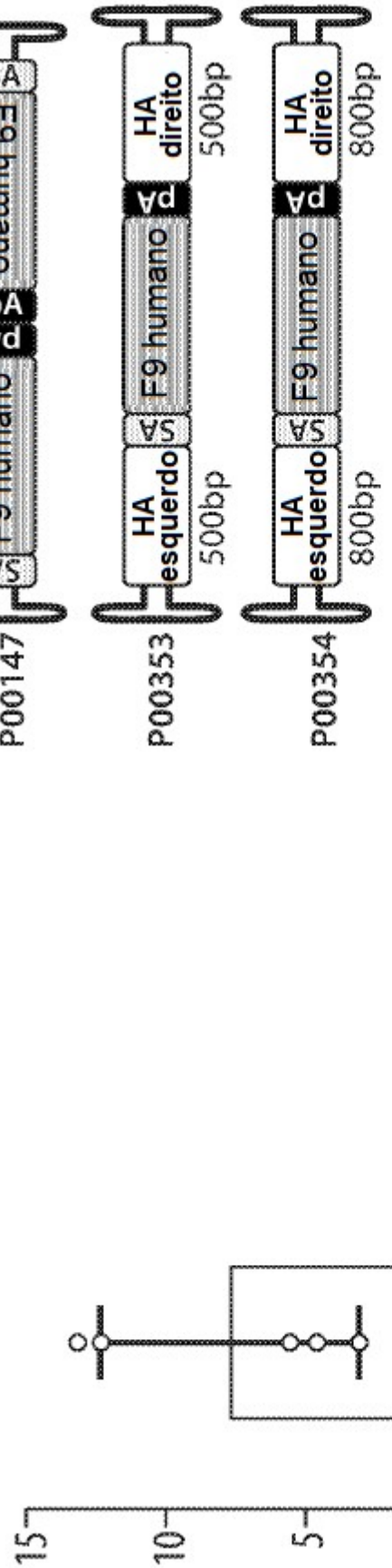

SA
F9 humano
pA
pA
F9 humano
SA
P00147
HA esquerdo
500bp
SA
F9 humano
pA
HA direito
500bp
P00353
HA esquerdo
800bp
SA
F9 humano
pA
HA direito
800bp
P00354
15
10
5
0
P00147
P00353
P00354

FIG. 4B

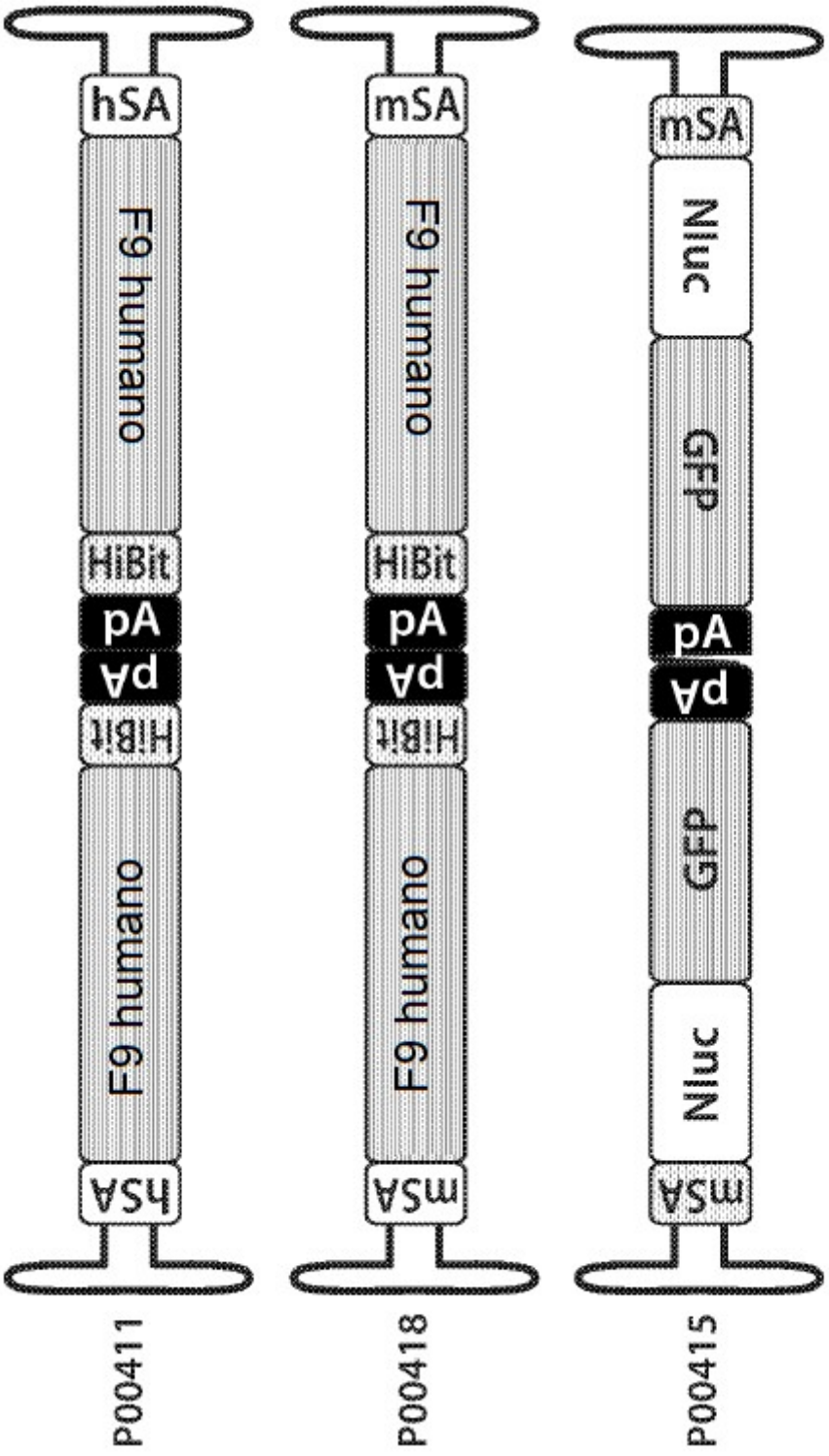

hSA
F9 humano
HiBit
pA
pA
HiBit
F9 humano
hSA
P00411
mSA
F9 humano
HiBit
pA
pA
HiBit
F9 humano
mSA
P00418
mSA
Nluc
GFP
pA
pA
GFP
Nluc
mSA
P00415

FIG. 5A

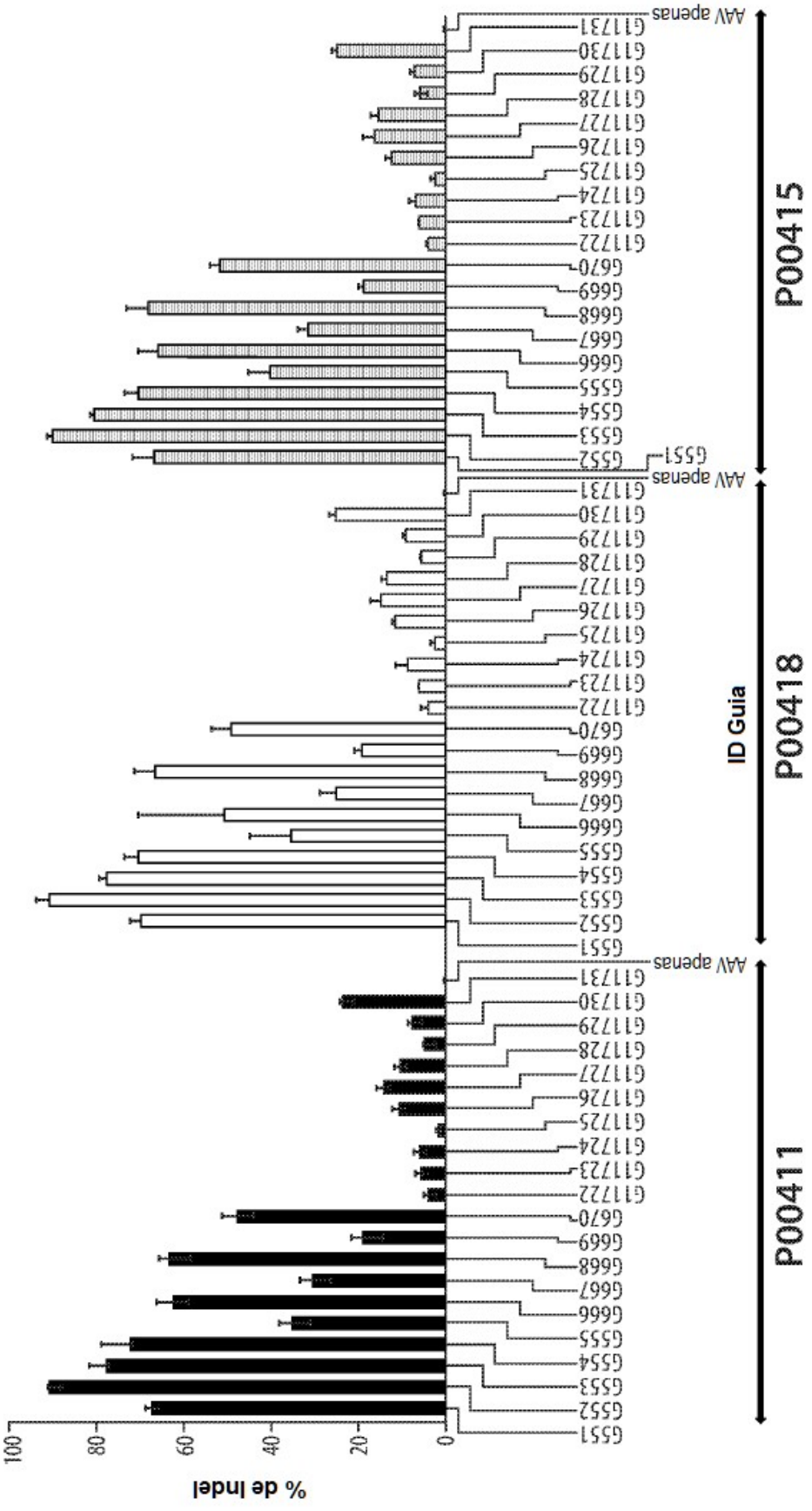

% de Indel
0
20
40
60
80
100
ID Guia
P00411
P00418
P00415
AAV apenas
G551
G552
G553
G554
G555
G666
G667
G668
G669
G670
G1722
G1723
G1724
G1725
G1726
G1727
G1728
G1729
G1730
G1731

