



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

(21) BR 112020017278-2 A2



(22) Data do Depósito: 27/02/2019

(43) Data da Publicação Nacional: 22/12/2020

(54) **Título:** VETORES DE VÍRUS ADENOASSOCIADO (AAV), VETORES DE AAV TENDO DESAMIDAÇÃO DE CAPSÍDEOS REDUZIDA E USOS PARA OS MESMOS

(51) **Int. Cl.:** C07K 14/005; C12N 15/864.

(30) **Prioridade Unionista:** 06/05/2018 US 62/667,587; 07/05/2018 US 62/667,881; 07/05/2018 US 62/667,888; 24/08/2018 US 62/722,382; 24/08/2018 US 62/722,388; (...).

(71) **Depositante(es):** THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA.

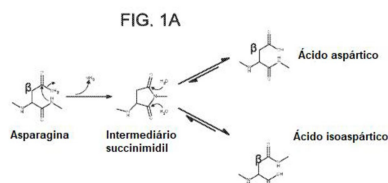
(72) **Inventor(es):** JAMES M. WILSON; APRIL TEPE; KEVIN TURNER; JOSHUA JOYNER SIMS; QIANG WANG.

(86) **Pedido PCT:** PCT US2019019861 de 27/02/2019

(87) **Publicação PCT:** WO 2019/169004 de 06/09/2019

(85) **Data da Fase Nacional:** 24/08/2020

(57) **Resumo:** A presente invenção refere-se a um vetor de vírus adenoassociado recombinante (rAAV) compreendendo um capsídeo de AAV tendo uma população heterogênea de proteínas vp1, uma população heterogênea de proteína vp2 e uma população heterogênea de proteínas vp3. O capsídeo contém aminoácidos modificados em comparação com a sequência de aminoácidos VP1 codificada, o capsídeo contendo resíduos de asparagina altamente desamidados no par asparagina-glicina, e compreendendo ainda vários outros resíduos de asparagina menos desamidada e opcionalmente de glutamina.



Relatório Descritivo da Patente de Invenção para “**VETORES DE VÍRUS ADENOASSOCIADO (AAV), VETORES DE AAV TENDO DESAMIDAÇÃO DE CAPSÍDEOS REDUZIDA E USOS PARA OS MESMOS**”.

DECLARAÇÃO DE PESQUISA PATROCINADA PELO GOVERNO FEDERAL

[001] Esta invenção foi feita com o apoio do governo sob o P01HL059407 concedido pelo National Institute of Health. O governo tem determinados direitos sobre a invenção.

REFERÊNCIA A UMA LISTAGEM DE SEQUÊNCIAS ENVIADA VIA EFS-WEB

[002] O conteúdo do texto arquivado da listagem de sequências denominada "18-8592PCT\_Sequence\_Listing\_ST25", que foi criado em 27 de fevereiro de 2019 e enviado eletronicamente via EFS-Web com o presente pedido, é incorporado neste documento por referência em sua totalidade.

FUNDAMENTOS DA INVENÇÃO

[003] O capsídeo do vírus adenoassociado (AAV) é icosaédrico na estrutura e é composto por 60 monômeros de proteína viral (VP) (VP1, VP2 e VP3) em uma razão de 1:1:10 (Xie Q, et al. *Proc Natl Acad Sci USA*. 2002; 99(16):10405-10). A totalidade da sequência da proteína VP3 (~ 535aa) está contida no terminal C de VP1 e VP2, e as sequências VP3 compartilhadas são principalmente responsáveis pela estrutura geral do capsídeo. Devido à flexibilidade estrutural das regiões únicas VP1/VP2 e à baixa representação dos monômeros VP1 e VP2 em relação aos monômeros VP3 no capsídeo montado, VP3 é a única proteína do capsídeo a ser resolvida por cristalografia de raios-x (Nam HJ, et al. *J Virol*. 2007; 81(22):12260-71). VP3 contém nove regiões hipervariáveis (HVRs) que são a fonte primária de variação de sequência entre sorotipos AAV (Govindasamy L, et al. *J Virol*. 2013; 87(20):11187-99).

Dada sua flexibilidade e localização na superfície do capsídeo, os HVRs são amplamente responsáveis pelas interações com as células-alvo, bem como com o sistema imunológico (Huang LY, et al. *J Virol.* 2016; 90(11):5219-30; Raupp C, et al. *J Virol.* 2012; 86(17):9396-408). Embora as estruturas de uma série de sorotipos sejam publicadas (Protein Data Bank (PDB) IDs 1LP3, 4RSO, 4V86, 3UX1, 3KIC, 2QA0, 2G8G do banco de dados do Research Collaboratory for Structural Bioinformatics (RCSB)) para as entradas de estrutura para AAV2, AAVrh.8, AAV6, AAV9, AAV3B, AAV8 e AAV4, respectivamente), há muito pouca informação na literatura a respeito de modificações na superfície desses capsídeos. A pesquisa sugere que a fosforilação intracelular do capsídeo ocorre em resíduos de tirosina específicos (Zhong L, et al. *Virology.* 2008; 381(2):194-202). Apesar dos sítios de glicosilação putativos na sequência VP3 primária, nenhum evento de glicosilação foi identificado em AAV2 (Murray S, et al. *J Virol.* 2006; 80(12):6171-6; Jin X, et al. *Hum Gene Ther Methods.* 2017; 28(5):255-267); outros sorotipos de AAV ainda não foram avaliados para glicosilação de capsídeo.

[004] Os vetores de terapia genética de AAV foram submetidos a menos do escrutínio de nível molecular que normalmente acompanha o desenvolvimento e fabricação de terapêuticas de proteínas recombinantes. As modificações pós-tradução do capsídeo de AAV (PTM) têm sido amplamente inexploradas, portanto, pouco se sabe sobre seu potencial para impactar a função, ou sobre estratégias para controlar os níveis de PTM em terapias de AAV fabricadas.

[005] Variações nas modificações pós-tradução de terapias com proteínas de terapia não genética complicaram seu desenvolvimento como drogas. Jenkins, N, Murphy, L e Tyther, R (2008). Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. *Mol Biotechnol* 39: 113-118; Houde, D, Peng, Y, Berkowitz, SA, and Engen, JR (2010). Post-translational modifications differentially

affect IgG1 conformation and receptor binding. *Mol Cell Proteomics* 9: 1716-1728. Por exemplo, a desamidação de aminoácidos selecionados modula a estabilidade e a resposta imune à vacina contra antraz à base de antígeno protetor recombinante. (Powell BS, et al. *Proteins*. 2007; 68(2):458-79; Verma A, et al. *Clin Vaccine Immunol*. 2016; 23(5):396-402). Em alguns casos, este processo é catalisado por desamidases virais ou bacterianas para modular as vias de sinalização da célula hospedeira ou respostas imunes inatas (Zhao J, et al. *J Virol*. 2016; 90(9):4262-8; Zhao J, et al. *Cell Host Microbe*. 2016; 20(6):770-84). Mais comumente, a desamidação endógena é um processo espontâneo independente de enzima. Embora o propósito da desamidação espontânea não tenha sido totalmente elucidado, estudos anteriores sugeriram que esse evento serve como um relógio molecular para indicar a idade relativa de uma proteína e regular sua dobra (Robinson NE e Robinson AB. *Proc Natl Acad Sci USA*. 2001; 98(3):944-9).

[006] A desamidação ocorre quando o grupo amida da asparagina ou menos frequentemente da glutamina sofre ataque nucleofílico de um átomo de nitrogênio adjacente e o grupo amida é perdido. Este processo leva a um intermediário succinimidil (Yang H e Zubarev RA. *Electrophoresis*. 2010; 31(11):1764-72) que, por meio de hidrólise, se resolve em uma mistura de ácido aspártico e ácido isoaspártico (ou ácido glutâmico e ácido isoglutâmico) (Catak S, et al. *J Phys Chem A*. 2009; 113(6):1111-20). Estudos de peptídeos sintéticos curtos estimam que essa hidrólise resulta em uma mistura 3:1 de ácido isoaspártico em ácido aspártico (Geiger T. and Clarke S. *J Biol Chem*. 1987; 262(2):785-94).

[007] Continua a haver uma necessidade de composições compreendendo construtos baseados em AAV para distribuição de moléculas heterólogas que têm ligação ao receptor estável e/ou capsídeos está-

veis, evitam anticorpos neutralizantes e/ou retêm pureza no armazenamento.

#### SUMÁRIO DA INVENÇÃO

[008] Em uma modalidade, é fornecida uma composição que inclui uma população mista de vírus adenoassociado recombinante (rAAV), cada um dos referidos rAAV compreendendo: (a) um capsídeo de AAV compreendendo cerca de 60 proteínas vp1 do capsídeo, proteínas vp2 e proteínas vp3, em que as proteínas vp1, vp2 e vp3 são: uma população heterogênea de proteínas vp1 que são produzidas a partir de uma sequência de ácido nucleico que codifica uma sequência de aminoácidos AAV vp1 selecionada, uma população heterogênea de proteínas vp2 que são produzidas a partir de uma sequência de ácido nucleico que codifica uma sequência de aminoácidos AAV vp2 selecionada, uma população heterogênea de proteínas vp3 que produziu a partir de uma sequência de ácido nucleico que codifica uma sequência de aminoácidos AAV vp3 selecionada, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares de asparagina-glicina no capsídeo de AAV e, opcionalmente, compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácido; e (b) um genoma de vetor no capsídeo de AAV, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

[009] Em certas modalidades, as asparaginas desamidadas são desamidadas em ácido aspártico, ácido isoaspártico, um par intercon-

vertido de ácido aspártico/ácido isoaspártico ou combinações dos mesmos. Em certas modalidades, o capsídeo compreende ainda glutamina(s) desamidada(s) que são desamidadas em ácido ( $\alpha$ )-glutâmico, ácido  $\gamma$ -glutâmico, um par interconvertido de ácido ( $\alpha$ )-glutâmico/ácido  $\gamma$ -glutâmico ou combinações dos mesmos.

[0010] Em um aspecto adicional, um vírus adenoassociado recombinante (rAAV) é fornecido, o qual compreende: (A) um capsídeo AAVrh79 compreendendo um ou mais de: (1) proteínas do capsídeo AAVrh79 compreendendo: uma população heterogênea de proteínas AAVrh79 vp1 selecionadas de: proteínas vp1 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 2, proteínas vp1 produzidas a partir da SEQ ID NO: 1 ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 2, uma população heterogênea de proteínas AAVrh79 vp2 selecionadas de: proteínas vp2 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 738 de SEQ ID NO: 2, proteínas vp2 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 412 a 2214 de SEQ ID NO: 1, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 412 a 2214 da SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 738 da SEQ ID NO: 2, uma população heterogênea de proteínas AAVrh79 vp3 selecionadas de: proteínas vp3 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 2, proteínas vp3 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 610 a

2214 da SEQ ID NO: 1, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 610 a 2214 da SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 2; e/ou (2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de SEQ ID NO: 2, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica o sequência de aminoácidos de pelo menos cerca de aminoácidos 138 a 738 de SEQ ID NO: 2, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 de SEQ ID NO: 2, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina em SEQ ID NO: 2 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos; e (B) um genoma de vetor no capsídeo AAVrh79, o genoma do vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

[0011] Em outro aspecto, um método de transdução de um tecido alvo é fornecido. Em uma modalidade, o método inclui administrar um AAV com um capsídeo de AAVrh79 conforme descrito neste documento. Em uma modalidade, um método de transdução de tecido hepático é fornecido, compreendendo administrar um AAV com o capsídeo AAVrh79. Em outra modalidade, um método de transdução de tecido

muscular é fornecido, compreendendo administrar um AAV com o capsídeo AAVrh79.

[0012] Em ainda outro aspecto, é fornecido um método para reduzir a desamidação de um capsídeo de AAVrh79. Em uma modalidade, o método inclui produzir um capsídeo AAVrh79 a partir de uma sequência de ácido nucleico contendo códons AAVrh79VP modificados, a sequência de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina - glicina localizados na posição N57, N263, N385 e/ou N514 em SEQ ID NO: 2, de modo que o códon modificado codifica um aminoácido diferente de glicina. Em outra modalidade, o método inclui produzir um capsídeo AAVrh79 a partir de uma sequência de ácido nucleico contendo códons AAVrh79vp modificados, a sequência de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina-glicina localizados na posição N94, N254, N305, N410, e/ou N479 de SEQ ID NO: 2.

[0013] Em uma modalidade adicional, é fornecido um rAAV8.AR2.08 que compreende: (A) um capsídeo AAV8.AR2.08 compreendendo um ou mais de: (1) proteínas do capsídeo AAV8.2.08 compreendendo: uma população heterogênea de proteínas AAV8.AR2.08vp1 selecionadas de: proteínas vp1 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18, proteínas vp1 produzidas a partir da SEQ ID NO: 17, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18, uma população heterogênea de proteínas AAV8.AR2.08vp2 selecionadas de: proteínas AAV8.AR2.08vp2 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos



aminoácidos 138 a 738 da SEQ ID NO: 18, proteínas vp2 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 411 a 2214 da SEQ ID NO: 17, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica aos pelo menos nucleotídeos 412 a 2214 da SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, uma população heterogênea de proteínas AAV8.AR2.08vp3 selecionada de: proteínas vp3 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 204 a 738 de SEQ ID NO: 18, proteínas vp3 produzidas a partir de uma sequência compreendendo pelo menos nucleotídeos 607 a 2214 de SEQ ID NO: 17, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 607 a 2214 de SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 204 a 738 de SEQ ID NO: 18; e/ou (2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de SEQ ID NO: 18, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica o sequência de aminoácidos de pelo menos cerca de aminoácidos 138 a 738 de SEQ ID NO: 18, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 de SEQ ID NO: 18 em que : as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 18 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos; e (B) um genoma de vetor no

capsídeo AAV8.AR2.08, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

[0014] Em outro aspecto, um método de transdução de um tecido alvo é fornecido. Em uma modalidade, o método inclui administrar um AAV tendo um capsídeo AAV8.AR2.08 conforme descrito neste documento. Em uma modalidade, um método de transdução de tecido hepático é fornecido, compreendendo administrar um AAV com o capsídeo AAV8.AR2.08. Em outra modalidade, um método de transdução de tecido muscular é fornecido, compreendendo administrar um AAV com o capsídeo AAV8.AR2.08.

[0015] Em ainda outro aspecto, um método para reduzir a desamidação de um capsídeo AAV8.Capsídeo AR2.08 é fornecido. Em uma modalidade, o método inclui produzir um capsídeo AAV8.AR2.08 de uma sequência de ácido nucleico contendo códons de AAV8.AR2.08 vp modificados, a sequência de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina - glicina localizados na posição N57, N263, N385, N514 e/ou N540 na SEQ ID NO: 18, de modo que o códon modificado codifica um aminoácido diferente da glicina. Em uma outra modalidade, o método inclui produzir um capsídeo AAV8.AR2.08 de uma sequência de ácido nucleico contendo códons de AAV8.AR2.08 vp modificados, a sequência de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina - glicina localizados na posição N94, N254, N305, N521, N590, Q601, N653 e/ou N665 da SEQ ID NO : 18.

[0016] Em certas modalidades, é fornecido um rAAV5.5.9 que compreende: (A) um capsídeo AAV5.5.9 compreendendo um ou mais de:

(1) proteínas do capsídeo AAVG5 compreendendo: uma população heterogênea de proteínas AAV5.5.9 vp1 selecionadas de: proteínas vp1 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 736 de SEQ ID NO: 10, proteínas vp1 produzidas a partir de SEQ ID NO: 9 ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de 1 a 736 da SEQ ID NO: 1, uma população heterogênea de proteínas AAV5.5.9 vp2 selecionadas de: proteínas AAVG5 vp2 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 736 da SEQ ID NO: 10, proteínas vp2 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 412 a 2211 da SEQ ID NO: 9, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 412 a 2211 da SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 736 da SEQ ID NO: 10, uma população heterogênea de proteínas AAV5.5.9 vp3 selecionadas de: proteínas AAV5.5.9 vp3 produzida pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 203 a 736 da SEQ ID NO: 10, proteínas vp3 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 607 a 2211 da SEQ ID NO: 9, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 607 a 2211 de SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 203 a 736 de SEQ ID NO: 10 ; e/ou (2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de

aminoácidos de SEQ ID NO: 10, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica o sequência de aminoácidos de pelo menos cerca de aminoácidos 138 a 736 de SEQ ID NO: 10, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 203 a 726 de SEQ ID NO: 10 em que : as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina - glicina em SEQ ID NO: 10 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos; e (B) um genoma de vetor no capsídeo de AAV5.5.9, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

[0017] Em outro aspecto, um método de transdução de um tecido alvo é fornecido. Em uma modalidade, o método inclui administrar um AAV com um capsídeo de AAV5.5.9 conforme descrito neste documento. Em uma modalidade, um método de transdução de tecido hepático é fornecido, compreendendo administrar um AAV com o capsídeo AAV5.5.9. Em outra modalidade, um método de transdução de tecido muscular é fornecido, compreendendo administrar um AAV com o capsídeo AAV5.5.9.

[0018] Em ainda outro aspecto, é fornecido um método para reduzir a desamidação de um capsídeo de AAV5.5.9. Em uma modalidade, o método inclui produzir um capsídeo AAV5.5.9 a partir de uma sequência de ácido nucleico contendo códons AAV5.5.9vp modificados, a sequên-

cia de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina - glicina localizados na posição N57, N319, N442 e/ou N502 em SEQ ID NO: 10, de modo que o códon modificado codifica um aminoácido diferente de glicina. Em outra modalidade, o método inclui produzir um capsídeo de AAV5.5.9 a partir de uma sequência de ácido nucleico contendo códons AAV5.5.9vp modificados, a sequência de ácido nucleico compreendendo códons de glicina independentemente modificados em um a quatro dos pares asparagina - glicina localizados na posição N35, N113, N204, N217, N243, N249, N293/294, N304, N399/400, N505, Q589, N618, N641, N653, N658 e/ou N699 de SEQ ID NO: 10.

[0019] Em outro aspecto, uma composição que compreende uma população mista de AAVrh79 recombinante, AAV8.AR2.08 ou AAV5.5.9, conforme descrito neste documento, é fornecida.

[0020] Em ainda outro aspecto, um AAV recombinante (rAAV), conforme descrito neste documento, é fornecido, para a distribuição de um produto de gene desejado a um sujeito em necessidade do mesmo.

[0021] Em outro aspecto, é fornecido um sistema de produção de rAAV útil para produzir um rAAV, conforme descrito neste documento. Em uma modalidade, o sistema inclui (a) um AAVrh79, AAV8.AR2.08, ou sequência de ácido nucleico do capsídeo AAV5.5.9 que codifica a sequência de aminoácidos predita de SEQ ID NO: 2, 10 ou 18; (b) uma molécula de ácido nucleico adequada para empacotamento no capsídeo AAV, a referida molécula de ácido nucleico compreendendo pelo menos uma repetição terminal invertida de AAV (ITR) e uma sequência de ácido nucleico não AAV que codifica um produto de gene operacionalmente ligado a sequências que dirigem a expressão do produto em uma célula hospedeira; e (c) funções rep de AAV e funções auxiliares suficientes para permitir o empacotamento da molécula de ácido nucleico no capsídeo de AAV recombinante.

[0022] Estes e outros aspectos da invenção serão evidentes a partir da seguinte descrição detalhada da invenção.

#### BREVE DESCRIÇÃO DOS DESENHOS

[0023] FIG. 1A – FIG. 1G. Análise eletroforética de isoformas AAV8 VP. (FIG. 1A) Diagrama ilustrando o mecanismo pelo qual resíduos de asparagina sofrem ataque nucleofílico por átomos de nitrogênio adjacentes, formando um intermediário succinimidil. Este intermediário então sofre hidrólise, resolvendo-se em uma mistura de ácido aspártico e ácido isoaspártico. O carbono beta é marcado como tal. O diagrama foi gerado no BIOVIA Draw 2018. (FIG. 1B) 1 µg de vetor AAV8 foi executado em um SDS-PAGE unidimensional desnaturante. (FIG. 1C) pontos isoeletricos das manchas do marcador pI da anidrase carbônica são mostrados. (FIG. 1D) 5 µg de vetor AAV8 foram analisados por eletroforese em gel bidimensional e corados com Coomassie Blue. As manchas 1-20 são marcadores pI da anidrase carbônica carbamilada. As regiões em caixas são as seguintes: a = VP1, b = VP2, c = VP3, d = marcador de tropomiosina interno (seta: mancha de tropomiosina de MW = 33kDa, pI = 5,2). A focagem isoeletrica foi realizada com um intervalo de pI de 4-8. FIG. 1E - FIG. 1G) Resultados de focagem isoeletrica realizada com um intervalo de pI de 4-8. 1e11 GC de wtAAV8 (FIG. 1E) ou vetor mutante (FIG. 1F e FIG. 1G), que foram analisados por eletroforese em gel 2D e corados com Sypro Ruby. Marcação de proteínas: A = VP1; B = VP2; C = VP3, D = marcador de conalbumina de clara de ovo de frango, E = marcador de turbonuclease. A focagem isoeletrica foi realizada com um intervalo de pI de 6-10. As manchas primárias da isoforma VP1/2/3 estão circulados e a distância de migração das manchas principais dos marcadores são indicadas por linhas verticais (turonuclease = tracejado, conalbumina = sólido).

[0024] FIG. 2A – FIG. 2E. Análise da desamidação de asparagina e

glutamina em proteínas do capsídeo AAV8. (FIG. 2A - FIG. 2B) Espectrometria de massa de ionização por eletropulverização (ESI) e massas teóricas e observadas do peptídeo 3+ (93-103) contendo Asn-94 (FIG. 2A) e Asp-94 (FIG. 2B) são mostradas. (FIG. 2C - FIG. 2D) espectrometria de massa ESI e massas teóricas e observadas do peptídeo 3+ (247-259) contendo Asn-254 (FIG. 2C) e Asp-254 (FIG. 2D) são mostradas. As mudanças de massa observadas para Asn-94 e Asn-254 foram de 0,982 Da e 0,986 Da, respectivamente, versus uma mudança de massa teórica de 0,984 Da. (FIG. 2E) A porcentagem de desamidação em resíduos específicos de asparagina e glutamina de interesse é mostrada para os peptídeos trípticos AAV8 purificados por diferentes métodos. As barras que indicam a desamidação nos resíduos de asparagina com N+1 glicinas estão hachuradas. Resíduos determinados como sendo pelo menos 2% desamidados, em pelo menos uma preparação analisada, foram incluídos. Os dados são representados como média  $\pm$  desvio-padrão.

[0025] FIG. 3A – FIG. 3E. Modelagem estrutural do monômero AAV8 VP3 e análise de sítios desamidados. (FIG. 3A) O monômero AAV8 VP3 (identificador de PDB: 3RA8) é mostrado em uma representação de bobina. A cor da fita indica o grau relativo de flexibilidade (azul = mais rígido/fator de temperatura normal, vermelho = mais flexível/fator de alta temperatura). As esferas indicam resíduos de interesse. Os diagramas expandidos são representações esféricas de resíduos de interesse e seus resíduos circundantes para demonstrar a estrutura da proteína local (Azul = nitrogênio, vermelho = oxigênio). Os resíduos sublinhados são aqueles em motivos NG. FIG. 3B - FIG. 3E: Modelos isoaspárticos de asparaginas desamidadas com N + 1 glicinas são mostrados. O mapa de densidade de elétrons 2FoFc (nível de 1 sigma) gerado a partir do refinamento da estrutura cristalina AAV8 (PDB ID: 3RA8) com

(FIG. 3B) um modelo de asparagina de N410 em comparação com modelos de ácido isoaspártico de (FIG. 3C) N263, (FIG. 3D) N514 e (FIG. 3E) N540. O mapa de densidade de elétrons é mostrado na grade magenta. O carbono beta é marcado como tal. A seta indica a densidade do elétron correspondente ao grupo R do resíduo de interesse.

[0026] FIG. 4A – FIG. 4D. Determinação dos fatores que influenciam a desamidação do capsídeo de AAV8. Uma preparação AAV8 foi (FIG. 4A) incubada a 70°C por três ou sete dias, (FIG. 4B) exposta a pH 2 ou pH 10 por sete dias, ou (FIG. 4C) preparada para espectrometria de massa usando D<sub>2</sub>O no lugar de H<sub>2</sub>O para determinar possíveis fontes de desamidação não intrínsecas à formação de capsídeo de AAV. (FIG. 4D) Um dot blot de vetor tratado como na FIG. 4A usando o anticorpo B1 (reage ao capsídeo desnaturado) e um anticorpo específico da conformação AAV8 (reage aos capsídeos intactos) para avaliar a integridade estrutural do capsídeo.

[0027] FIG. 5A – FIG. 5B. Frequências de desamidação em proteínas não AAV. As porcentagens de desamidação são mostradas para duas proteínas recombinantes não AAV contendo motivos NG suscetíveis de serem desamidados, anidrase carbônica humana (FIG. 5A) e fenilalanina-hidroxilase de rato (FIG. 5B), para comparação com as porcentagens de desamidação de AAV.

[0028] FIG. 6. Comparação da desamidação AAV8 por cento calculada usando tubos de análise de dados de duas instituições. A porcentagem de desamidação em resíduos específicos de asparagina e glutamina de interesse é mostrada para os peptídeos trípticos AAV8 avaliados em duas instituições diferentes.

[0029] FIG. 7A - FIG. 7C ilustra substituições funcionais de asparagina em sítios não NG com alta variabilidade entre os lotes. (FIG. 7A) Os títulos de wtAAV8 e vetores mutantes foram produzidos por transfecção tripla em pequena escala em células 293, conforme medido por



PCR quantitativa (qPCR). Os títulos são relatados em relação ao controle wtAAV8. As eficiências de transdução foram medidas conforme descrito na FIG. 8B. Os títulos e as eficiências de transdução são normalizados para o valor do controle wtAAV8. (FIG. 7B) Imagens representativas da luciferase no dia 14 pós-injeção são mostradas para camundongos que receberam vetor mutante de capsídeo wtAAV8.CB7.ffluc e N499Q. (FIG. 7C) A expressão de luciferase no dia 14 dos períodos de estudo de camundongos C57BL/6 injetados por via intravenosa com wtAAV8 ou vetores mutantes (n = 3 ou 4) foi medida por imageamento de luciferase e relatada em unidades de fluxo total. Todos os dados são representados como média + desvio-padrão.

[0030] FIG. 8A e FIG. 8B mostram os resultados da análise *in vitro* do impacto da desamidação genética no desempenho do vetor. (FIG. 8A) Os títulos de wtAAV8 e vetores mutantes de desamidação genética produzidos por transfecção tripla em pequena escala em células 293, conforme medido por PCR quantitativa (qPCR). Os títulos são relatados em relação ao controle wtAAV8. Sítios NG com alta desamidação (barras padronizadas), sítios com baixa desamidação (barras brancas) e sítios altamente variáveis (barras pretas) são apresentados com wtAAV8 e um controle negativo. (FIG. 8B) Eficiência de transdução de vetores mutantes AAV8 produzindo luciferase de pirilampo relatada em relação ao controle wtAAV8. A eficiência de transdução é medida em unidades de luminescência geradas por GC adicionado às células HUH7 e é determinada realizando transduções com vetor bruto em diluições múltiplas. Os dados de eficiência de transdução são normalizados para a referência em peso. Todos os dados são representados como média  $\pm$  desvio-padrão.

[0031] FIG. 9A - FIG. 9D ilustram que a perda de atividade do vetor ao longo do tempo está correlacionada à desamidação progressiva. (FIG. 9A) Produção de vetor (cópias do genoma resistentes a DNaseI,

GC) para um curso de tempo de células HEK 293 triplamente transfectadas produzindo vetor AAV8 que empacota um gene repórter de luciferase. Os níveis de GC são normalizados para o valor máximo observado. (FIG. 9B) O vetor de curso de tempo purificado foi usado para transduzir células Huh7. A eficiência de transdução (unidades de luminescência por GC adicionadas às células alvo) foi medida como na FIG. 8B usando múltiplas diluições de amostras de vetor de curso de tempo purificadas. Barras de erro representam o desvio-padrão de pelo menos 10 réplicas técnicas para cada tempo de amostra. Desamidação de sítios AAV8 NG (FIG. 9C) e sítios não NG (FIG. 9D) para o vetor coletado 1, 2 e 5 dias após a transfecção.