FIG. 5B

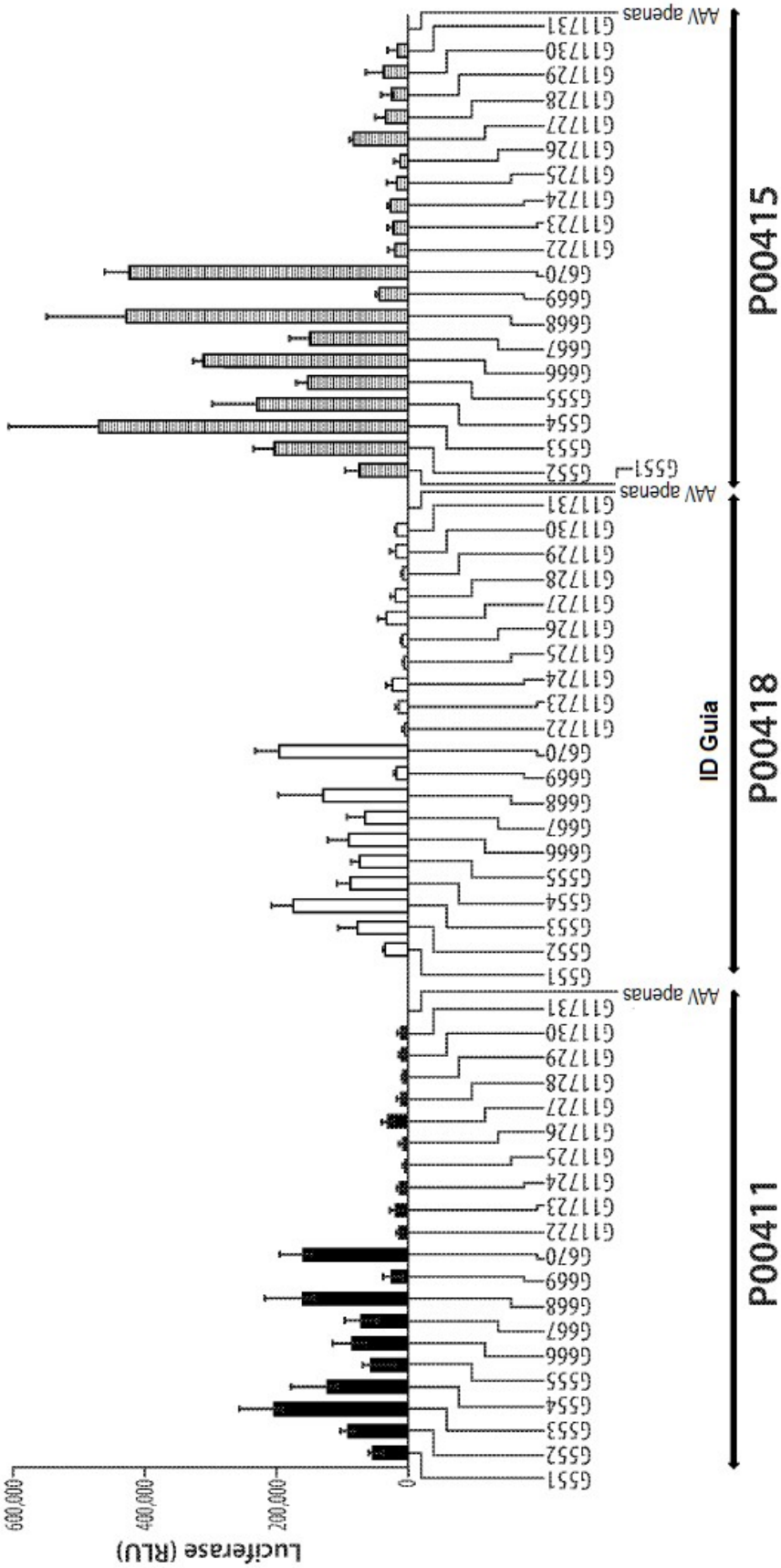

Luciferase (RLU)
600,000
400,000
200,000
0
G551
G552
G553
G554
G555
G666
G667
G668
G669
G670
G1722
G1723
G1724
G1725
G1726
G1727
G1728
G1729
G1730
G1731
AAV apenas
P00411
G551
G552
G553
G554
G555
G666
G667
G668
G669
G670
G1722
G1723
G1724
G1725
G1726
G1727
G1728
G1729
G1730
G1731
AAV apenas
P00418
ID Guia
G551
G552
G553
G554
G555
G666
G667
G668
G669
G670
G1722
G1723
G1724
G1725
G1726
G1727
G1728
G1729
G1730
G1731
AAV apenas
P00415

FIG. 5C

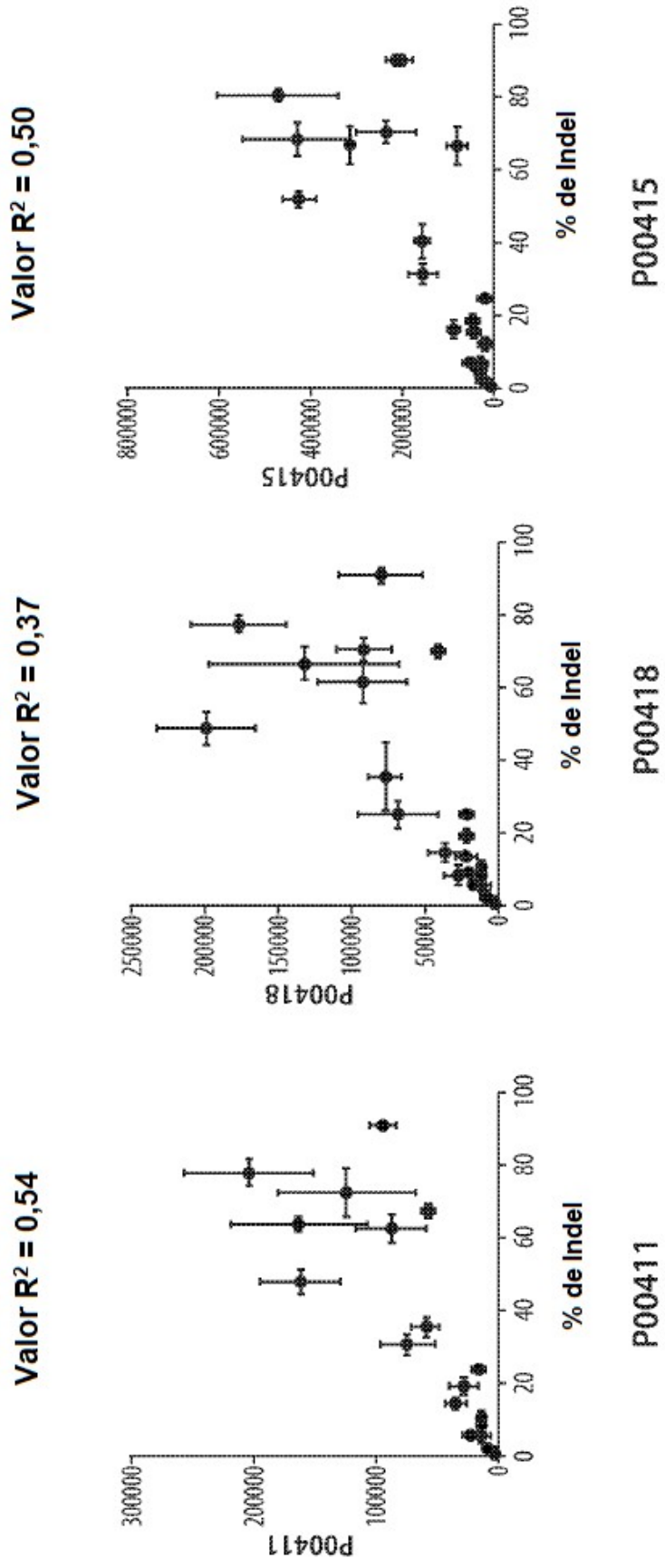

Valor R² = 0,54
P00411
% de Indel
P00411
Valor R² = 0,37
P00418
% de Indel
P00418
Valor R² = 0,50
P00415
% de Indel
P00415

FIG. 5D

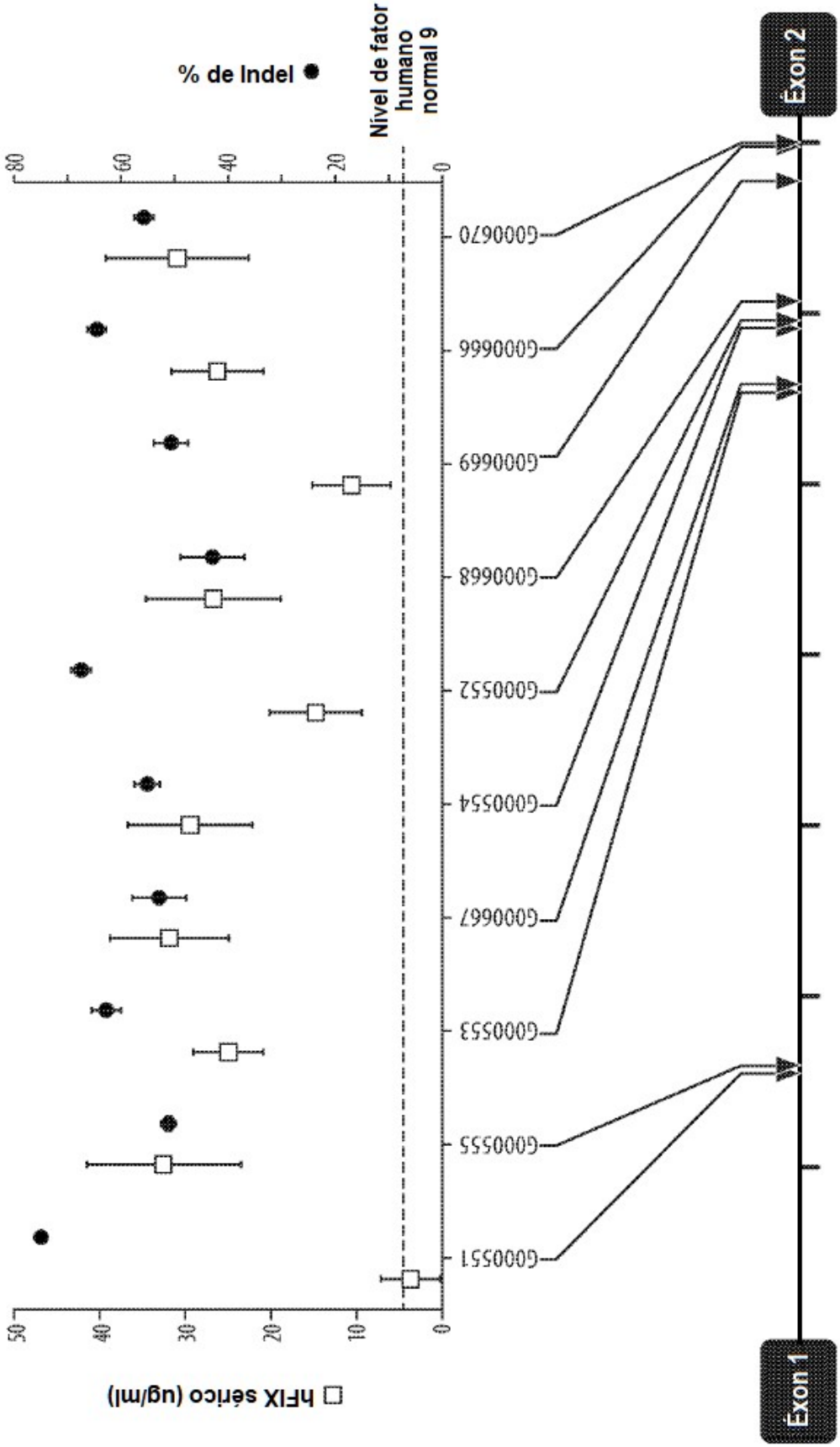

Éxon 1
Éxon 2
6000551
6000555
6000553
6000667
6000554
6000552
6000668
6000669
6000666
6000670
Nível de fator humano normal 9
% de Indel
□ hFIX sérico (ug/ml)
0
10
20
30
40
50
0
20
40
60
80