[0032] FIG. 10A - FIG. 10D ilustra o impacto da estabilização das asparaginas no desempenho do vetor. A FIG. 10A mostra os títulos de wtAAV8 e vetores mutantes na posição +1 produzidos por transfecção tripla em pequena escala em células 293, conforme medido por PCR quantitativa (qPCR). Os títulos são relatados em relação ao controle wtAAV8. A FIG 10B mostra a eficiência de transdução de vetores mutantes AAV8 produzindo luciferase de pirilampo relatada em relação ao controle wtAAV8. A eficiência de transdução foi medida como na FIG. 8B usando material vetorial bruto. Um teste t de duas amostras (\*p<0,005) foi executado para determinar a significância entre wtAAV8 e a eficiência de transdução mutante para G264A/G515A e G264A/G541A. A FIG. 10C mostra a expressão de luciferase no dia 14 do período de estudo na região do fígado de camundongos C57BL/6 injetados por via intravenosa com wtAAV8 ou vetores mutantes (n = 3 a 5) medidos por imageamento de luciferase e relatados em unidades de fluxo total. A FIG. 10D mostra os títulos e a eficiência de transdução de vetores mutantes AAV8 de múltiplos sítios que produzem luciferase de pirilampo relatados em relação ao controle wtAAV8. Todos os dados são representados como média ± desvio-padrão.

[0033] FIG. 11A – FIG. 11C. Análise da desamidação de asparagina e glutamina em proteínas do capsídeo AAV9. (FIG. 11A) 1e11 GCs de wtAAV9 foram analisados por eletroforese em gel 2D e corados com Sypro Ruby. Marcação de proteínas: A = VP1; B = VP2; C = VP3, D = marcador de conalbumina de clara de ovo de frango, E = marcador de turbonuclease. A focagem isoelétrica foi realizada com um intervalo de pI de 6-10. (FIG. 11B) A porcentagem de desamidação em resíduos específicos de asparagina e glutamina de interesse é mostrada para os peptídeos trípticos AAV9 purificados por diferentes métodos. As barras que indicam a desamidação nos resíduos de asparagina com N+1 glicinas estão hachuradas. Resíduos determinados como sendo pelo menos 2% desamidados, em pelo menos uma preparação analisada, foram incluídos. Os dados são representados como média  $\pm$  desvio-padrão. (FIG. 11C) O modelo isoaspártico de N512 é mostrado no mapa de densidade de elétrons 2FoFc gerado por refinamento não polarizado da estrutura cristalina AAV9 (PDB ID: 3UX1). A seta indica a densidade do elétron correspondente ao grupo R do resíduo N512.

[0034] FIG. 11D – FIG. 11F. Determinação dos fatores que influenciam a desamidação do capsídeo de AAV9. (FIG. 11D) Duas preparações de AAV9 foram incubadas a 70°C três ou sete dias ou (FIG. 11F) expostas a pH 2 ou pH 10 por sete dias para determinar possíveis fontes de desamidação não intrínsecas à formação de capsídeo de AAV. Os dados são representados como média  $\pm$  desvio-padrão. (FIG. 11F) Um dot blot de vetor tratado como na FIG. 11D usando o anticorpo B1 (reage ao capsídeo desnaturado) para avaliar a integridade estrutural do capsídeo.

[0035] FIG. 11G e FIG. 11H ilustram a análise *in vitro* do impacto da desamidação genética no desempenho do vetor para AAV9. (FIG. 11G) Os títulos de wtAAV9 e vetores mutantes de desamidação genética foram produzidos por transfecção tripla em pequena escala em células

293, conforme medido por PCR quantitativa (qPCR). Os títulos são relatados em relação ao controle wtAAV9. Sítios NG com alta desamidação (barras padronizadas), sítios com baixa desamidação (barras brancas) e sítios altamente variáveis (barras pretas) são apresentados com wtAAV8 e um controle negativo. (FIG. 11H) A eficiência de transdução de vetores mutantes AAV9 que produzem luciferase de pirilampo é relatada em relação ao controle wtAAV9. Todos os dados são representados como média  $\pm$  desvio-padrão.

[0036] FIG. 11I - FIG. 11K mostram a potência *in vitro* do vetor AAV9 ao longo do tempo. (FIG. 11I) Produção de vetor (cópias do genoma resistentes a DNaseI, GC) para um curso de tempo de células HEK 293 triplamente transfectadas produzindo vetor AAV9 que empacota um gene repórter de luciferase. Os níveis de GC são normalizados para o valor máximo observado. (FIG. 11J) O vetor de curso de tempo bruto foi usado para transduzir células Huh7. (FIG. 11K) As eficiências de transdução do vetor coletado 1 dia após a transfecção vs 5 dias após a transfecção são mostradas para amostras de vetor em bruto e purificadas. A eficiência de transdução é expressa como atividade de luciferase/GC, normalizada para o valor no dia 1.

[0037] A FIG 12A fornece um alinhamento das sequências de aminoácidos de AAV5.5.9 [SEQ ID NO: 10] (também às vezes chamado de AAVG5), AAV9 [SEQ ID NO: 4] e AAVPHP.B [SEQ ID NO: 12], preparado usando Clustal Omega 1.2.2 e seus parâmetros padrão para alinhamento. As FIGS 12B - 12E fornecem um alinhamento das sequências de nucleotídeos de AAV5.5.9 [SEQ ID NO: 9], PHP.B [SEQ ID NO: 11], AAV9 [SEQ ID NO: 3] e AAVhu68 [SEQ ID NO: 14].

[0038] A FIG 13A fornece um alinhamento das sequências de aminoácidos do mutante AAV8Triple (AAV8T) [SEQ ID NO: 16], AAV8.AR2.08 [SEQ ID NO: 18] (às vezes também chamado de AAVG3

ou AR2 ou AAV.AR2) e AAV8 [SEQID NO: 20], preparado usando Clustal Omega 1.2.2 e seus parâmetros padrão para alinhamento. As FIGS 13B - 13D fornecem um alinhamento das sequências de nucleotídeos do Mutante Triplo AAV8 [SEQ ID NO: 15], AAV8.AR2.08 [SEQ ID NO: 17], AAV8 [SEQ ID NO: 19].

[0039] A FIG 14A fornece um alinhamento das sequências de aminoácidos de AAVrh79 [SEQ ID NO: 2] (também às vezes chamado de AAVG2), AAVrh10 [SEQ ID NO: 24] e AAVhu37 [SEQ ID NO: 22], usando Clustal Omega 1.2.2 e seus parâmetros padrão para alinhamento. As FIGS 14B - 14D fornecem um alinhamento das sequências de nucleotídeos de AAVrh79 [SEQ ID NO: 1], AAVrh10 [SEQ ID NO: 22] e AAVhu37 [SEQ ID NO: 21].

[0040] FIGS 15A e 15B ilustram o rendimento da produção para AAV8triple, AAVhu68, AAV9, AAV9 e AAVrh79 em pequena escala ou preparação em mega escala do vetor referenciado.

[0041] A FIG 16 fornece a pureza de produção das preparações em mega escala da FIG 15B.

[0042] FIGs 17A a17D mostrar a expressão da luciferase no fígado e tecido muscular após a administração intramuscular (IM) de  $3 \times 10^{11}$  GC/camundongo no músculo gastrocnêmio de camundongos machos C57BL/6 (n = 5/grupo) usando vetores que expressam luciferase de pi-rilampo.

[0043] A FIG. 18A mostra o plasmídeo usado para as experiências de código de barras no Exemplo 5. A FIG. 18B mostra as quantidades de cada variante de código de barras AAV injetada em 6 camundongos pretos. Os animais foram sacrificados, amostras de tecido colhidas e DNA isolado de cada um deles. A distribuição total do vetor para os três animais é mostrada na FIG. 18C. A frequência real vs. teórica da mistura de vetores injetados é mostrada na FIG. 18D.

[0044] FIGs. 19A-FIG. 20C mostram os resultados dos experimentos de biodistribuição de código de barras do Exemplo 5. Amostras de tecido individuais foram analisadas para frequência de código de barras individual na amostra vs. mistura de injeção para genômico e cDNA. Os resultados são mostrados para o músculo (FIG. 19A, 19B); coração (FIG. 19C e FIG. 19D) e fígado (FIG. 19E e FIG. 19F). As alterações de dobra em comparação com as frequências teóricas são mostradas na FIG. 20A - FIG. 20C.

[0045] As FIGs. 21 e 22 mostram a biodistribuição AAV8.AR2.08 em camundongos em comparação com AAV8. Os resultados mostram que AAV8.AR2.08 é mais específico do fígado do que AAV8.

[0046] A FIG. 23 compara AAV8 vs. AAV8.AR2.08 vs. AAVrh79 para título e rendimento em relação à capacidade de fabricação.

[0047] A FIG. 24 mostra a biodistribuição AAV8.AR2.08 nos tecidos (barra mais à esquerda) em comparação com AAV8 (barras do meio e direita).

[0048] As FIGs 25-28 mostram os resultados após a administração de vetores AAV a primatas não humanos. A FIG. 25 fornece detalhes para vetores e animais usados para estudos. A FIG. 26 quantifica os níveis de GC e GFP detectados no fígado de animais que receberam vetores AAV8, AAVrh79 ou AAV8.AR2.08. A FIG. 27 resume os níveis de expressão de GFP em fígados HNP. A FIG. 28 mostra o nível de vetor detectado em tecido de HNP aos quais foi administrado vetores AAV8, AAVrh79 ou AAV8.AR2.08.

[0049] A FIG. 29 mostra a biodistribuição do vetor AAVrh79 detectado em vários tecidos.

#### DESCRIÇÃO DETALHADA DA INVENÇÃO

[0050] São fornecidos aqui vírus adenoassociados recombinantes (rAAV) com sequência e heterogeneidade de carga em cada uma das três populações de proteínas do capsídeo VP1, VP2 e VP3 encontradas

no capsídeo de um AAV recombinante e composições contendo os mesmos. São fornecidos neste documento novos rAAV, bem como métodos para reduzir a desamidação e, opcionalmente, outras modificações de monômero de capsídeo. Além disso, são fornecidos neste documento o rAAV modificado com modificações diminuídas, que são úteis para fornecer rAAV com capsídeos que retêm maior estabilidade, potência e/ou pureza.

[0051] Um "AAV recombinante" ou "rAAV" é uma partícula viral resistente a DNase contendo dois elementos, um capsídeo de AAV e um genoma de vetor contendo pelo menos sequências de codificação não AAV empacotadas dentro do capsídeo de AAV. A menos que especificado de outra forma, este termo pode ser usado alternadamente com a frase "vetor rAAV". O rAAV é um "vírus de replicação defeituosa" ou "vetor viral", uma vez que carece de qualquer gene rep de AAV funcional ou gene cap de AAV funcional e não pode gerar progênie. Em certas modalidades, as únicas sequências de AAV são as sequências de repetição terminal invertida (ITRs) de AAV, normalmente localizadas nas extremidades 5' e 3' extremas do genoma do vetor, a fim de permitir que o gene e as sequências regulatórias localizadas entre as ITRs sejam empacotadas dentro do capsídeo AAV.

[0052] Como usado neste documento, um "genoma de vetor" se refere à sequência de ácido nucleico empacotada dentro do capsídeo de rAAV que forma uma partícula viral. Essa sequência de ácido nucleico contém sequências de repetição terminal invertida (ITRs) de AAV. Nos exemplos aqui descritos, um genoma de vetor contém, no mínimo, de 5'a 3', um AAV 5' ITR, sequência(s) de codificação e um AAV 3' ITR. ITRs de AAV2, um AAV de fonte diferente do capsídeo, ou diferente de ITRs de comprimento total podem ser selecionados. Em certas modalidades, as ITRs são da mesma fonte de AAV que o AAV, que fornece a função rep durante a produção ou um AAV de transcomplementação.

Além disso, outras ITRs podem ser usadas. Além disso, o genoma do vetor contém sequências regulatórias que dirigem a expressão dos produtos de gene. Componentes adequados de um genoma de vetor são discutidos em mais detalhes neste documento.

[0053] Um rAAV é composto de um capsídeo de AAV e um genoma de vetor. Um capsídeo AAV é um conjunto de uma população heterogênea de vp1, uma população heterogênea de vp2 e uma população heterogênea de proteínas vp3. Como usado neste documento, quando usado para se referir a proteínas do capsídeo vp, o termo "heterogêneo" ou qualquer variação gramatical do mesmo, refere-se a uma população que consiste em elementos que não são iguais, por exemplo, tendo monômeros vp1, vp2 ou vp3 (proteínas) com diferentes sequências de aminoácidos modificados.

[0054] Como usado neste documento, o termo "população heterogênea", conforme usado em conexão com as proteínas vp1, vp2 e vp3 (alternativamente chamadas de isoformas), refere-se a diferenças na sequência de aminoácidos das proteínas vp1, vp2 e vp3 dentro de um capsídeo. O capsídeo de AAV contém subpopulações dentro das proteínas vp1, dentro das proteínas vp2 e dentro das proteínas vp3 que possuem modificações dos resíduos de aminoácidos preditos. Essas subpopulações incluem, no mínimo, certos resíduos de asparagina desamidada (N ou Asn). Por exemplo, certas subpopulações compreendem pelo menos uma, duas, três ou quatro posições de asparaginas (N) altamente desamidadas em pares asparagina-glicina e, opcionalmente, compreendendo ainda outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácido e outras modificações opcionais.