FIG. 6

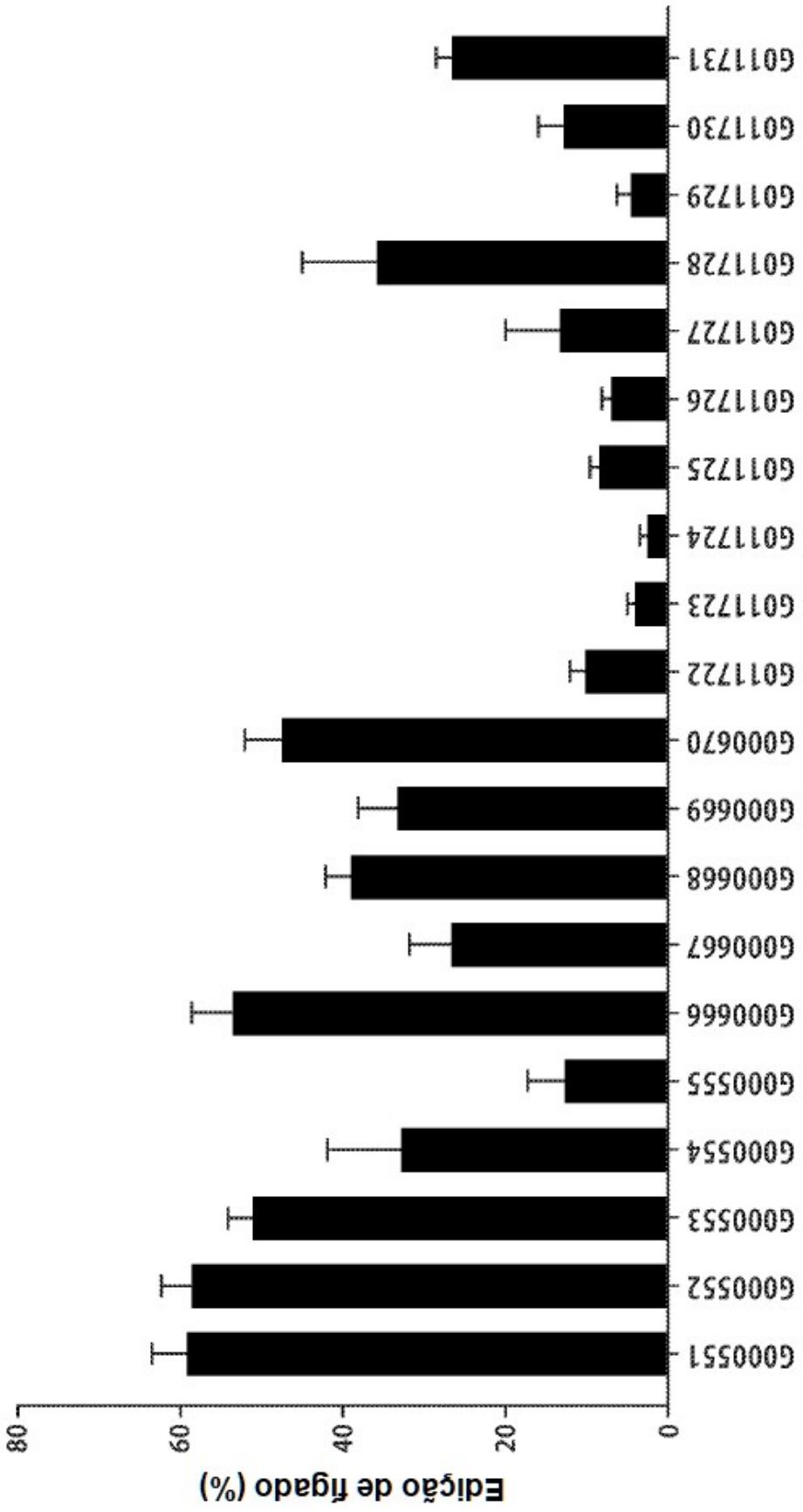

Edição de fígado (%)
0
20
40
60
80
6000551
6000552
6000553
6000554
6000555
6000666
6000667
6000668
6000669
6000670
6011722
6011723
6011724
6011725
6011726
6011727
6011728
6011729
6011730
6011731

FIG. 7A

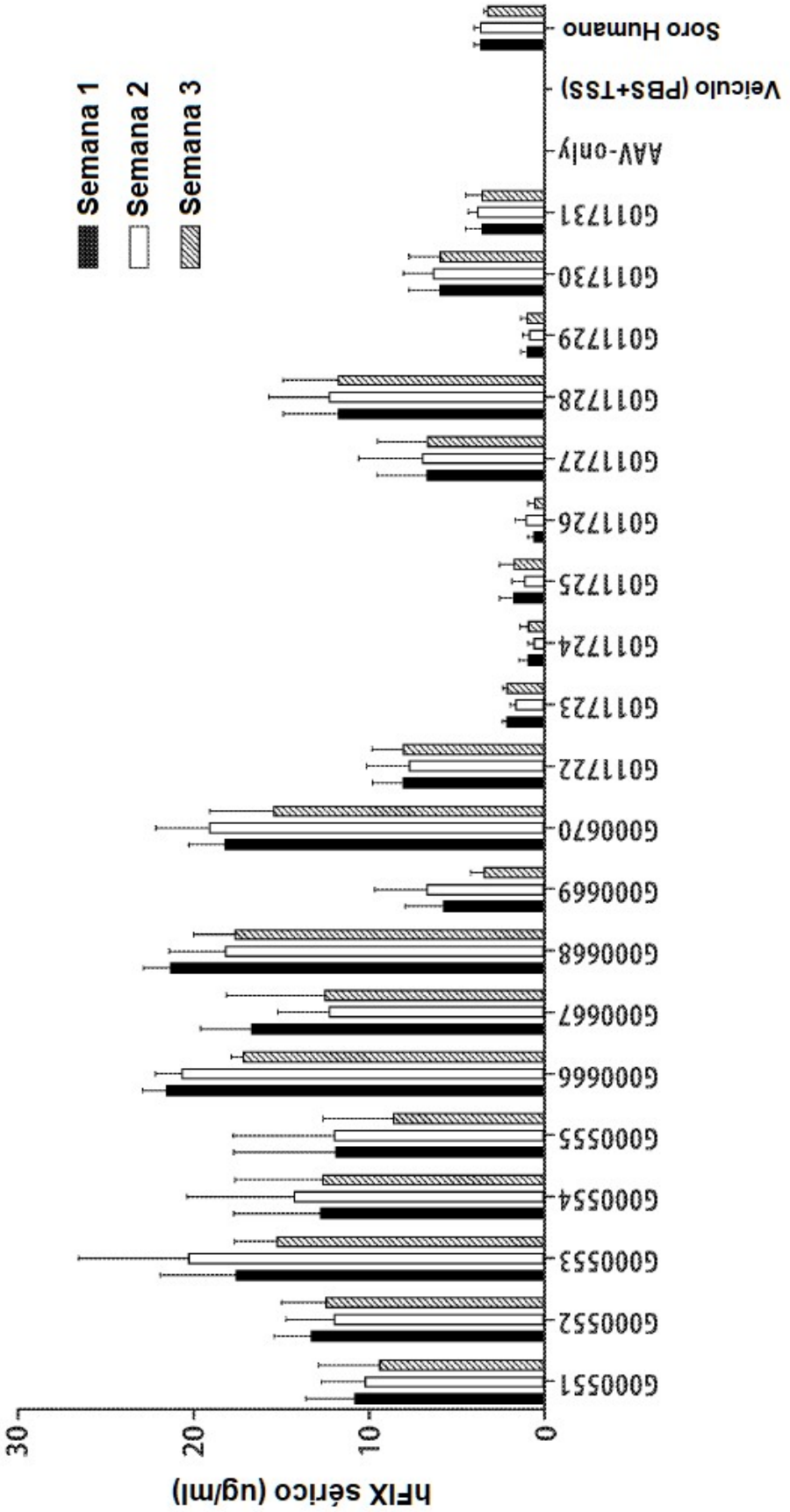

Semana 1
Semana 2
Semana 3
hFIX sérico (ug/ml)
0
10
20
30
G000551
G000552
G000553
G000554
G000555
G000666
G000667
G000668
G000669
G000670
G011722
G011723
G011724
G011725
G011726
G011727
G011728
G011729
G011730
G011731
AAV-only
Veículo (PBS+TSS)
Soro Humano

FIG. 7B

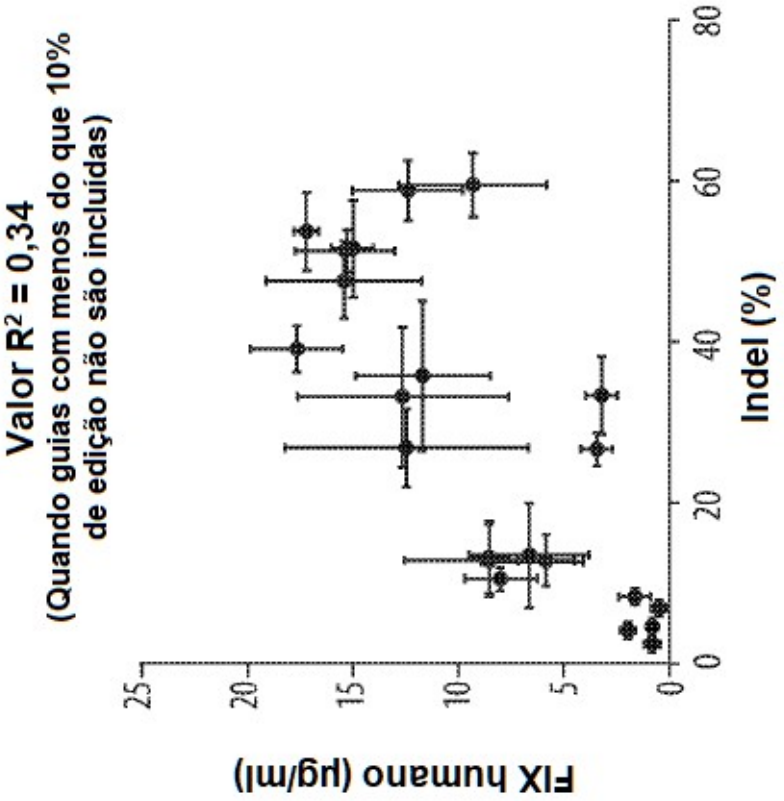

Valor R^2 = 0,34
(Quando guias com menos do que 10% de edição não são incluídas)
FIX humano (µg/ml)
Indel (%)
0
5
10
15
20
25
0
20
40
60
80

FIG. 7D

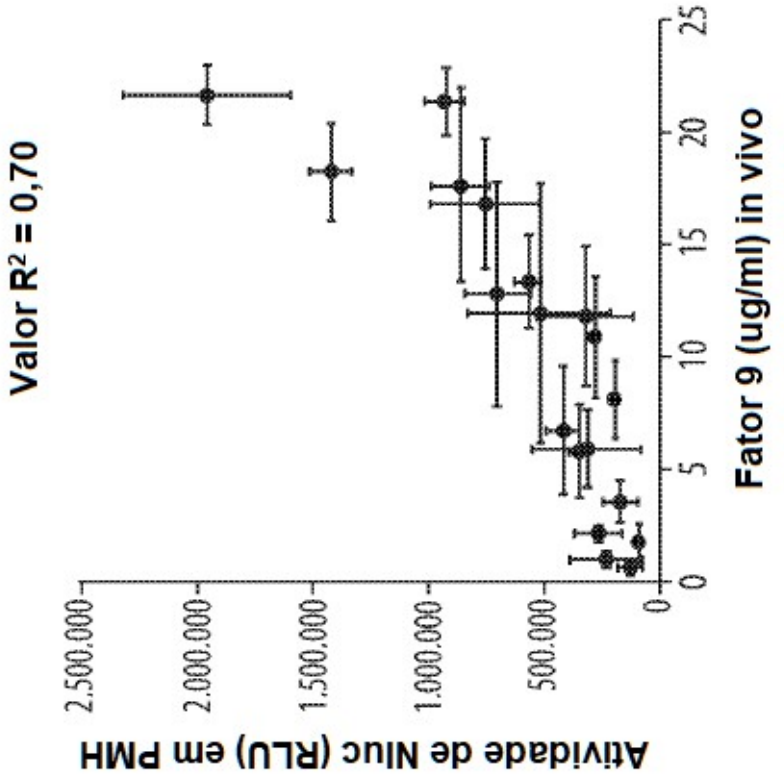

Valor R^2 = 0,70
Atividade de Nluc (RLU) em PMH
Fator 9 (ug/ml) in vivo
0
500.000
1.000.000
1.500.000
2.000.000
2.500.000
0
5
10
15
20
25

FIG. 7C

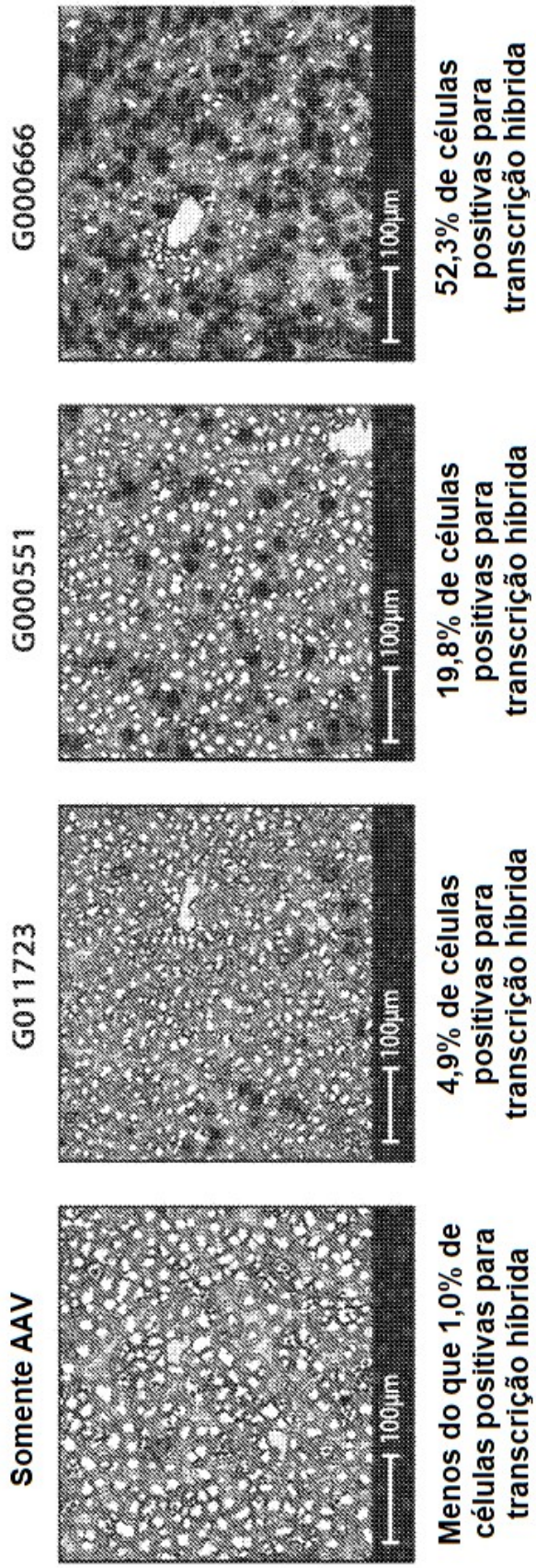

Somente AAV
G011723
G000551
G000666
100µm
100µm
100µm
100µm
Menos do que 1,0% de células positivas para transcrição híbrida
4,9% de células positivas para transcrição híbrida
19,8% de células positivas para transcrição híbrida
52,3% de células positivas para transcrição híbrida

FIG. 8A

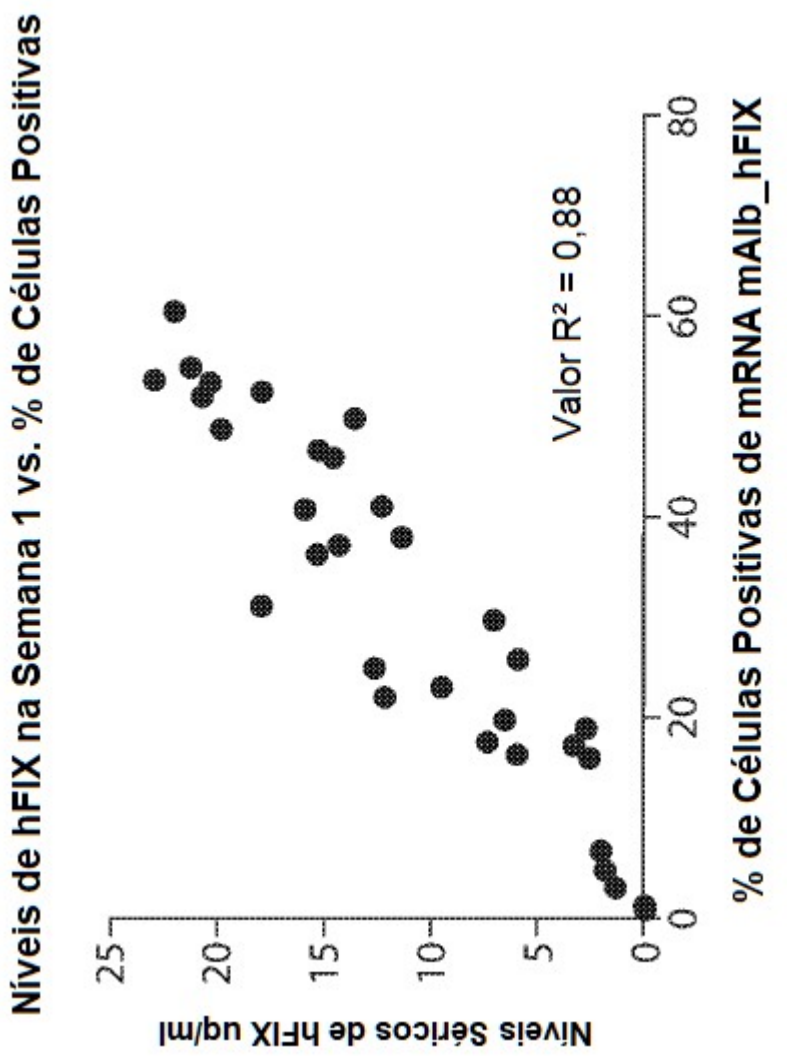

Níveis de hFIX na Semana 1 vs. % de Células Positivas
Níveis Séricos de hFIX ug/ml
25
20
15
10
5
0
0
20
40
60
80
% de Células Positivas de mRNA mAlb_hFIX
Valor R² = 0,88