[0055] Como usado neste documento, uma "subpopulação" de proteínas vp refere-se a um grupo de proteínas vp que tem pelo menos uma característica definida em comum e que consiste em pelo menos um



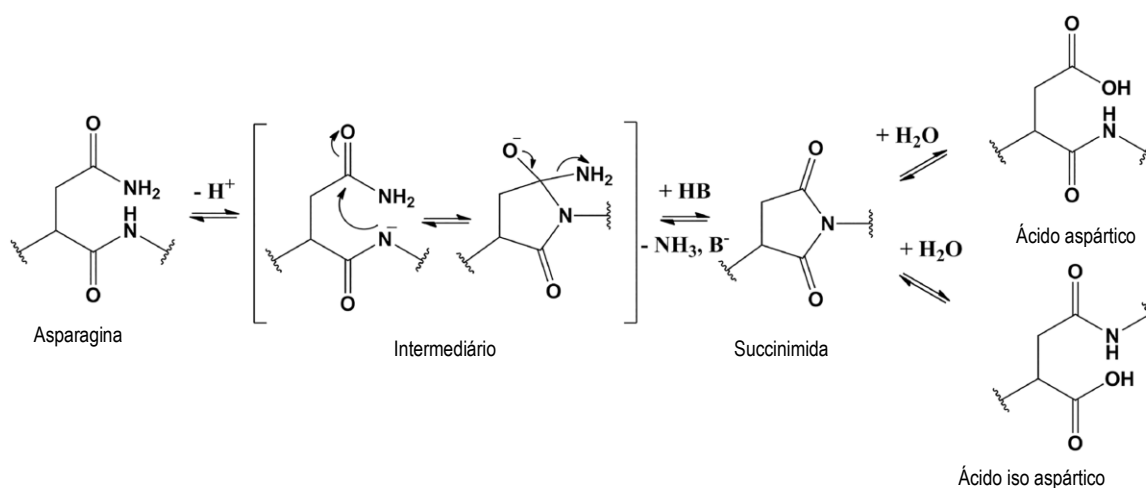
membro do grupo a menos do que todos os membros do grupo de referência, a menos que especificado de outra forma. Por exemplo, uma "subpopulação" de proteínas vp1 pode ser pelo menos uma (1) proteína vp1 e menos do que todas as proteínas vp1 em um capsídeo AAV montado, a menos que especificado de outra forma. Uma "subpopulação" de proteínas vp3 pode ser uma (1) proteína vp3 para menos do que todas as proteínas vp3 em um capsídeo de AAV montado, a menos que especificado de outra forma. Por exemplo, as proteínas vp1 podem ser uma subpopulação de proteínas vp; as proteínas vp2 podem ser uma subpopulação separada de proteínas vp, e vp3 é ainda uma subpopulação adicional de proteínas vp em um capsídeo de AAV montado. Em outro exemplo, as proteínas vp1, vp2 e vp3 podem conter subpopulações com diferentes modificações, por exemplo, pelo menos uma, duas, três ou quatro asparaginas altamente desamidadas, por exemplo, em pares asparagina - glicina.

[0056] A menos que especificado de outra forma, altamente desamidado refere-se a pelo menos 45% desamidado, pelo menos 50% desamidado, pelo menos 60% desamidado, pelo menos 65% desamidado, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99% ou até cerca de 100% desamidado em uma posição de aminoácido referenciada, em comparação com a sequência de aminoácidos predita na posição de aminoácido de referência (por exemplo, pelo menos 80% das asparaginas no aminoácido 57 da SEQ ID NO: 2 podem ser desamidadas com base nas proteínas vp1 totais podem ser desamidadas com base nas proteínas vp1, vp2 e vp3 totais). Essas porcentagens podem ser determinadas usando gel 2D, técnicas de espectrometria de massa ou outras técnicas adequadas.

[0057] Sem desejar estar limitado pela teoria, acredita-se que a de-

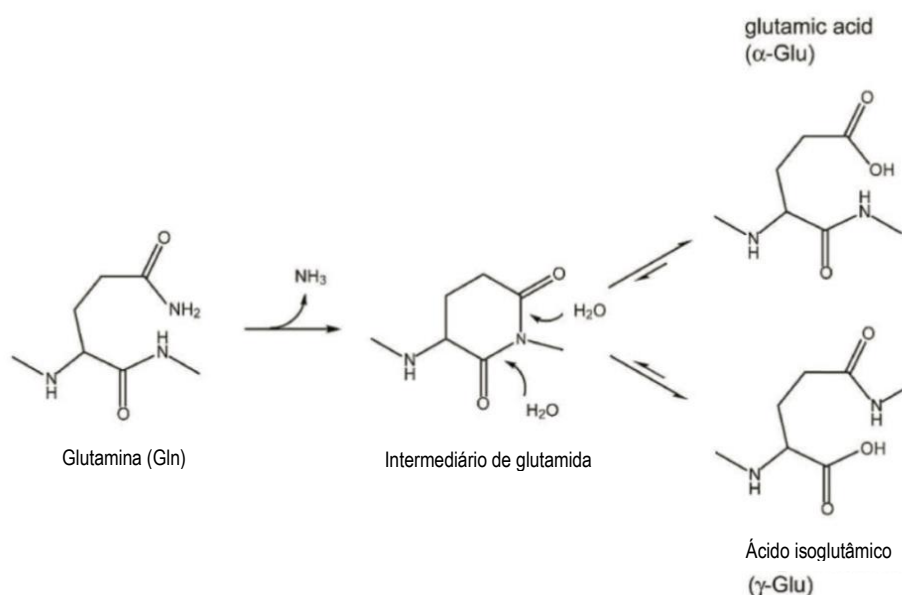
desamidação de pelo menos resíduos altamente desamidados nas proteínas vp no capsídeo de AAV seja principalmente não enzimática, sendo causada por grupos funcionais dentro da proteína do capsídeo que desamidam asparaginas selecionadas, e em menor grau, resíduos de glutamina. A montagem eficiente do capsídeo da maioria das proteínas vp1 de desamidação indica que esses eventos ocorrem após a montagem do capsídeo ou que a desamidação em monômeros individuais (vp1, vp2 ou vp3) é bem tolerada estruturalmente e não afeta amplamente a dinâmica da montagem. A desamidação extensiva na região única de VP1 (VP1-u) (~aa 1-137), geralmente considerada como localizada internamente antes da entrada celular, sugere que a desamidação de VP pode ocorrer antes da montagem do capsídeo.

[0058] Sem desejar estar limitado pela teoria, a desamidação de N pode ocorrer através do átomo de nitrogênio da espinha dorsal de seu resíduo C-terminal conduz um ataque nucleofílico ao átomo de carbono do grupo amida da cadeia lateral de Asn. Pensa-se que se forma um resíduo de succinimida de anel fechado intermediário. O resíduo de succinimida então conduz a hidrólise rápida para levar ao produto final ácido aspártico (Asp) ou ácido iso-aspártico (IsoAsp). Portanto, em certas modalidades, a desamidação da asparagina (N ou Asn) leva a um Asp ou IsoAsp, que pode se interconverter através do intermediário succinimida, por exemplo, conforme ilustrado abaixo.



[0059] Conforme fornecido neste documento, cada N desamidado em VP1, VP2 ou VP3 pode ser independentemente ácido aspártico (Asp), ácido isoaspártico (isoAsp), aspartato e/ou uma mistura de interconversão de Asp e isoAsp, ou combinações dos mesmos. Qualquer razão adequada de ácido  $\alpha$ - e isoaspártico pode estar presente. Por exemplo, em certas modalidades, a razão pode ser de 10:1 a 1:10 aspártico para isoaspártico, cerca de 50:50 aspártico:isoaspártico ou cerca de 1:3 aspártico:isoaspártico ou outra razão selecionada.

[0060] Em certas modalidades, uma ou mais glutamina (Q) podem desamidar em ácido glutâmico (Glu), ou seja, ácido  $\alpha$ -glutâmico, ácido  $\gamma$ -glutâmico (Glu) ou uma mistura de ácido  $\alpha$  e  $\gamma$ -glutâmico, que pode se interconverter através de um intermediário comum de glutarimidina. Qualquer razão adequada de ácido  $\alpha$ - e  $\gamma$ -glutâmico pode estar presente. Por exemplo, em certas modalidades, a razão pode ser de 10:1 a 1:10  $\alpha$  a  $\gamma$ , cerca de 50:50  $\alpha$ : $\gamma$ , ou cerca de 1:3  $\alpha$ : $\gamma$ , ou outra razão selecionada.



[0061] Assim, um rAAV inclui subpopulações dentro do capsídeo de rAAV das proteínas vp1, vp2 e/ou vp3 com aminoácidos desamidados, incluindo, no mínimo, pelo menos uma subpopulação compreendendo

pelo menos uma asparagina altamente desamidada. Além disso, outras modificações podem incluir isomerização, particularmente em posições de resíduo de ácido aspártico (D ou Asp) selecionadas. Em ainda outras modalidades, as modificações podem incluir uma amidação em uma posição Asp.

[0062] Em certas modalidades, um capsídeo de AAV contém subpopulações de vp1, vp2 e vp3 tendo pelo menos 1, pelo menos 2, pelo menos 3, pelo menos 4, pelo menos 5 a pelo menos cerca de 25 posições de resíduos de aminoácidos desamidados, dos quais pelo menos 1 a 10%, pelo menos 10 a 25%, pelo menos 25 a 50%, pelo menos 50 a 70%, pelo menos 70 a 100%, pelo menos 75 a 100%, pelo menos 80-100% ou pelo menos 90- 100% são desamidados em comparação com a sequência de aminoácidos codificada das proteínas vp. A maioria destes pode ser resíduos de N. No entanto, os resíduos Q também podem ser desamidados.

[0063] Como usado neste documento, "sequência de aminoácidos codificada" refere-se ao aminoácido que é predito com base na tradução de um códon de DNA conhecido de uma sequência de ácido nucleico referenciada sendo traduzida para um aminoácido. A tabela a seguir ilustra os códons de DNA e vinte aminoácidos comuns, mostrando o código de uma letra (SLC) e o código de três letras (3LC).

Aminoácido	SLC	3 LC	Códons de DNA
Isoleucina	I	Ile	ATT, ATC, ATA
Leucina	L	Leu	CTT, CTC, CTA, CTG, TTA, TTG
Valina	V	Val	GTT, GTC, GTA, GTG
Fenilalanina	F	Phe	TTT, TTC
Metionina	M	Met	ATG
Cisteína	C	Cys	TGT, TGC
Alanina	A	Ala	GCT, GCC, GCA, GCG
Glicina	G	Gly	GGT, GGC, GGA, GGG
Prolina	P	Pro	CCT, CCC, CCA, CCG
Treonina	T	Thr	ACT, ACC, ACA, ACG
Serina	S	Ser	TCT, TCC, TCA, TCG, AGT, AGC
Tirosina	Y	Tyr	TAT, TAC

Triptofano	W	Trp	TGG
Glutamina	Q	Gln	CAA, CAG
Asparagina	N	Asn	AAT, AAC
Histidina	H	His	CAT, CAC
Ácido glutâmico	E	Glu	GAA, GAG
Ácido aspártico	D	Asp	GAT, GAC
Lisina	K	Lys	AAA, AAG
Arginina	R	Arg	CGT, CGC, CGA, CGG, AGA, AGG
Códons de parada	Parada		TAA, TAG, TGA

[0064] Em certas modalidades, um rAAV tem um capsídeo de AAV com proteínas vp1, vp2 e vp3 tendo subpopulações que compreendem combinações de dois, três, quatro, cinco ou mais resíduos desamidados nas posições estabelecidas nas tabelas fornecidas aqui e aqui incorporadas por referência.

[0065] A desamidação no rAAV pode ser determinada usando eletroforese em gel 2D e/ou espectrometria de massa e/ou técnicas de modelagem de proteína. A cromatografia online pode ser realizada com uma coluna Acclaim PepMap e um sistema Thermo UltiMate 3000 RSLC (Thermo Fisher Scientific) acoplado a um Q Exactive HF com uma fonte NanoFlex (Thermo Fisher Scientific). Os dados de MS são adquiridos usando um método top-20 dependente de dados para o Q Exactive HF, escolhendo dinamicamente os íons precursores ainda não sequenciados mais abundantes das varreduras de pesquisa (200-2000 m/z). O sequenciamento é realizado por meio de fragmentação de dissociação colisional de alta energia com um valor alvo de íons  $1e5$  determinado com controle de ganho automático preditivo e um isolamento de precursores foi realizado com uma janela de 4 m/z. Varreduras de pesquisa foram adquiridas com uma resolução de 120.000 em m/z 200. A resolução para os espectros de HCD pode ser definida para 30.000 em m/z 200 com um tempo máximo de injeção de íons de 50 ms e uma energia de colisão normalizada de 30. O nível de RF da lente S pode ser definido em 50, para dar a transmissão ideal da região m/z ocupada pelos peptídeos do digest. Os íons precursores podem ser excluídos com estados

de carga simples, não atribuídos ou seis e mais altos da seleção de fragmentação. O software BioPharma Finder 1.0 (Thermo Fischer Scientific) pode ser usado para análise dos dados adquiridos. Para o mapeamento de peptídeos, as pesquisas são realizadas usando um banco de dados FASTA de proteína de entrada única com carbamidometilação definida como uma modificação fixa; e oxidação, desamidação e fosforilação definidas como modificações variáveis, uma precisão de massa de 10 ppm, uma alta especificidade de protease e um nível de confiança de 0,8 para espectros de MS/MS. Exemplos de proteases adequadas podem incluir, por exemplo, tripsina ou quimotripsina. A identificação por espectrometria de massa de peptídeos desamidados é relativamente simples, pois a desamidação adiciona à massa da molécula intacta +0,984 Da (a diferença de massa entre os grupos -OH e -NH<sub>2</sub>). A percentagem de desamidação de um peptídeo particular é determinada pela área de massa do peptídeo desamidado dividida pela soma da área dos peptídeos desamidados e nativos. Considerando o número de sítios de desamidação possíveis, as espécies isobáricas que são desamidadas em diferentes sítios podem comigrar em um único pico. Consequentemente, fragmentos de íons originários de peptídeos com múltiplos sítios de desamidação potenciais podem ser usados para localizar ou diferenciar múltiplos sítios de desamidação. Nestes casos, as intensidades relativas dentro dos padrões de isótopos observados podem ser usadas para determinar especificamente a abundância relativa dos diferentes isômeros de peptídeo desamidado. Este método presume que a eficiência de fragmentação para todas as espécies isoméricas é a mesma e independente do sítio de desamidação. Será entendido por aqueles versados na técnica que uma série de variações desses métodos ilustrativos pode ser usada. Por exemplo, espectrômetros de massa adequados podem incluir, *por exemplo*, um espectrômetro de massa de tempo de voo quadrupolo (QTOF), como um Waters Xevo ou Agilent

6530 ou um instrumento orbitrap, como o Orbitrap Fusion ou Orbitrap Velos (Thermo Fisher). Os sistemas de cromatografia líquida adequados incluem, *por exemplo*, o sistema Acquity UPLC da Waters ou os sistemas Agilent (séries 1100 ou 1200). O software de análise de dados adequado pode incluir, *por exemplo*, MassLynx (Waters), Pinpoint and Peptide (Thermo Fischer Scientific), Mascot (Matrix Science), Peaks DB (Bioinformatics Solutions). Ainda outras técnicas podem ser descritas, por exemplo, em X. Jin et al, Hu Gene Therapy Methods, Vol. 28, No. 5, pp. 255-267, publicado online em 16 de junho de 2017.

[0066] Além de desamidações, outras modificações podem ocorrer, não resultando na conversão de um aminoácido em um resíduo de aminoácido diferente. Essas modificações podem incluir resíduos acetilados, isomerizações, fosforilações ou oxidações.