FIG. 8B

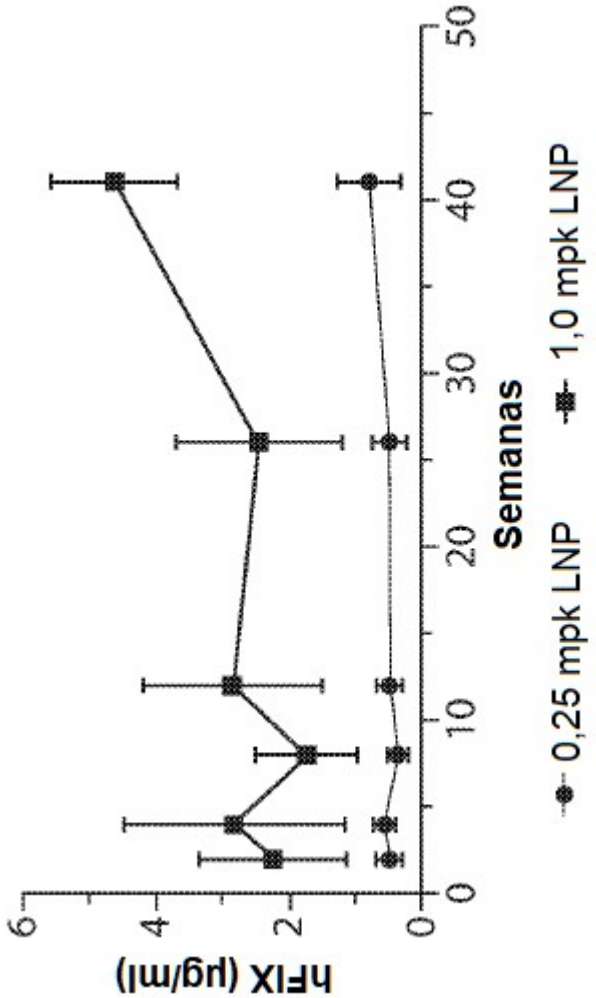


FIG. 9A

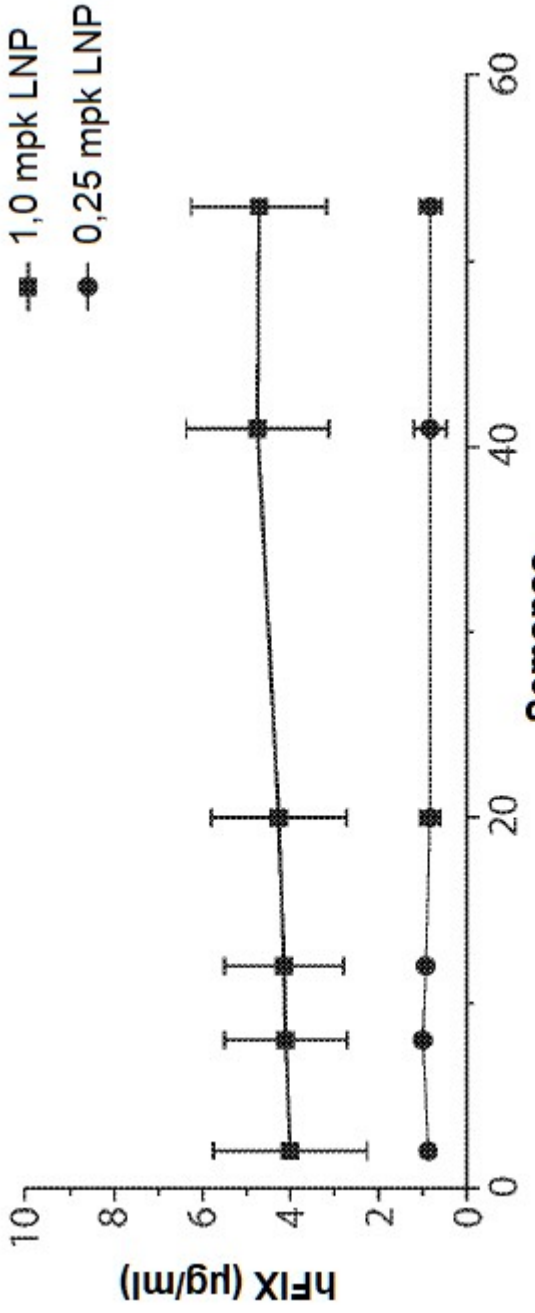


FIG. 9B

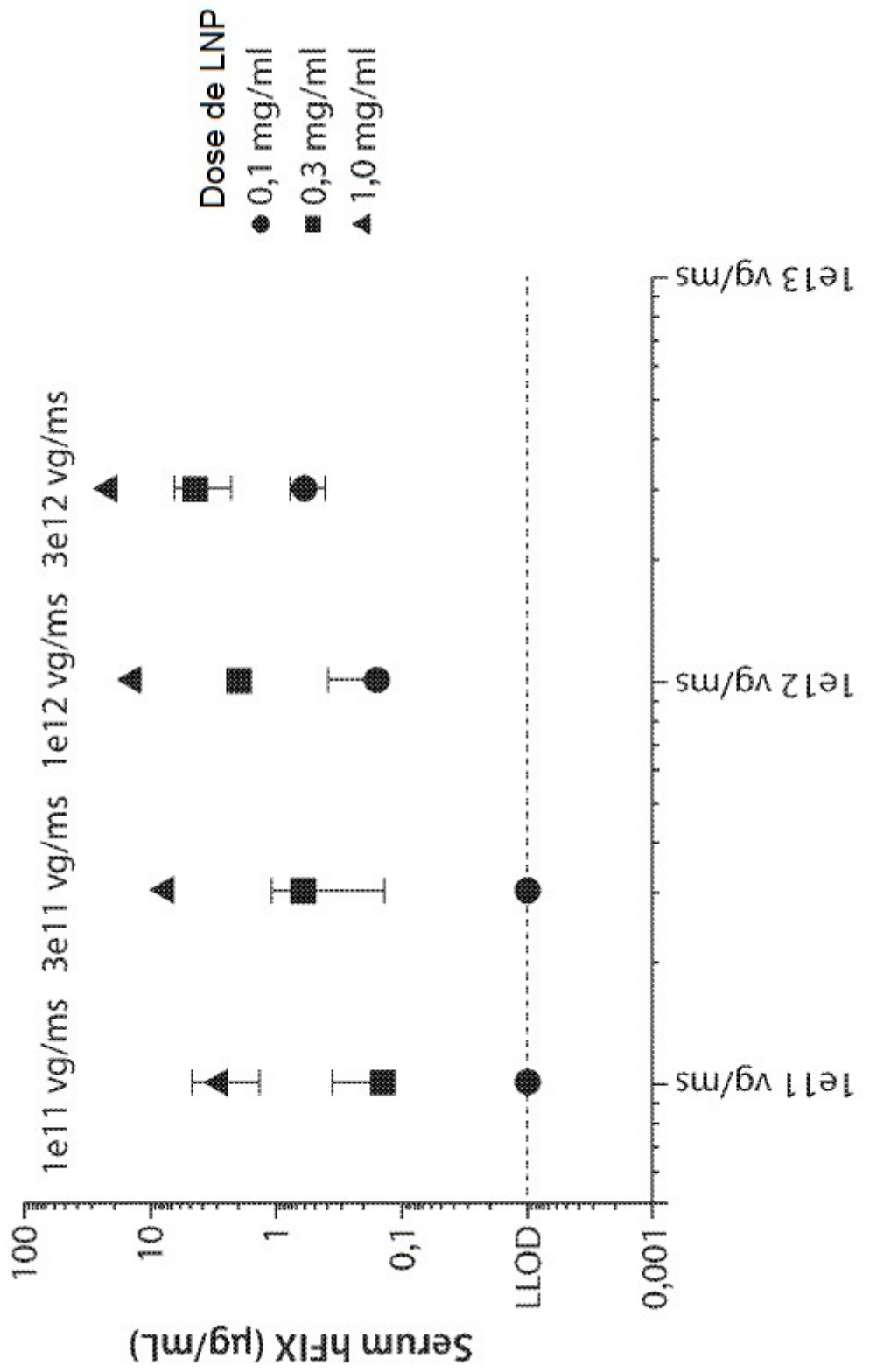

Dose de LNP
0,1 mg/ml
0,3 mg/ml
1,0 mg/ml
1e11 vg/ms
3e11 vg/ms
1e12 vg/ms
3e12 vg/ms
100
10
1
0,1
LLOD
0,001
Serum hFIX (µg/mL)
1e11 vg/ms
1e12 vg/ms
1e13 vg/ms

FIG. 10A

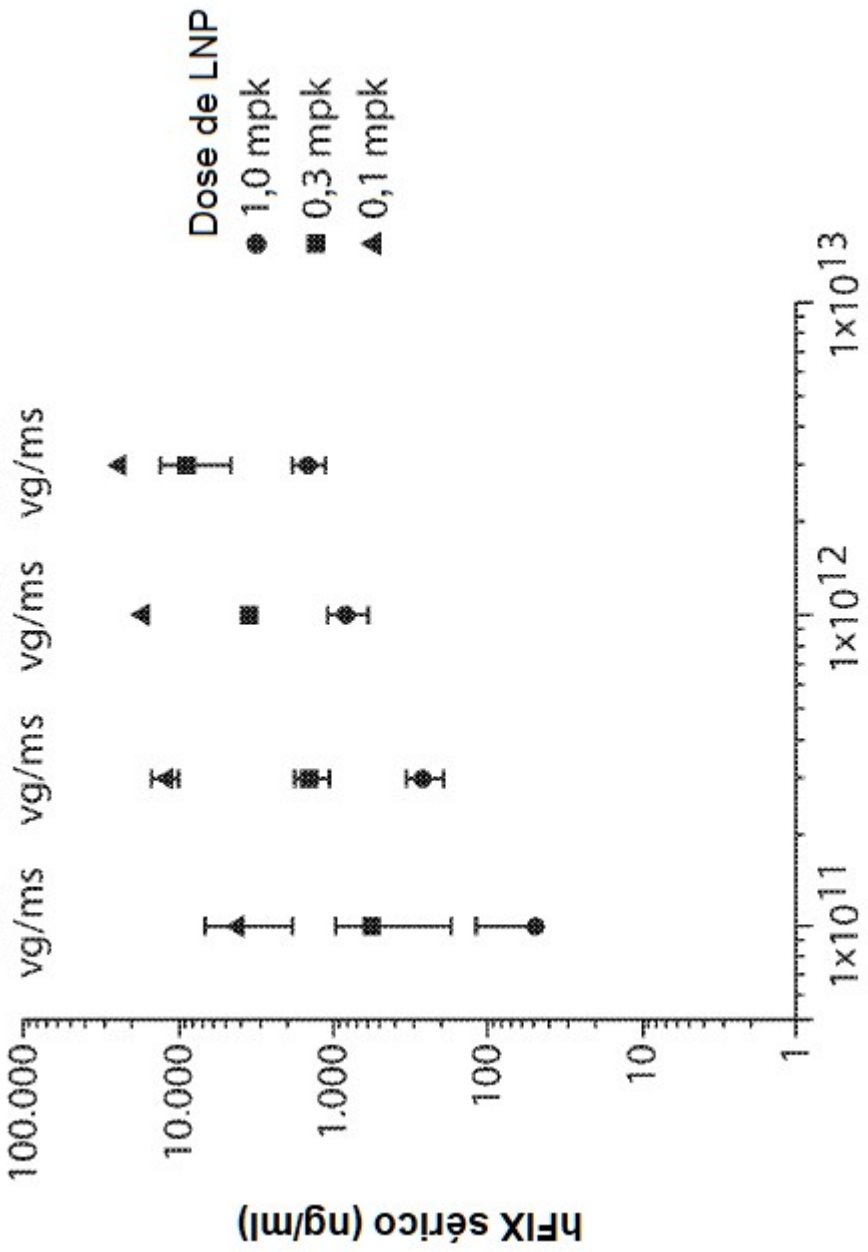

hFIX sérico (ng/ml)
100.000
10.000
1.000
100
10
1
1x10^11
1x10^12
1x10^13
AAV (vg/camundongo)
1e11 vg/ms
3e11 vg/ms
1e12 vg/ms
3e12 vg/ms
Dose de LNP
1,0 mpk
0,3 mpk
0,1 mpk

FIG. 10B

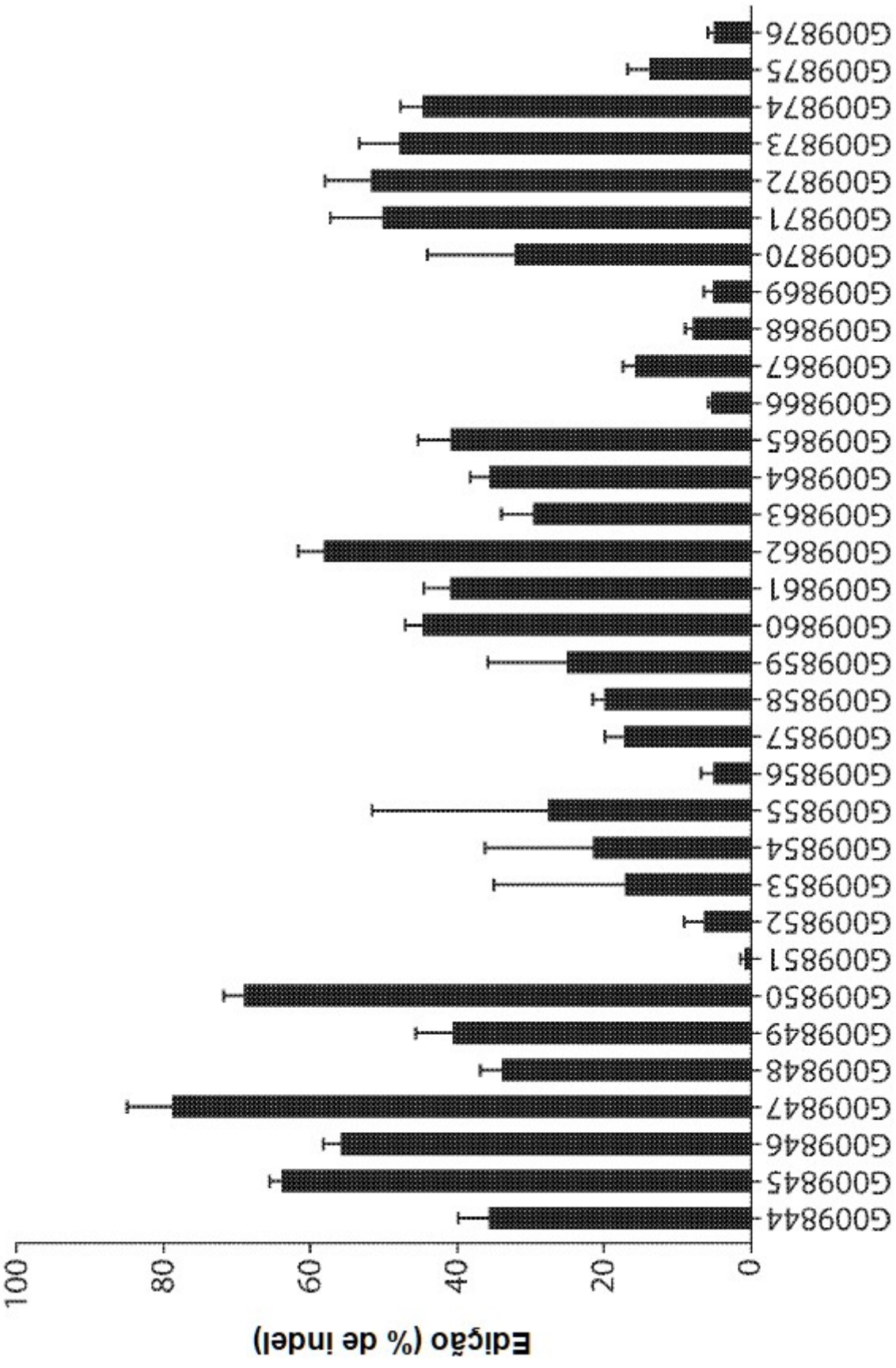

Edição (% de indel)
0
20
40
60
80
100
G009844
G009845
G009846
G009847
G009848
G009849
G009850
G009851
G009852
G009853
G009854
G009855
G009856
G009857
G009858
G009859
G009860
G009861
G009862
G009863
G009864
G009865
G009866
G009867
G009868
G009869
G009870
G009871
G009872
G009873
G009874
G009875
G009876

FIG. 11A

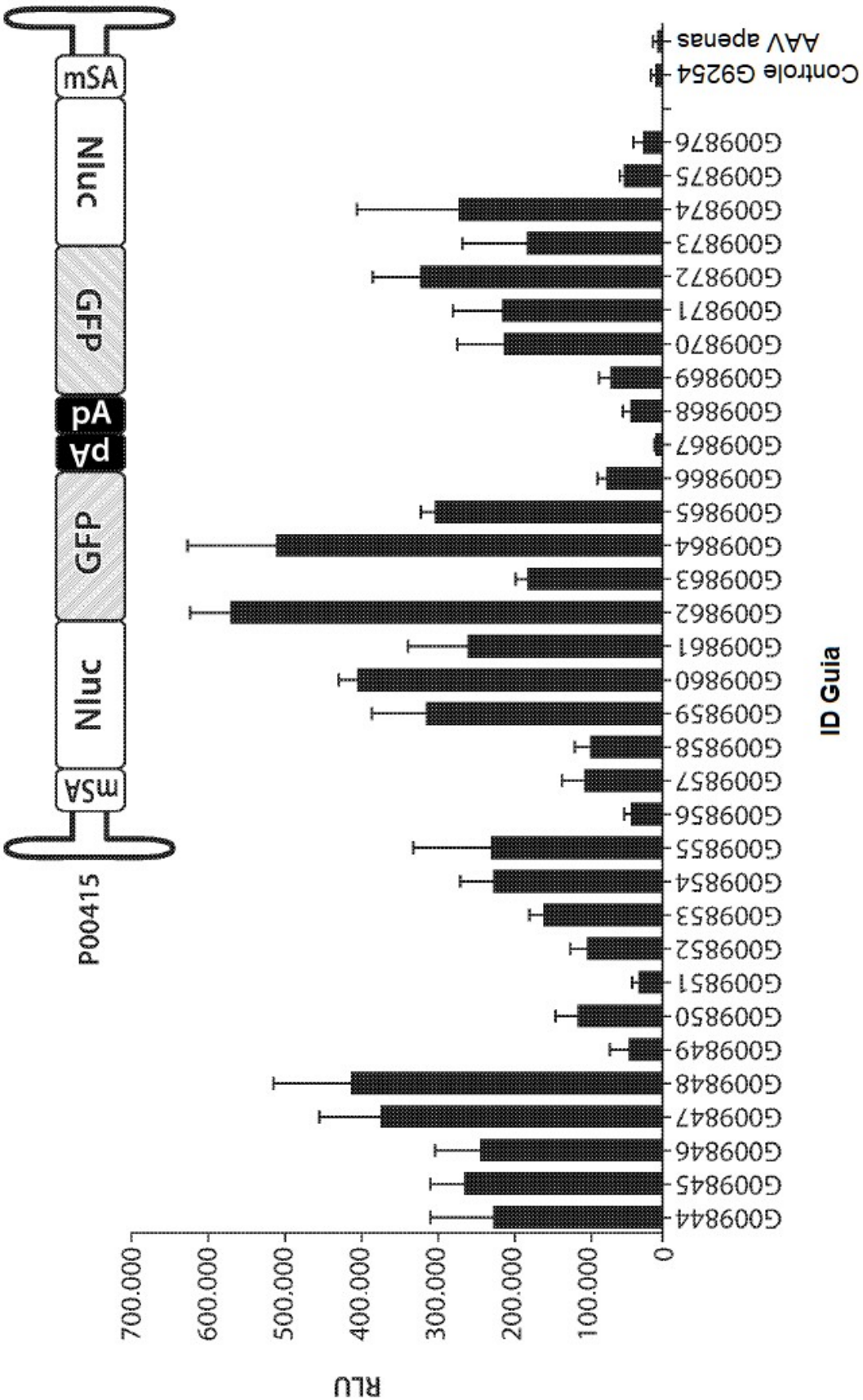

mSA
Nluc
GFP
pA
pA
GFP
Nluc
mSA
P00415
RLU
700.000
600.000
500.000
400.000
300.000
200.000
100.000
0
ID Guia
Controle G9254
AAV apenas
G009844
G009845
G009846
G009847
G009848
G009849
G009850
G009851
G009852
G009853
G009854
G009855
G009856
G009857
G009858
G009859
G009860
G009861
G009862
G009863
G009864
G009865
G009866
G009867
G009868
G009869
G009870
G009871
G009872
G009873
G009874
G009875
G009876

FIG. 11B

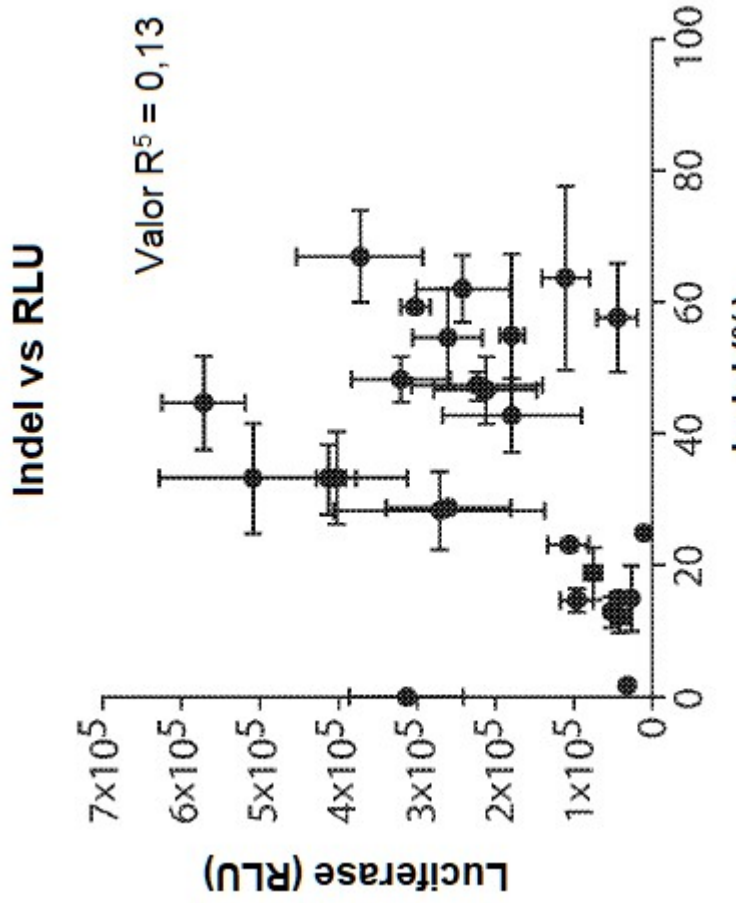

Indel vs RLU
Valor R^5 = 0,13
Luciferase (RLU)
7x10^5
6x10^5
5x10^5
4x10^5
3x10^5
2x10^5
1x10^5
0
0
20
40
60
80
100
Indel (%)