[0067] Modulação da Desamidação: Em certas modalidades, o AAV é modificado para alterar a glicina em um par asparagina-glicina, para reduzir a desamidação. Em outras modalidades, a asparagina é alterada para um aminoácido diferente, por exemplo, uma glutamina que desamida a uma taxa mais lenta; ou a um aminoácido que carece de grupos amida (por exemplo, glutamina e asparagina contêm grupos amida); e/ou a um aminoácido que carece de grupos amina (por exemplo, lisina, arginina e histidina contêm grupos amina). Como usado neste documento, aminoácidos sem amida ou grupos laterais de amina referem-se a, por exemplo, glicina, alanina, valina, leucina, isoleucina, serina, treonina, cistina, fenilalanina, tirosina ou triptofano e/ou prolina. As modificações, tais como descritas, podem ser em um, dois ou três dos pares asparagina-glicina encontrados na sequência de aminoácidos AAV codificada. Em certas modalidades, tais modificações não são feitas em todos os quatro pares asparagina - glicina. Assim, um método para reduzir a desamidação de AAV e/ou variantes de AAV engenheiradas com

taxas de desamidação mais baixas. Adicionalmente, ou alternativamente, um ou mais outros aminoácidos amida podem ser alterados para um aminoácido não amida para reduzir a desamidação do AAV. Em certas modalidades, um capsídeo de AAV mutante, conforme descrito neste documento, contém uma mutação em um par asparagina-glicina, de modo que a glicina seja alterada para uma alanina ou uma serina. Um capsídeo de AAV mutante pode conter um, dois ou três mutantes em que o AAV de referência contém nativamente quatro pares NG. Em certas modalidades, um capsídeo de AAV pode conter um, dois, três ou quatro mutantes, onde o AAV de referência contém nativamente cinco pares de NG. Em certas modalidades, um capsídeo de AAV mutante contém apenas uma única mutação em um par NG. Em certas modalidades, um capsídeo de AAV mutante contém mutações em dois pares NG diferentes. Em certas modalidades, um capsídeo de AAV mutante contém mutação em dois pares NG diferentes que estão localizados em um local estruturalmente separado no capsídeo de AAV. Em certas modalidades, a mutação não está na região única de VP1. Em certas modalidades, uma das mutações está na região única de VP1. Opcionalmente, um capsídeo de AAV mutante não contém modificações nos pares NG, mas contém mutações para minimizar ou eliminar a desamidação em uma ou mais asparaginas, ou uma glutamina, localizada fora de um par NG.

[0068] Em certas modalidades, é fornecido um método para aumentar a potência de um vetor rAAV que compreende a engenharia de um capsídeo de AAV que elimina um ou mais dos NGs no capsídeo de AAV de tipo selvagem. Em certas modalidades, a sequência de codificação para o "G" do "NG" é engenheirada para codificar outro aminoácido. Em certos exemplos abaixo, um "S" ou um "A" é substituído. No entanto, podem ser selecionadas outras sequências de codificação de aminoácidos adequadas. Veja, por exemplo, as tabelas abaixo nas quais com



base na numeração de AAV8, a sequência de codificação para pelo menos uma das seguintes posições: N57+1, N263+1, N385+1, N514+1, N540+1, é modificada, ou como mostrado nas tabelas abaixo. Em certas modalidades, os mutantes AAV8 evitam alterar os pares NG nas posições N57, N94, N263, N305, Q467, N479 e/ou N653. Em certas modalidades, outros AAVs evitam a mutação nas posições N correspondentes, conforme determinado com base em um alinhamento com AAV8, usando a numeração de AAV8 como referência.

[0069] Estas modificações de aminoácidos podem ser feitas por técnicas convencionais de engenharia genética. Por exemplo, uma sequência de ácido nucleico contendo códons vp de AAV modificados pode ser gerada em que um a três dos códons que codificam glicina em pares arginina - glicina são modificados para codificar um aminoácido diferente de glicina. Em certas modalidades, uma sequência de ácido nucleico contendo códons de arginina modificados pode ser engenheirada em um a três dos pares arginina-glicina, de modo que o códon modificado codifique um aminoácido diferente de arginina. Cada códon modificado pode codificar um aminoácido diferente. Alternativamente, um ou mais dos códons alterados podem codificar o mesmo aminoácido. Em certas modalidades, estas sequências de ácido nucleico AAVrh79, AAV8.AR2.08 ou AAV5.5.9 modificadas podem ser usadas para gerar um rAAV mutante possuindo um capsídeo com desamidação inferior do que o capsídeo AAVrh79, AAV8.AR2.08 ou AAV5.5.9 nativo. Tal rAAV mutante pode ter imunogenicidade reduzida e/ou aumentar a estabilidade no armazenamento, particularmente armazenamento na forma de suspensão.

[0070] Também são fornecidas neste documento as sequências de ácido nucleico que codificam os capsídeos de AAV com desamidação reduzida. É do conhecimento da técnica conceber sequências de ácidos

nucleicos que codificam esse capsídeo do AAV, incluindo DNA (genômico ou cDNA) ou RNA (*por exemplo*, mRNA). Essas sequências de ácido nucleico podem ser otimizadas por códons para expressão em um sistema selecionado (isto é, tipo de célula) pode ser projetado por vários métodos. Essa otimização pode ser realizada usando métodos que estão disponíveis on-line (*por exemplo*, GeneArt), métodos publicados ou uma empresa que fornece serviços de otimização de codificação, *por exemplo*, como DNA2.0 (Menlo Park, CA). Um método de otimização por códon é descrito, *por exemplo*, na Publicação de Patente Internacional WO 2015/012924, que é aqui incorporado por referência na sua totalidade. *Ver também, por exemplo*, a Publicação de Patente U.S. 2014/0032186 e a Publicação de Patente U.S. 2006/0136184. Adequadamente, o comprimento total da estrutura de leitura aberta (ORF) para o produto é modificado. No entanto, em algumas modalidades, apenas um fragmento da ORF pode ser alterado. Utilizando um destes métodos, pode-se aplicar as frequências a qualquer sequência polipeptídica e produzir um fragmento de ácido nucleico de uma região de codificação otimizada por códons que codifica o polipeptídeo. Estão disponíveis várias opções para realizar alterações reais aos códons ou para sintetizar as regiões de codificação otimizadas por códons concebidas como aqui descrito. Tais modificações ou sínteses podem ser realizadas usando manipulações biológicas moleculares padrão e de rotina bem conhecidas pelos versados na técnica. Numa abordagem, uma série de pares de oligonucleotídeos complementares de 80-90 nucleotídeos cada um em comprimento e abrangendo o comprimento da sequência desejada são sintetizados por métodos-padrão. Estes pares de oligonucleotídeos são sintetizados de modo a que, após recozimento, formem fragmentos de fita dupla de 80-90 pares de bases, *contendo extremidades coesivas, por exemplo*, cada oligonucleotídeo no par é sintetizado para estender 3, 4, 5, 6, 7, 8, 9 10, ou mais bases além da região que é complementar

ao outro oligonucleotídeo no par. As extremidades de fita simples de cada par de oligonucleotídeos são concebidas para recozer com a extremidade de fita simples de outro par de oligonucleotídeos. Os pares de oligonucleotídeos são autorizados a recozer e aproximadamente cinco a seis destes fragmentos de fita dupla são então permitidos recozerem juntos através das extremidades coesivas de fita simples, e depois são ligados e clonados num vetor de clonagem bacteriano padrão, por exemplo, um Vetor TOPO® disponível na Invitrogen Corporation, Carlsbad, Calif. O construto é, então, sequenciado por métodos-padrão. Vários destes construtos consistem em 5 a 6 fragmentos de fragmentos de 80 a 90 pares de bases ligados entre si, isto é, são preparados fragmentos de cerca de 500 pares de bases, de tal modo que a sequência inteira desejada é representada numa série de construtos de plasmídeo. Os insertos destes plasmídeos são, então, cortados com enzimas de restrição apropriadas e ligadas em conjunto para formar o construto final. O construto final é, então, clonado num vetor de clonagem bacteriana padrão e sequenciado. Métodos adicionais seriam imediatamente evidentes para o versado na técnica. Além disso, a síntese genética está prontamente disponível comercialmente.

[0071] Em certas modalidades, são fornecidos capsídeos de AAV que têm uma população heterogênea de isoformas de capsídeo de AAV (isto é, VP1, VP2, VP3) que contêm múltiplas posições "NG" altamente desamidadas. Em certas modalidades, as posições altamente desamidadas estão nos locais identificados abaixo, com referência à sequência de aminoácidos de VP1 de comprimento total predita. Em outras modalidades, o gene do capsídeo é modificado de modo que o "NG" referenciado seja removido e um "NG" mutante seja projetado em outra posição.

[0072] Em certas modalidades, a população mista de rAAV resulta

de um sistema de produção usando uma única sequência de ácido nucleico do capsídeo de AAV que codifica uma sequência de aminoácidos de VP1 de AAV predita de um tipo de AAV. No entanto, o processo de produção e fabricação fornece a população heterogênea de proteínas do capsídeo descritas acima.

[0073] Em certas modalidades, um novo capsídeo de AAVrh79 isolado é fornecido. A sequência de ácido nucleico que codifica o AAV é fornecida na SEQ ID NO: 1 e a sequência de aminoácidos codificada é fornecida na SEQ ID NO: 2.

[0074] Em certas modalidades, um rAAV compreende um capsídeo AAVrh79. Um capsídeo AAVrh79 compreende uma população heterogênea de proteínas AAVrh79 vp1, proteínas AAVrh79 vp2 e proteínas AAVrh79 vp3. Em uma modalidade, o capsídeo de AAVrh79 é produzido pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 de SEQ ID NO: 2. Opcionalmente, as sequências que coexpressam a proteína vp3 de uma sequência de ácido nucleico excluindo a região única de vp1 (cerca de aa 1 a 137) ou a região única de vp2 (cerca de aa 1 a 203), proteínas vp1 produzidas a partir de SEQ ID NO: 1, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de 1 a 738 de SEQ ID NO: 2. Em outras modalidades, as proteínas AAVrh79 vp2 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 738 de SEQ ID NO: 2, proteínas vp2 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 412 a 2214 da SEQ ID NO: 1, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 412 a 2214 da SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 138 a 738 de

SEQ ID NO: 2, proteínas AAVrh79 vp3 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 204 a 738 de SEQ ID NO: 2, proteínas vp3 produzidas a partir de uma sequência que compreende pelo menos os nucleotídeos 610 a 2214 da SEQ ID NO: 1, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 610 a 2214 da SEQ ID NO: 1 que codificam a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 204 a 738 de SEQ ID NO: 2.

[0075] Em certas modalidades, um capsídeo AAVrh79 compreende: uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de SEQ ID NO: 2, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca de aminoácidos 138 a 738 de SEQ ID NO: 2, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 de SEQ ID NO: 2.

[0076] As proteínas AAVrh79vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina em SEQ ID NO: 2 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácido. Níveis elevados de desamidação nos pares NG N57, N263, N385 e/ou N514 são observados, em relação ao número de SEQ ID NO: 2. A desamidação foi observada em outros resíduos, como mostrado na tabela abaixo e nos exemplos. Em certas modalidades, AAVrh79 pode ter outros resíduos desamidados, por exemplo, tipicamente em menos de 10% e/ou pode

ter outras modificações, incluindo metilação (por exemplo, ~R487) (tipicamente menos que 5%, mais tipicamente menos que 1% em um dado resíduo), isomerização (por exemplo, em D97) (tipicamente menos que 5%, mais tipicamente menos que 1% em um determinado resíduo), fosforilação (por exemplo, quando presente, na faixa de cerca de 10 a cerca de 60%, ou cerca de 10 a cerca de 30%, ou cerca de 20 a cerca de 60%) (por exemplo, em um ou mais de S149, ~S153, ~S474, ~T570, ~S665) ou oxidação (por exemplo, em um ou mais de W248, W307, W307, M405, M437, M473, W480, W480, W505, M526, M544, M561, W621, M637, e/ou W697). Opcionalmente, o W pode oxidar em quinurenina.

Tabela A - Desamidação AAVrh79

Desamidação de AAVrh79 baseada na numeração VP1	% de Desamidação
N57+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N94+Desamidação	5 - 15, cerca de 10
~N254+Desamidação	10 - 20
~N263+Desamidação	75 - 100
~N305+Desamidação	1 - 5
~N385+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
~N410+Desamidação	1 - 25,
N479+Desamidação	1 - 5, 1-3
~N514+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
~Q601+Desamidação	0-1
N653+Desamidação	0 - 2

[0077] Em certas modalidades, um capsídeo AAVrh79 é modificado em uma ou mais das posições identificadas na tabela anterior, nos intervalos fornecidos abaixo, conforme determinado usando espectrometria de massa com uma enzima tripsina. Em certas modalidades, uma ou mais das seguintes posições ou a glicina após o N é modificada conforme descrito neste documento. Os números de resíduos são baseados na sequência AAVrh79 aqui fornecida. Ver, SEQ ID NO: 2.

[0078] Em certas modalidades, a sequência de ácido nucleico que codifica a proteína do capsídeo AAVrh79 vp1 é fornecida na SEQ ID NO: 1. Em outras modalidades, uma sequência de ácido nucleico de 70% a 99,9% de identidade com a SEQ ID NO: 1 pode ser selecionada

para expressar as proteínas do capsídeo AAVrh79. Em certas outras modalidades, a sequência de ácido nucleico é pelo menos cerca de 75% idêntica, pelo menos 80% idêntica, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97% idêntica ou pelo menos 99% a 99,9% idêntica à SEQ ID NO: 1. No entanto, outras sequências de ácido nucleico que codificam a sequência de aminoácidos de SEQ ID NO: 2 podem ser selecionadas para uso na produção de capsídeos de rAAV. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 1 ou uma sequência pelo menos 70% a 99% idêntica, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica à SEQ ID NO: 1 que codifica a SEQ ID NO: 2. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 1 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a cerca de nt412 a cerca de nt 2214 de SEQ ID NO: 1 que codifica a proteína de capsídeo vp2 (cerca de aa 138 a 738) de SEQ ID NO: 2. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de cerca de nt610 a cerca de nt 2214 da SEQ ID NO: 1 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a nt SEQ ID NO: 1 que codifica a proteína de capsídeo vp3 (cerca de aa 204 a 738) de SEQ ID NO: 2.

[0079] A invenção também abrange sequências de ácido nucleico que codificam o AAVrh79 mutante, em que um ou mais resíduos foram alterados a fim de diminuir a desamidação ou outras modificações que são identificadas aqui. Essas sequências de ácido nucleico podem ser usadas na produção de capsídeos de rAAVrh79 mutantes.