FIG. 11C

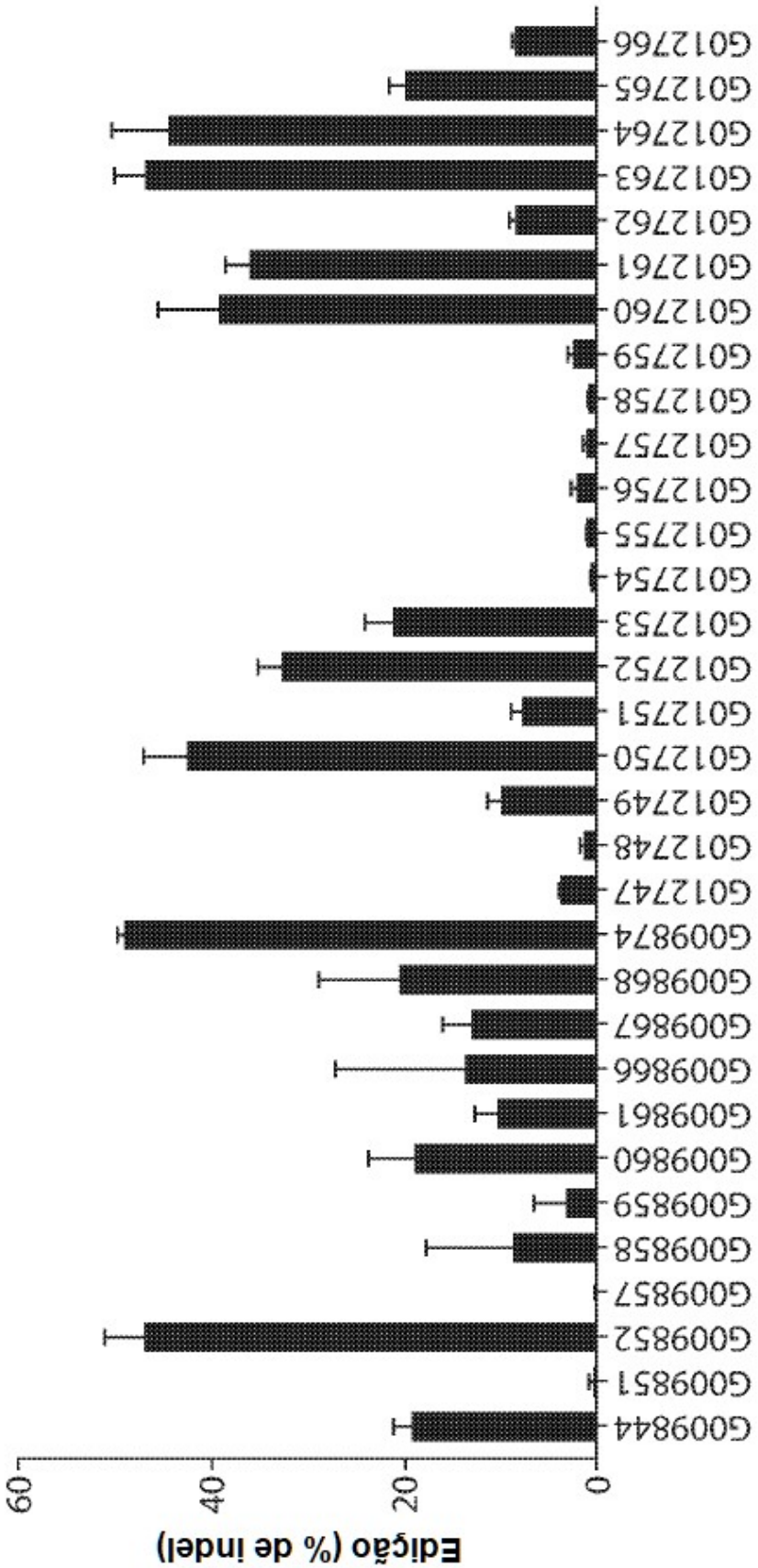

Edição (% de indel)
0
20
40
60
G009844
G009851
G009852
G009857
G009858
G009859
G009860
G009861
G009866
G009867
G009868
G009874
G012747
G012748
G012749
G012750
G012751
G012752
G012753
G012754
G012755
G012756
G012757
G012758
G012759
G012760
G012761
G012762
G012763
G012764
G012765
G012766

FIG. 12A

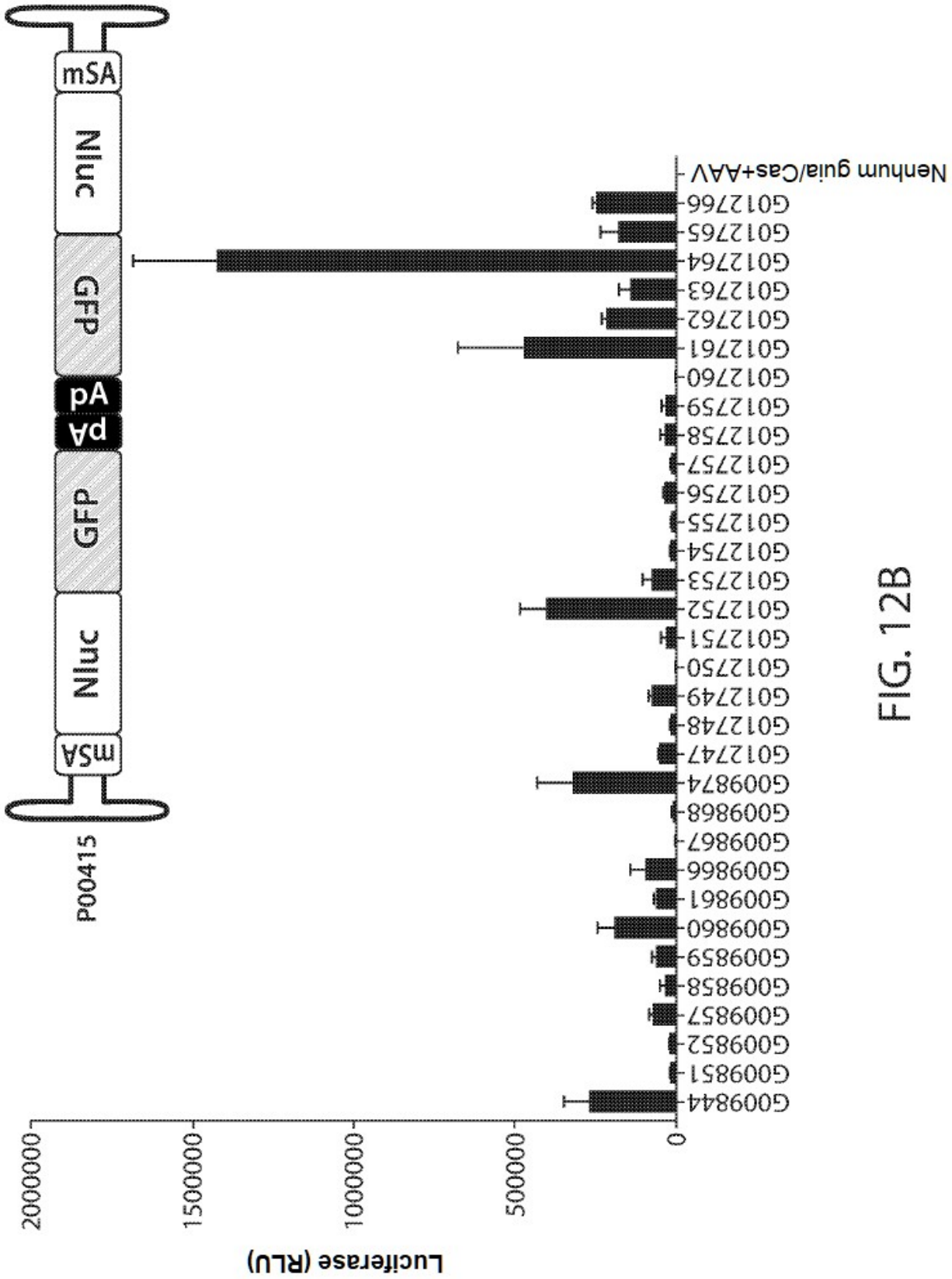

mSA
NLuc
GFP
pA
pA
GFP
Nluc
mSA
P00415
Luciferase (RLU)
2000000
1500000
1000000
500000
0
Nenhum guia/Cas+AAV
G012766
G012765
G012764
G012763
G012762
G012761
G012760
G012759
G012758
G012757
G012756
G012755
G012754
G012753
G012752
G012751
G012750
G012749
G012748
G012747
G009874
G009868
G009867
G009866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