[0080] Em certas modalidades, um novo capsídeo AAV8.AR2.08 é

fornecido. A sequência de ácido nucleico que codifica o AAV é fornecida na SEQ ID NO: 17 e a sequência de aminoácidos codificada é fornecida na SEQ ID NO: 18. Em uma modalidade, um vírus adenoassociado recombinante (rAAV) tem um capsídeo AAV8.AR2.08. Um alinhamento das sequências de aminoácidos de AAV8T, AAV8.AR2.08 e AAV8 é fornecido na FIG. 13A. Um alinhamento das sequências de ácido nucleico de AAV8T, AAV8.AR2.08 e AAV8 é fornecido na FIG. 13B-13D.

[0081] Em certas modalidades, um capsídeo AAV8.AR2.08 compreende proteínas de capsídeo AAV8.AR2.08 compreendendo: proteínas AAV8.AR2.08vp1 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18, proteínas vp1 produzidas a partir da SEQ ID NO: 17, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18, proteínas AAV8.AR2.08vp2 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, proteínas vp2 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 412 a 2214 da SEQ ID NO: 17, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica aos pelo menos nucleotídeos 412 a 2214 da SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, proteínas AAV8.AR2.08vp3 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 18, proteínas vp3 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 610 a 2214 da SEQ ID NO: 17, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos



70% idêntica aos pelo menos nucleotídeos 612 a 2214 da SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 18.

[0082] Adicionalmente, ou alternativamente, um capsídeo AAV8.AR2.08 compreende: uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos da SEQ ID NO: 18, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 da SEQ ID NO: 18, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 18 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em um mudança de aminoácidos. AAV8.AR2.08 é caracterizado por ter resíduos altamente desamidados, por exemplo, nas posições N57, N263, N385, N514 e N540 com base na numeração do AAV8.AR2.08VP1 [SEQ ID NO: 18]. Além disso, os resíduos nas posições a seguir à tabela e a tabela detalhada no pedido mostram as desamidações que foram observadas no capsídeo AAV8.AR2.08.

[0083] Em certas modalidades, um capsídeo AAV8.AR2.08 é modificado em uma ou mais das seguintes posições, nas faixas fornecidas abaixo, conforme determinado usando espectrometria de massa com uma enzima tripsina. Em certas modalidades, uma ou mais das seguintes posições ou a glicina após o N é modificada conforme descrito neste documento. Por exemplo, em certas modalidades, um G pode ser modificado para um S ou um A, por exemplo, na posição 58, 264, 386, 515

ou 541. Redução significativa na desamidação é observada quando NG57/58 é alterado para NS 57/58 ou NA57/58. No entanto, em certas modalidades, um aumento na desamidação é observado quando NG é alterado para NS ou NA. Em certas modalidades, um N de um par NG é modificado para um Q enquanto retém o G. Em certas modalidades, ambos os aminoácidos de um par NG são modificados. Em certas modalidades, N385Q resulta em redução significativa da desamidação nesse local. Em certas modalidades, o N499Q resulta em um aumento significativo da desamidação nesse local.

[0084] Além da desamidação, outras modificações podem incluir isomerização (por exemplo, em um ou mais de D442 e/ou D584) (1-15%), fosforilações (por exemplo, em um ou mais de ~S149, ~T417, ~T454, ~T493, S600 e/ou ~T663) e/ou oxidações (por exemplo, em uma ou mais das posições ~W22, ~M204, ~M212, W248, W307, M405, M437, M473, W480, W505,

[0085] M526, M561, M607, ~W609, W621, M637, W697). Ainda outras posições podem ter essas ou outras modificações (por exemplo, acetilação ou desamidações adicionais).

Tabela B - Desamidação de AAV8.AR2.08

Desamidação de AAV8. baseada na numeração VP1 SEQ ID NO: 18	%
N57+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N94+Desamidação	1-15
~N254+Desamidação	1-15
~N263+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
~N305+Desamidação	1-15
~N385+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
~N514+Desamidação	65 - 100
~N521+Desamidação	1-10
~ N540+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N590+Desamidação	0-5
Q601+Desamidação	0-5
N653+Desamidação	0-5
N665+Desamidação	0-5

[0086] Em certas modalidades, a sequência de ácido nucleico que codifica a proteína de capsídeo AAV8.AR2.08vp1 é fornecida na SEQ ID NO: 17. Em outras modalidades, uma sequência de ácido nucleico de 70% a 99,9% de identidade com a SEQ ID NO: 17 pode ser selecionada para expressar as proteínas de capsídeo AAV8.AR2.08. Em certas outras modalidades, a sequência de ácido nucleico é pelo menos cerca de 75% idêntica, pelo menos 80% idêntica, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97% idêntica ou pelo menos 99% a 99,9% idêntica à SEQ ID NO: 17. No entanto, outras sequências de ácido nucleico que codificam a sequência de aminoácidos de SEQ ID NO: 18 podem ser selecionadas para uso na produção de capsídeos de rAAV8.AR2.08. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 17 ou uma sequência pelo menos 70% a 99% idêntica, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica à SEQ ID NO: 17 que codifica a SEQ ID NO: 18. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 17 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a cerca de nt412 a cerca de nt 2214 de SEQ ID NO: 17 que codifica a proteína de capsídeo vp2 (cerca de aa 138 a 738) de SEQ ID NO: 18. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de cerca de nt 607 a cerca de nt 2214 da SEQ ID NO: 17 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a nt SEQ ID NO: 17 que codifica a proteína de capsídeo vp3 (cerca de aa 204 a 738) de SEQ ID NO: 18.

[0087] A invenção também abrange sequências de ácido nucleico

que codificam o AAV8.2.08 mutante, em que um ou mais resíduos foram alterados a fim de diminuir a desamidação ou outras modificações que são identificadas aqui. Essas sequências de ácido nucleico podem ser usadas na produção de rAAV8.2.08 mutante.

[0088] Em certas modalidades, um novo capsídeo AAV5.5.9 é fornecido. A sequência de ácido nucleico que codifica o AAV é fornecida na SEQ ID NO: 9 e a sequência de aminoácidos codificada é fornecida na SEQ ID NO: 10. Um alinhamento das sequências de aminoácidos de AAV5.5.9, AAV9 e AAVPHP.B é mostrado na FIG. 12A. Um alinhamento das sequências de ácido nucleico de AAV5.5.9, AAV9 e AAVPHP.B é mostrado na Fig. 12B-12E. Em uma modalidade, um vírus adenoassociado recombinante (rAAV) tem um capsídeo AAV5.5.9 compreendendo: proteínas do capsídeo AAV5.5.9 compreendendo: uma população heterogênea de proteínas AAV5.5.9 vp1 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 726 de SEQ ID NO: 10, proteínas vp1 produzidas a partir de SEQ ID NO: 9 ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de 1 a 726 de SEQ ID NO: 1; uma população heteróloga de proteínas AAV5.5.9 vp2 produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 137 a 726 de SEQ ID NO: 10, proteínas vp2 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 409 a 2178 de SEQ ID NO: 9, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 577 a 2178 de SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 137 a 726 de SEQ ID NO: 10, e uma população heteróloga de proteínas AAV5.5.9 vp3

produzidas pela expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca de aminoácidos 193 a 726 de SEQ ID NO: 10, proteínas vp3 produzidas a partir de uma sequência compreendendo pelo menos os nucleotídeos 577 a 2178 da SEQ ID NO: 9, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 577 a 2178 da SEQ ID NO: 9 que codificam a sequência de aminoácidos de pelo menos cerca dos aminoácidos 193 a 726 da SEQ ID NO: 10.

[0089] Adicionalmente ou alternativamente, um capsídeo AAV5.5.9 compreende: uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos da SEQ ID NO: 10, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca dos aminoácidos 137 a 726 da SEQ ID NO: 10, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 193 a 726 da SEQ ID NO: 10, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 10 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos.

Tabela C - Desamidação de AAV5.5.9

AAV5.5.9 baseada na numeração VP1 SEQ ID NO: 10	%
N35+Desamidação	0-15, 1 - 10
-N57+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N113+Desamidação	0-15
-N204+Desamidação	0-20, 1-20
N217+Desamidação	0-5, 1-5
-N243+Desamidação	0 - 25, 1-25

AAV5.5.9 baseada na numeração VP1 SEQ ID NO: 10	%
Q249+Desamidação	1-20
N293/294+Desamidação	10 - 45, 15 - 40
N304+Desamidação	1-10
N309+Desamidação	1-2
Q311+Desamidação	1-2
-N319+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N399/400+Desamidação	5 - 40, 10- 40, 15 - 35
-N442+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N467+Desamidação	1-5
N502+Desamidação	65-90, 70-95, 80-95, 75 - 100, 80-100, ou 90-100
N505+Desamidação	5-25, 10-25
-Q589+Desamidação	5-30, 10 - 30
N618+Desamidação	1 - 15, 5 - 10
- N641+Desamidação	1 -15, 5 - 10
N653+Desamidação	1 - 15, 5-10
-N658+Desamidação	5 - 40, 10 - 30
N694+Desamidação	0-5
-N699+Desamidação	1 - 10

[0090] Em certas modalidades, a sequência de ácido nucleico que codifica a proteína do capsídeo AAV5.5.9 vp1 é fornecida na SEQ ID NO: 9. Em outras modalidades, uma sequência de ácido nucleico de 70% a 99,9% de identidade com a SEQ ID NO: 9 pode ser selecionada para expressar as proteínas do capsídeo AAV5.5.9. Em certas outras modalidades, a sequência de ácido nucleico é pelo menos cerca de 75% idêntica, pelo menos 80% idêntica, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97% idêntica ou pelo menos 99% a 99,9% idêntica à SEQ ID NO: 9. No entanto, outras sequências de ácido nucleico que codificam a sequência de aminoácidos de SEQ ID NO: 10 podem ser selecionadas para uso na produção de capsídeos de rAAV5.5.9. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 10 ou uma sequência pelo menos 70% a 99% idêntica, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica à SEQ ID NO: 9 que codifica a SEQ ID NO:

10. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de SEQ ID NO: 9 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a cerca de nt409 a cerca de nt2178 de SEQ ID NO: 9 que codifica a proteína de capsídeo vp2 (cerca de aa 137 a 726) de SEQ ID NO: 10. Em certas modalidades, a sequência de ácido nucleico tem a sequência de ácido nucleico de cerca de nt577 a cerca de nt2178 da SEQ ID NO: 9 ou uma sequência pelo menos 70% a 99%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 97%, pelo menos 99%, idêntica a nt577 a 2178 SEQ ID NO: 9 que codifica a proteína de capsídeo vp3 (cerca de aa 193 a 726) de SEQ ID NO: 10.

[0091] A invenção também abrange sequências de ácido nucleico que codificam o AAV5.5.9 mutante, em que um ou mais resíduos foram alterados a fim de diminuir a desamidação ou outras modificações que são identificadas aqui. Essas sequências de ácido nucleico podem ser utilizadas na produção de rAAV5.5.9 mutante.

#### I. Vetores rAAV

[0092] Como indicado acima, as novas sequências e proteínas de AAV são úteis na produção de rAAV e também são úteis em vetores de AAV recombinantes que podem ser vetores de distribuição antissenso, vetores de terapia genética ou vetores de vacina. Além disso, os capsídeos de AAV engenheirados aqui descritos podem ser usados para engenheirar vetores de rAAV para distribuição de uma série de moléculas de ácido nucleico adequadas a células e tecidos alvo.

[0093] As sequências genômicas que são empacotadas em um capsídeo de AAV e distribuídas a uma célula hospedeira são tipicamente compostas por, no mínimo, um transgene e suas sequências regulatórias e repetições terminais invertidas de AAV (ITRs). Ambos AAV

de fita simples e AAV autocomplementar (sc) estão incluídos no rAAV. O transgene é uma sequência de codificação de ácido nucleico, heterogênea às sequências do vetor, que codifica um polipeptídeo, proteína, molécula de RNA funcional (*por exemplo*, miRNA, inibidor de miRNA) ou outro produto de gene de interesse. A sequência de codificação de ácido nucleico está operacionalmente ligada a componentes regulatórios de uma maneira que permite a transcrição, tradução e/ou expressão do transgene em uma célula de um tecido alvo.

[0094] As sequências AAV do vetor compreendem tipicamente as sequências de repetição terminal invertida 5' e 3' de ação cis (*Ver, por exemplo*, B. J. Carter, in "Handbook of Parvoviruses", ed., P. Tijsser, CRC Press, pp. 155 168 (1990)). As sequências ITR têm cerca de 145 bp de comprimento. De preferência, substancialmente todas as sequências que codificam as ITRs são utilizadas na molécula, embora algum grau de modificação menor dessas sequências seja permissível. A capacidade de modificar essas sequências ITR está dentro da habilidade na técnica. (*Ver, por exemplo*, textos como Sambrook *et al*, "Molecular Cloning. A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory, New York (1989); and K. Fisher et al., J. Virol., 70:520 532 (1996)). Um exemplo de tal molécula empregada na presente invenção é um plasmídeo de "ação cis" contendo o transgene, no qual a sequência do transgene selecionada e os elementos regulatórios associados são flanqueados pelas sequências ITR 5' e 3' AAV. Em uma modalidade, as ITRs são de um AAV diferente daquele que fornece um capsídeo, resultando em um vetor pseudotipado. Em uma modalidade, as sequências ITR de AAV2. Uma versão abreviada da ITR 5', denominada  $\Delta$ ITR, foi descrita na qual o sítio de resolução de terminal e sequência-D (trs) são eliminados. Em outras modalidades, são utilizadas as ITRs AAV 5'e 3' de comprimento total. No entanto, as ITRs de outras fontes de AAV podem ser selecionadas. Onde a fonte das ITRs é de AAV2 e o capsídeo



AAV é de outra fonte de AAV, o vetor resultante pode ser denominado pseudotipado. No entanto, outras configurações desses elementos podem ser adequadas.

[0095] Além dos elementos principais identificados acima para o vetor AAV recombinante, o vetor também inclui elementos de controle convencionais necessários que estão operacionalmente ligados ao transgene de uma maneira que permite sua transcrição, tradução e/ou expressão em uma célula transfectada com o vetor de plasmídeo ou infectado com o vírus produzido pela invenção. Como usado neste documento, sequências “operacionalmente ligadas” incluem ambas as sequências de controle de expressão que são contíguas com o gene de interesse e sequências de controle de expressão que atuam in trans ou a uma distância para controlar o gene de interesse.