FIG. 12B

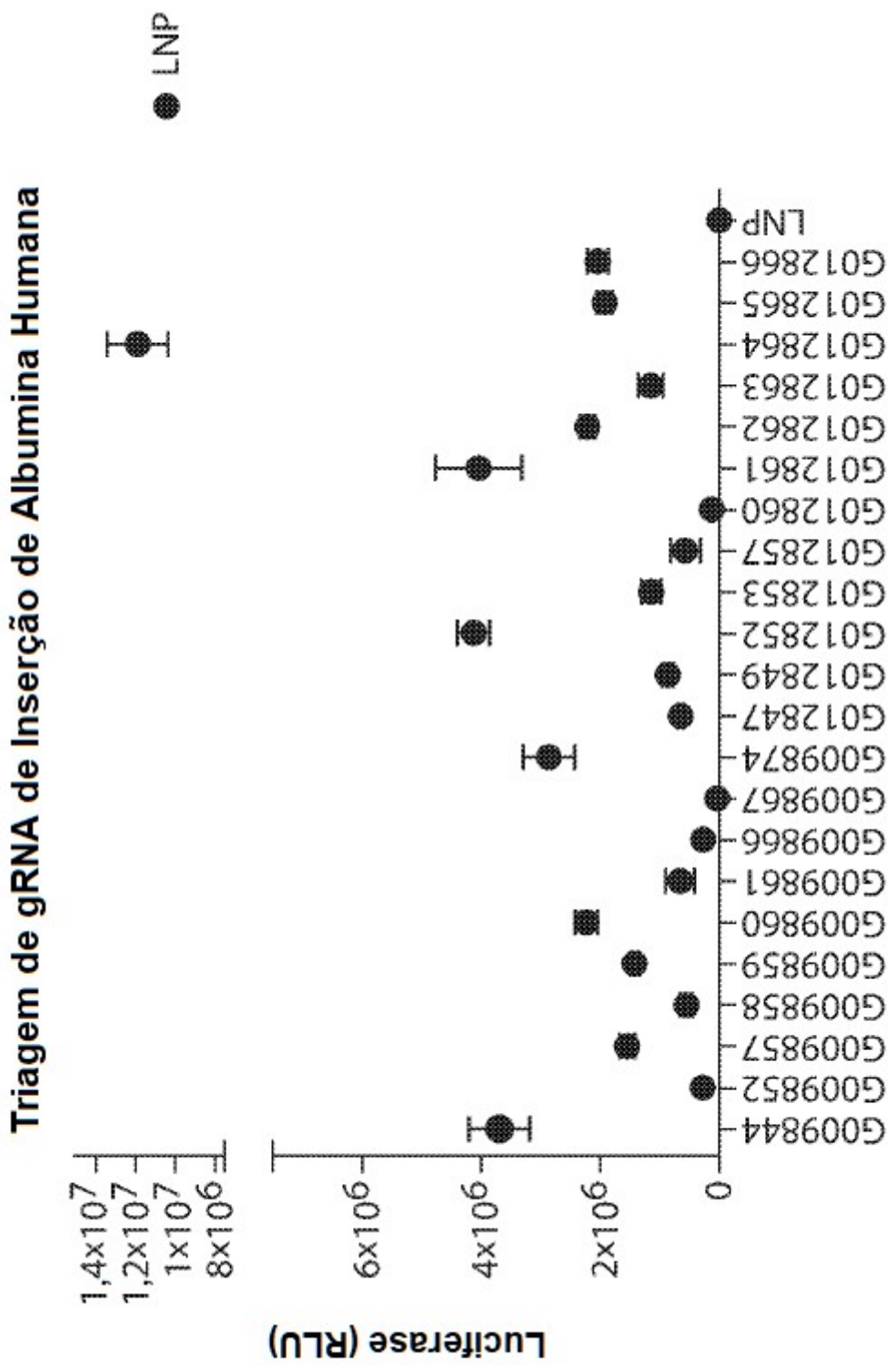

Triagem de gRNA de Inserção de Albumina Humana
LNP
Luciferase (RLU)
1,4x10^7
1,2x10^7
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
G009844
G009852
G009857
G009858
G009859
G009860
G009861
G009866
G009867
G009874
G012847
G012849
G012852
G012853
G012857
G012860
G012861
G012862
G012863
G012864
G012865
G012866
LNP

FIG. 12C

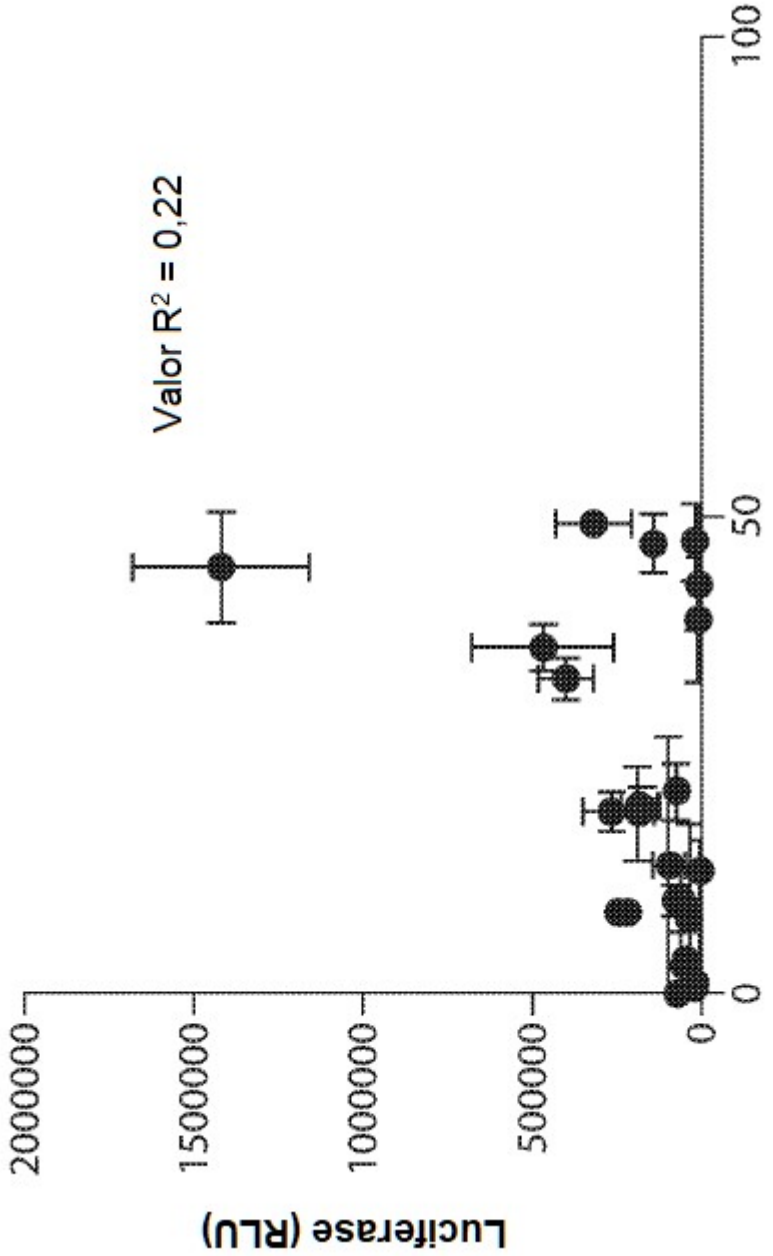

Indel vs RLU
Valor R^2 = 0,22
Luciferase (RLU)
2000000
1500000
1000000
500000
0
0
50
100
Indel (%)

FIG. 12D

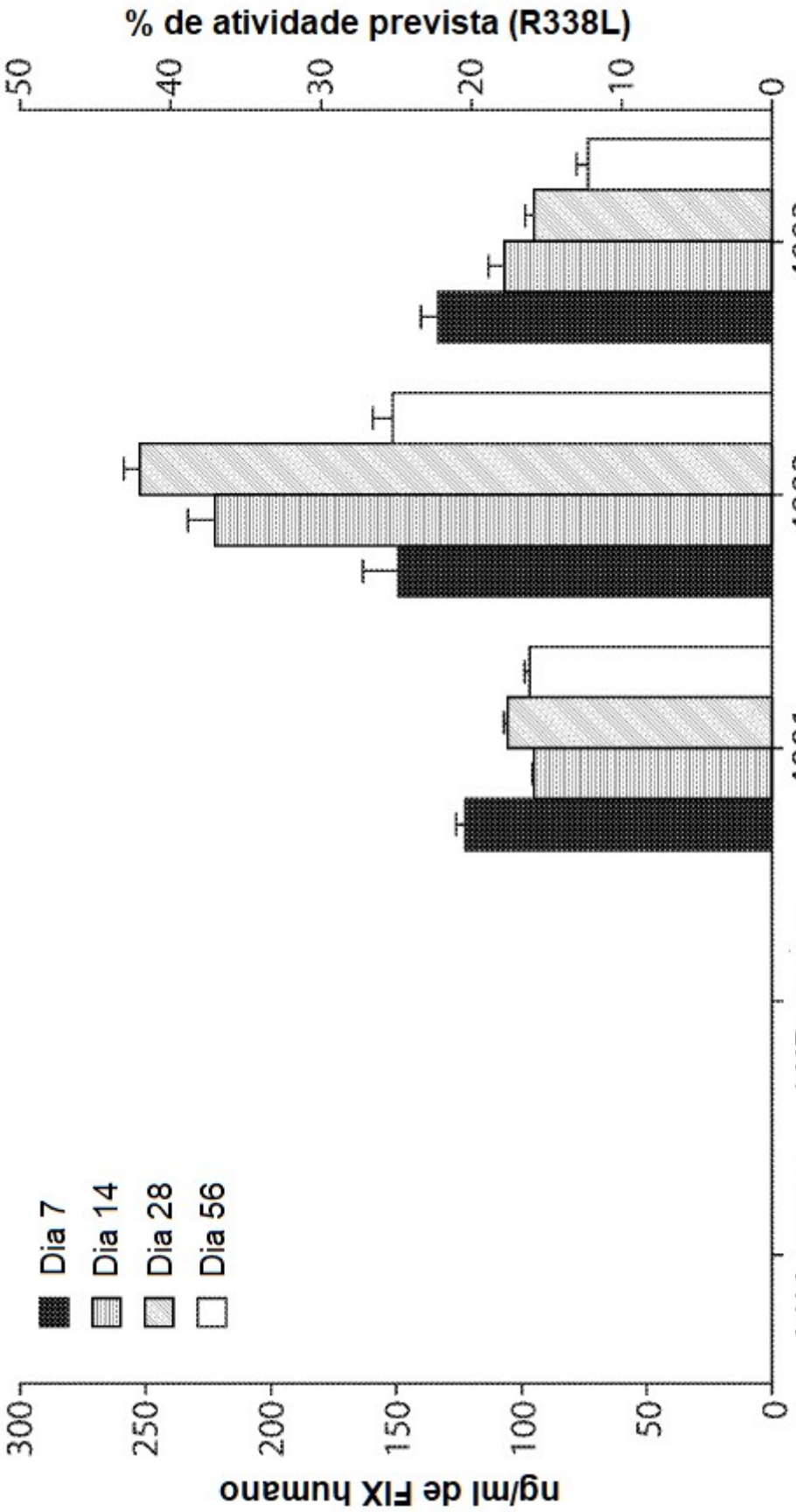

% de atividade prevista (R338L)
50
40
30
20
10
0
ng/ml de FIX humano
300
250
200
150
100
50
0
Dia 7
Dia 14
Dia 28
Dia 56
AAV apenas
LNP apenas
4001
4002
4003

FIG. 13

RESUMO

“CONSTRUTOS DE ÁCIDO NUCLEICO E MÉTODOS PARA USO”

Trata-se de construtos de ácido nucleico que permitem a inserção e/ou expressão de uma sequência de interesse, como um transgene. Composições e métodos para usar tais construtos para expressão de um polipeptídeo ou agente terapêutico, por exemplo, também são fornecidos.

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

Código de Controle

Campo 1



Campo 2



Outras Informações:

- Nome do Arquivo: 213.368-7 Listagem de sequências.txt
- Data de Geração do Código: 15/06/2021
- Hora de Geração do Código: 16:54:49
- Código de Controle:
  - Campo 1: D022ADF71B18479C
  - Campo 2: A1290E374137C19D