[0096] Os elementos de controle regulatórios tipicamente contêm uma sequência promotora como parte das sequências de controle de expressão, por exemplo, localizada entre a sequência 5' ITR selecionada e a sequência de codificação. Os promotores constitutivos, promotores reguláveis [ver, por exemplo, WO 2011/126808 e WO 2013/04943], promotores específicos de tecido, ou um promotor responsivo a pistas fisiológicas podem ser utilizados nos vetores aqui descritos. O(s) promotor(es) podem ser selecionados a partir de fontes diferentes, *por exemplo*, intensificador/promotor imediato-precoce de citomegalovírus humano (CMV), o intensificador/promotor precoce de SV40, o promotor de poliomavírus JC, proteína básica de mielina (MBP) ou promotores de proteína glial fibrilar acídica (GFAP), promotor associado à latência (LAP) do vírus herpes simplex (HSV-1), promotor de repetição do terminal longo (LTR) do vírus do rouse sarcoma (RSV), promotor específico de neurônio (NSE), promotor do fator de crescimento derivado de plaquetas (PDGF), hSYN, promotor do hormônio de concentração de melanina (MCH), CBA, promotor de metaloproteína de matriz (MPP) e

o promotor de beta-actina de frango. Em uma modalidade, o promotor é um promotor específico do fígado, tal como aquele denominado LSP aqui exemplificado.

[0097] Além de um promotor, um vetor pode conter uma ou mais outras sequências de iniciação, terminação, intensificadoras de transcrição apropriadas, sinais de processamento de RNA eficientes, tais como sinais de emenda e poliadenilação (poliA); sequências que estabilizam o mRNA citoplasmático, por exemplo WPRE; sequências que intensificam a eficiência da tradução (ou seja, sequência de consenso Kozak); sequências que intensificam a estabilidade da proteína; e quando desejado, sequências que intensificam a secreção do produto codificado. Um exemplo de um intensificador adequado é o intensificador de CMV. Outros intensificadores adequados incluem aqueles que são apropriados para as indicações desejadas do tecido alvo. Em uma modalidade, o cassete de expressão compreende um ou mais intensificadores de expressão. Em uma modalidade, o cassete de expressão contém dois ou mais intensificadores de expressão. Esses intensificadores podem ser iguais ou podem diferir uns dos outros. Por exemplo, um intensificador pode incluir um intensificador precoce imediato de CMV. Este intensificador pode estar presente em duas cópias localizadas adjacentes uma à outra. Alternativamente, as cópias duplas do intensificador podem ser separadas por uma ou mais sequências. Em ainda outra modalidade, o cassete de expressão contém ainda um íntron, por exemplo, o íntron de beta-actina de frango. Outros íntrons adequados incluem aqueles conhecidos na técnica, por exemplo, como são descritos em WO 2011/126808. Exemplos de sequências poliA adequadas incluem, por exemplo, SV40, SV50, hormônio de crescimento bovino (bGH), hormônio de crescimento humano e poliAs sintéticos. Opcionalmente, uma ou mais sequências podem ser selecionadas para estabilizar o mRNA. Um exemplo de tal sequência é uma sequência WPRE modificada, que

pode ser projetada a montante da sequência poliA e a jusante da sequência de codificação [ver, por exemplo, MA Zanta-Boussif, et al, Gene Therapy (2009) 16: 605-619.

[0098] Estes rAAVs são particularmente adequados para distribuição de genes para fins terapêuticos e para imunização, incluindo a indução de imunidade protetora. Além disso, as composições da invenção também podem ser usadas para a produção de um produto de gene desejado *in vitro*. Para a produção *in vitro*, um produto desejado (*por exemplo*, uma proteína) pode ser obtido a partir de uma cultura desejada após a transfecção de células hospedeiras com um rAAV contendo a molécula que codifica o produto desejado e cultivando a cultura de células em condições que permitem a expressão. O produto expresso pode então ser purificado e isolado, conforme desejado. As técnicas adequadas para transfecção, cultura de células, purificação e isolamento são conhecidas dos versados na técnica.

#### Transgenes terapêuticos

[0099] Produtos úteis codificados pelo transgene incluem uma variedade de produtos de gene que substituem um gene defeituoso ou deficiente, inativam ou "knock-out", ou "knock-down" ou reduzem a expressão de um gene que está se expressando em um nível indesejavelmente alto, ou distribuindo um produto de gene que tem um efeito terapêutico desejado. Na maioria das modalidades, a terapia será "terapia genética somática", ou seja, transferência de genes para uma célula do corpo que não produz espermatozoides ou óvulos. Em certas modalidades, os transgenes expressam proteínas têm a sequência de sequências humanas nativas. No entanto, em outras modalidades, proteínas sintéticas são expressas. Essas proteínas podem ser destinadas ao tratamento de humanos, ou em outras modalidades, projetadas para o tratamento de animais, incluindo animais de companhia, como populações caninas ou felinas, ou para o tratamento de gado ou outros animais que

entram em contato com populações humanas.

[00100] Exemplos de produtos de genes adequados podem incluir aqueles associados com hipercolesterolemia familiar, distrofia muscular, fibrose cística e doenças raras ou órfãs. Exemplos de tais doenças raras podem incluir atrofia muscular espinhal (SMA), Doença de Huntington, Síndrome de Rett (por exemplo, proteína 2 de ligação a metil-CpG (MeCP2); UniProtKB - P51608), Esclerose Lateral Amiotrófica (ALS), distrofia Muscular do Tipo Duchenne, Friedrichs Ataxia (por exemplo, frataxina), progranulina (PRGN) (associada a degenerações cerebrais não relacionadas ao Alzheimer, incluindo demência frontotemporal (FTD), afasia não fluente progressiva (PNFA) e demência semântica), entre outras. Veja, por exemplo, [www.orpha.net/consor/cgi-bin/Disease\\_Search\\_List.php](http://www.orpha.net/consor/cgi-bin/Disease_Search_List.php); [rarediseases.info.nih.gov/diseases](http://rarediseases.info.nih.gov/diseases).

[00101] Exemplos de genes adequados podem incluir, por exemplo, hormônios e fatores de crescimento e diferenciação incluindo, sem limitação, insulina, glucagon, peptídeo -1 semelhante ao glucagon (GLP1), hormônio do crescimento (GH), hormônio da paratireoide (PTH), fator de liberação do hormônio do crescimento (GRF), hormônio folículo estimulante (FSH), hormônio luteinizante (LH), gonadotrofina coriônica humana (hCG), fator de crescimento endotelial vascular (VEGF), angiopoi-etinas, angiostatina, fator estimulador de colônia de granulócitos (GCSF), eritropoietina (EPO) (incluindo, por exemplo, epo humano, canino ou felino), fator de crescimento do tecido conjuntivo (CTGF), fatores neutróficos, incluindo, por exemplo, fator de crescimento de fibroblasto básico (bFGF), fator de crescimento de fibroblasto ácido (aFGF), fator de crescimento epidérmico (EGF), fator de crescimento derivado de plaquetas (PDGF), fatores de crescimento de insulina I e II (IGF-I e IGF-II), qualquer um da superfamília do fator de crescimento de transformação  $\alpha$ , incluindo TGF $\alpha$ , ativinas, inibinas ou qualquer uma das proteínas morfogênicas ósseas (BMP) BMPs 1-15, qualquer um da família de fatores

de crescimento hereglúina/neuregulina/ARIA/fator de diferenciação neu (NDF), fator de crescimento do nervo (NGF), fator neurotrófico derivado do cérebro (BDNF), neurotrofinas NT-3 e NT-4/5, fator neurotrófico ciliar (CNTF), fator neurotrófico derivado da linhagem de células gliais (GDNF), neurturina, agrina, qualquer um da família das semaforinas/collapsinas, netrin-1 e netrin-2, fator de crescimento de hepatócitos (HGF), efrinas, nogina, hedgehog sônico e tirosina hidroxilase.

[00102] Outros produtos transgênicos úteis incluem proteínas que regulam o sistema imunológico, incluindo, sem limitação, citocinas e linfocinas, tais como trombopoietina (TPO), interleucinas (IL) IL-1 a IL-36 (incluindo, *por exemplo*, interleucinas humanas IL-1, IL-1 $\alpha$ , IL-1 $\beta$ , IL-2, IL-3, IL-4, IL-6, IL-8, IL-12, IL-11, IL-12, IL-13, IL-18, IL-31, IL-35), proteína quimioatraente de monócitos, fator inibidor de leucemia, fator de estimulação de colônia de granulócitos-macrófagos, ligando Fas, fatores de necrose tumoral  $\alpha$  e  $\beta$ , interferons  $\alpha$ ,  $\beta$  e  $\gamma$ , fator de células-tronco, ligando flk-2/flt3. Os produtos de gene produzidos pelo sistema imunológico também são úteis na invenção. Estes incluem, sem limitações, imunoglobulinas IgG, IgM, IgA, IgD e IgE, imunoglobulinas quiméricas, anticorpos humanizados, anticorpos de cadeia única, receptores de células T, receptores de células T quiméricos, receptores de células T de cadeia simples, moléculas de MHC de classe I e classe II, bem como imunoglobulinas engenheiradas e moléculas de MHC. Por exemplo, em certas modalidades, os anticorpos rAAV podem ser projetados para distribuir anticorpos caninos ou felinos, por exemplo, tais como anti-IgE, anti-IL31, anti-CD20, anti-NGF, anti-GnRH. Produtos de gene úteis também incluem proteínas regulatórias do complemento, como proteínas regulatórias do complemento, proteína cofator de membrana (MCP), fator de aceleração do decaimento (DAF), CR1, CF2, CD59 e inibidor da esterase C1 (C1-INH).

[00103] Ainda outros produtos de genes úteis incluem qualquer um

dos receptores para os hormônios, fatores de crescimento, citocinas, linfocinas, proteínas regulatórias e proteínas do sistema imunológico. A invenção abrange receptores para a regulação do colesterol e/ou modulação lipídica, incluindo o receptor de lipoproteína de baixa densidade (LDL), receptor de lipoproteína de alta densidade (HDL), receptor de lipoproteína de densidade muito baixa (VLDL) e receptores sequestrantes. A invenção também abrange produtos de genes, tais como membros da superfamília de receptores de hormônio esteroide, incluindo receptores de glicocorticoides e receptores de estrogênio, receptores de vitamina D e outros receptores nucleares. Além disso, os produtos de gene úteis incluem fatores de transcrição, como *jun*, *fos*, *max*, *mad*, fator de resposta do soro (SRF), AP-1, AP2, *myb*, MyoD e miogenina, proteínas contendo caixa ETS, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, proteínas de ligação de caixa CCAAT, fator de regulação de interferon (IRF-1), proteína tumoral de Wilms, proteína de ligação ETS, STAT, proteínas de ligação de caixa GATA, *por exemplo*, GATA-3 e a família forkhead de proteínas de hélice alada.

[00104] Outros produtos de genes úteis incluem carbamoil sintetase I, ornitina transcarbamilase (OTC), arginossuccinato sintetase, arginossuccinato liase (ASL) para o tratamento de deficiência de arginossuccinato liase, arginase, fumarilacetato hidrolase, fenilalanina hidroxilase fetal, alfa-1 antitripsina, alfa-fetoproteína rhesus (AFP), gonadotrofina coriônica rhesus (CG), glicose-6-fosfatase, porfobilinogênio desaminase, cistationa beta-sintase, cetoácido descarboxilase de cadeia ramificada, albumina, isovaleril-coA desidrogenase, propionil CoA carboxilase, metil malonil CoA mutase, glutaril CoA desidrogenase, insulina, beta-glucosidase, piruvato carboxilato, fosforilase hepática, fosforilase quinase, glicina descarboxilase, proteína H, proteína T, uma sequência reguladora transmembranar de fibrose cística (CFTR), e um produto do gene da

distrofina [*por exemplo*, uma mini- ou microdistrofina]. Ainda outros produtos de genes úteis incluem enzimas, tais como podem ser úteis na terapia de substituição de enzima, que é útil em uma variedade de condições resultantes da atividade deficiente da enzima. Por exemplo, enzimas que contêm manose-6-fosfato podem ser utilizadas em terapias para doenças de armazenamento lisossômico (*por exemplo*, um gene adequado inclui aquele que codifica  $\beta$ -glucuronidase (GUSB)).

[00105] Em certas modalidades, o rAAV pode ser usado em sistemas de edição de genes, cujo sistema pode envolver um rAAV ou coadministração de múltiplos estoques de rAAV. Por exemplo, o rAAV pode ser projetado para fornecer SpCas9, SaCas9, ARCUS, Cpf1 e outros construtos de edição de genes adequados.

[00106] Ainda outros produtos de genes úteis incluem aqueles usados para o tratamento de hemofilia, incluindo hemofilia B (incluindo Fator IX) e hemofilia A (incluindo Fator VIII e suas variantes, tais como a cadeia leve e a cadeia pesada do heterodímero e o domínio B-deletado; Patente U.S. 6.200.560 e Patente U.S. 6.221.349). Em algumas modalidades, o minigene compreende primeiros 57 pares de bases da cadeia pesada do Fator VIII que codifica a sequência de sinal de 10 aminoácidos, bem como a sequência de poliadenilação do hormônio de crescimento humano (hGH). Em modalidades alternativas, o minigene compreende ainda os domínios A1 e A2, bem como 5 aminoácidos do terminal N do domínio B e/ou 85 aminoácidos do terminal C do domínio B, bem como os domínios A3, C1 e C2. Em ainda outras modalidades, os ácidos nucleicos que codificam a cadeia pesada e a cadeia leve do Fator VIII são fornecidos em um único minigene separado por 42 ácidos nucleicos que codificam para 14 aminoácidos do domínio B [Patente U.S. 6.200.560].

[00107] Outros produtos de genes úteis incluem polipeptídeos de ocorrência não natural, tais como polipeptídeos quiméricos ou híbridos

com uma sequência de aminoácidos de ocorrência não natural contendo inserções, deleções ou substituições de aminoácidos. Por exemplo, imunoglobulinas de cadeia simples podem ser úteis em certos pacientes imunocomprometidos. Outros tipos de sequências de genes de ocorrência não natural incluem moléculas antissenso e ácidos nucleicos catalíticos, como ribozimas, que podem ser usados para reduzir a superexpressão de um alvo.

[00108] A redução e/ou modulação da expressão de um gene é particularmente desejável para o tratamento de condições hiperproliferativas caracterizadas por células hiperproliferativas, como são cânceres e psoríase. Os polipeptídeos-alvo incluem aqueles polipeptídeos que são produzidos exclusivamente ou em níveis mais elevados em células hiperproliferativas em comparação com células normais. Os antígenos alvo incluem polipeptídeos codificados por oncogenes, tais como myb, myc, fyn, e o gene de translocação bcr/abl, ras, src, P53, neu, trk e EGRF. Além de produtos oncogene como antígenos alvo, polipeptídeos-alvo para tratamentos anticâncer e regimes de proteção incluem regiões variáveis de anticorpos feitos por linfomas de células B e regiões variáveis de receptores de células T de linfomas de células T que, em algumas modalidades, também são usados como antígenos alvo para doenças autoimunes. Outros polipeptídeos associados a tumor podem ser usados como polipeptídeos-alvo, tais como polipeptídeos que são encontrados em níveis mais elevados em células tumorais, incluindo o polipeptídeo reconhecido pelo anticorpo monoclonal 17-1A e polipeptídeos de ligação a folato.

[00109] Outros polipeptídeos e proteínas terapêuticas adequadas incluem aqueles que podem ser úteis para o tratamento de indivíduos que sofrem de doenças e distúrbios autoimunes, conferindo uma resposta imune protetora de ampla base contra alvos que estão associados à autoimunidade, incluindo receptores celulares e células que produzem



anticorpos "auto" direcionados. As doenças autoimunes mediadas por células T incluem artrite reumatoide (AR), esclerose múltipla (MS), síndrome de Sjögren, sarcoidose, diabetes mellitus dependente de insulina (IDDM), tireoidite autoimune, artrite reativa, espondilite anquilosante, esclerodermia, polimiosite, dermatomiosite, psoríase, granulomatose de Wegener, doença de Crohn e colite ulcerosa. Cada uma dessas doenças é caracterizada por receptores de células T (TCRs) que se ligam a antígenos endógenos e iniciam a cascata inflamatória associada a doenças autoimunes.

[00110] Outros genes ilustrativos que podem ser distribuídos via rAAV incluem, sem limitação, glicose-6-fosfatase, associada à doença de armazenamento de glicogênio ou deficiência tipo 1A (GSD1), fosfoenolpiruvato-carboxicinase (PEPCK), associada à deficiência de PEPCK; tipo quinase dependente de ciclina 5 (CDKL5), também conhecido como serina / treonina quinase 9 (STK9) associada a convulsões e grave comprometimento do desenvolvimento neurológico; galactose-1 fosfato uridil transferase, associada à galactosemia; fenilalanina hidroxilase, associada à fenilcetonúria (PKU); alfa-cetoácido desidrogenase de cadeia ramificada, associada à doença da urina do xarope de bordo; fumarilacetoacetato hidrolase, associado à tirosinemia tipo 1; metilmalonil-CoA mutase, associada à acidemia metilmalônica; acil CoA desidrogenase de cadeia média, associada à deficiência de acetil CoA de cadeia média; ornitina transcarbamilase (OTC), associada com deficiência de ornitina transcarbamilase; ácido argininossuccínico sintetase (ASS1), associado a citrulinemia; deficiência de lecitina-colesterol aciltransferase (LCAT); acidemia ametilmalônica (MMA); doença de Niemann-Pick, tipo C1); acidemia propiônica (PA); proteína do receptor de lipoproteína de baixa densidade (LDLR), associada à hipercolesterolemia familiar (FH); UDP-glucouronosiltransferase, associada à doença

de Crigler-Najjar; adenosina desaminase, associada a doença de imunodeficiência combinada grave; hipoxantina guanina fosforibosil transferase, associada à síndrome de gota e Lesch-Nyan; biotimidase, associada à deficiência de biotimidase; alfa-galactosidase A (a-Gal A) associada à doença de Fabry); ATP7B associado à doença de Wilson; beta-glucocerebrosidase, associada à doença de Gaucher tipo 2 e 3; proteína de membrana de peroxissoma de 70 kDa, associada à síndrome de Zellweger; arilsulfatase A (ARSA) associada a leucodistrofia metacromática, enzima galactocerebrosidase (GALC) associada à doença de Krabbe, alfa-glucosidase (GAA) associada à doença de Pompe; gene da esfingomielinase (SMPD1) associado à doença de Nieman Pick tipo A; argininosuccinato sintase associada a citrulinemia tipo II de início na idade adulta (CTLN2); carbamoil-fosfato sintase 1 (CPS1) associada a distúrbios do ciclo da ureia; proteína do neurônio motor de sobrevivência (SMN), associada à atrofia muscular espinhal; ceramidase associada à lipogranulomatose de Farber; b-hexosaminidase associada à gangliosidose GM2 e doenças de Tay-Sachs e Sandhoff; aspartilglucosaminidase associada a aspartilglucosaminúria;  $\alpha$ -fucosidase associada à fucosidose;  $\alpha$ -manosidase associada à alfa-manosidose; porfobilinogênio desaminase, associado à porfiria aguda intermitente (AIP); alfa-1 antitripsina para o tratamento da deficiência de alfa-1 antitripsina (enfisema); eritropoietina para o tratamento de anemia devido à talassemia ou insuficiência renal; fator de crescimento endotelial vascular, angiotensinogênio, e fator de crescimento de fibroblasto para o tratamento de doenças isquêmicas; trombosmodulina e inibidor da via do fator de tecido para o tratamento de vasos sanguíneos obstruídos como visto em, por exemplo, aterosclerose, trombose ou embolias; aminoácido aromático descarboxilase (AADC) e tirosina hidroxilase (TH) para o tratamento da doença de Parkinson; o receptor beta adrenérgico, antissenso para, ou uma forma mutante de fosfolambano, o retículo sarco(endo)plasmático

adenosina trifosfatase-2 (SERCA2), e a adenilil ciclase cardíaca para o tratamento de insuficiência cardíaca congestiva; um gene supressor de tumor, como p53, para o tratamento de vários cânceres; uma citocina, tal como uma das várias interleucinas para o tratamento de doenças inflamatórias e imunes e cânceres; distrofina ou minidistrofina e utrofina ou miniutrofina para o tratamento de distrofias musculares; e insulina ou GLP-1 para o tratamento de diabetes.

[00111] Genes e doenças adicionais de interesse incluem, por exemplo, doenças relacionadas ao gene da distonina, como Neuropatia Hereditária Sensorial e Autonômica do Tipo VI (o gene DST codifica a distonina; vetores AAV duplos podem ser necessários devido ao tamanho da proteína (~7570 aa); doenças relacionadas ao SCN9A, nas quais os mutantes de perda de função causam incapacidade de sentir dor e os mutantes de ganho de função causam condições de dor, como eritromelalgia. Outra condição é Charcot-Marie-Tooth tipo 1F e 2E devido a mutações no gene NEFL (cadeia leve do neurofilamento), caracterizado por uma neuropatia periférica motora e sensorial progressiva com expressão clínica e eletrofisiológica variável.

[00112] Em certas modalidades, o rAAV aqui descrito pode ser usado no tratamento de distúrbios de mucopolissacaridoses (MPS). Tal rAAV pode conter transportar uma sequência de ácido nucleico que codifica  $\alpha$ -L-iduronidase (IDUA) para o tratamento de MPS I (síndromes de Hurler, Hurler-Scheie e Scheie); uma sequência de ácido nucleico que codifica iduronato-2-sulfatase (IDS) para o tratamento de MPS II (síndrome de Hunter); uma sequência de ácido nucleico que codifica sulfamidase (SGSH) para o tratamento de MPSIII A, B, C e D (síndrome de Sanfilippo); uma sequência de ácido nucleico que codifica N-acetilgalactosamina-6-sulfato sulfatase (GALNS) para o tratamento de MPS IV A e B (síndrome de Morquio); uma sequência de ácido nucleico que codifica

arilsulfatase B (ARSB) para o tratamento de MPS VI (síndrome de Maroteaux-Lamy); uma sequência de ácido nucleico que codifica hialuronidase para o tratamento de MPSI IX (deficiência de hialuronidase) e uma sequência de ácido nucleico que codifica beta-glucuronidase para o tratamento de MPS VII (síndrome de Sly).

#### Transgenes imunogênicos

[00113] Em algumas modalidades, um vetor rAAV compreendendo um ácido nucleico que codifica um produto de gene associado ao câncer (por exemplo, supressores de tumor) pode ser usado para tratar o câncer, administrando um rAAV que abriga o vetor rAAV a um sujeito com câncer. Em algumas modalidades, um vetor rAAV compreendendo um ácido nucleico que codifica um pequeno ácido nucleico interferente (por exemplo, shRNAs, miRNAs) que inibe a expressão de um produto de gene associado ao câncer (por exemplo, oncogenes) pode ser usado para tratar o câncer, administrando um rAAV que abriga o vetor rAAV a um sujeito com câncer. Em algumas modalidades, um vetor rAAV compreendendo um ácido nucleico que codifica um produto de gene associado ao câncer (ou um RNA funcional que inibe a expressão de um gene associado ao câncer) pode ser usado para fins de pesquisa, por exemplo, para estudar o câncer ou para identificar terapêuticas que tratam o câncer. O que se segue é uma lista não limitativa de genes exemplificativos conhecidos por estarem associados ao desenvolvimento de câncer (por exemplo, oncogenes e supressores de tumor): AARS, ABCB1, ABCC4, ABI2, ABL1, ABL2, ACK1, ACP2, ACY1, ADSL, AK1, AKR1C2, AKT1, ALB, ANPEP, ANXA5, ANXA7, AP2M1, APC, ARHGAP5, ARHGEF5, ARID4A, ASNS, ATF4, ATM, ATP5B, ATP5O, AXL, BARD1, BAX, BCL2, BHLHB2, BLMH, BRAF, BRCA1, BRCA2, BTK, CANX, CAP1, CAPN1, CAPNS1, CAV1, CBF, CBLB, CCL2, CCND1, CCND2, CCND3, CCNE1, CCT5, CCYR61, CD24, CD44, CD59, CDC20, CDC25, CDC25A, CDC25B, CDC2L5, CDK10, CDK4,

CDK5, CDK9, CDKL1, CDKN1A, CDKN1B, CDKN1C, CDKN2A, CDKN2B, CDKN2D, CEBPG, CENPC1, CGRRF1, CHAF1A, CIB1, CKMT1, CLK1, CLK2, CLK3, CLNS1A, CLTC, COL1A1, COL6A3, COX6C, COX7A2, CRAT, CRHR1, CSF1R, CSK, CSNK1G2, CTNNA1, CTNNB1, CTPS, CTSC, CTSD, CUL1, CYR61, DCC, DCN, DDX10, DEK, DHCR7, DHRS2, DHX8, DLG3, DVL1, DVL3, E2F1, E2F3, E2F5, EGFR, EGR1, EIF5, EPHA2, ERBB2, ERBB3, ERBB4, ERCC3, ETV1, ETV3, ETV6, F2R, FASTK, FBN1, FBN2, FES, FGFR1, FGR, FKBP8, FN1, FOS, FOSL1, FOSL2, FOXG1A, FOXO1A, FRAP1, FRZB, FTL, FZD2, FZD5, FZD9, G22P1, GAS6, GCN5L2, GDF15, GNA13, GNAS, GNB2, GNB2L1, GPR39, GRB2, GSK3A, GSPT1, GTF2I, HDAC1, HDGF, HMMR, HPRT1, HRB, HSPA4, HSPA5, HSPA8, HSPB1, HSPH1, HYAL1, HYOU1, ICAM1, ID1, ID2, IDUA, IER3, IFITM1, IGF1R, IGF2R, IGFBP3, IGFBP4, IGFBP5, IL1B, ILK, ING1, IRF3, ITGA3, ITGA6, ITGB4, JAK1, JARID1A, JUN, JUNB, JUND, K-ALPHA-1, KIT, KITLG, KLK10, KPNA2, KRAS2, KRT18, KRT2A, KRT9, LAMB1, LAMP2, LCK, LCN2, LEP, LITAF, LRPAP1, LTF, LYN, LZTR1, MADH1, MAP2K2, MAP3K8, MAPK12, MAPK13, MAPKAPK3, MAPRE1, MARS, MAS1, MCC, MCM2, MCM4, MDM2, MDM4, MET, MGST1, MICB, MLLT3, MME, MMP1, MMP14, MMP17, MMP2, MNDA, MSH2, MSH6, MT3, MYB, MYBL1, MYBL2, MYC, MYCL1, MYCN, MYD88, MYL9, MYLK, NEO1, NF1, NF2, NFKB1, NFKB2, NFSF7, NID, NINE, NMBR, NME1, NME2, NME3, NOTCH1, NOTCH2, NOTCH4, NPM1, NQO1, NR1D1, NR2F1, NR2F6, NRAS, NRG1, NSEP1, OSM, PA2G4, PABPC1, PCNA, PCTK1, PCTK2, PCTK3, PDGFA, PDGFB, PDGFRA, PDPK1, PEA15, PFDN4, PFDN5, PGAM1, PHB, PIK3CA, PIK3CB, PIK3CG, PIM1, PKM2, PKMYT1, PLK2, PPARD, PPARG, PPIH, PPP1CA, PPP2R5A, PRDX2, PRDX4, PRKAR1A, PRKCBP1, PRNP, PRSS15, PSMA1, PTCH, PTEN, PTGS1, PTMA, PTN, PTPRN, RAB5A, RAC1, RAD50, RAF1, RALBP1, RAP1A, RARA, RARB, RASGRF1,

RB1, RBBP4, RBL2, REA, REL, RELA, RELB, RET, RFC2, RGS19, RHOA, RHOB, RHOC, RHOD, RIPK1, RPN2, RPS6 KB1, RRM1, SARS, SELENBP1, SEMA3C, SEMA4D, SEPP1, SERPINH1, SFN, SFPQ, SFRS7, SHB, SHH, SIAH2, SIVA, SIVA TP53, SKI, SKIL, SLC16A1, SLC1A4, SLC20A1, SMO, esfingomielina fosfodiesterase 1 (SMPD1), SNAI2, SND1, SNRPB2, SOCS1, SOCS3, SOD1, SORT1, SPINT2, SPRY2, SRC, SRPX, STAT1, STAT2, STAT3, STAT5B, STC1, TAF1, TBL3, TBRG4, TCF1, TCF7L2, TFAP2C, TFDP1, TFDP2, TGFA, TGFB1, TGFBI, TGFBR2, TGFBR3, THBS1, TIE, TIMP1, TIMP3, TJP1, TK1, TLE1, TNF, TNFRSF10A, TNFRSF10B, TNFRSF1A, TNFRSF1B, TNFRSF6, TNFSF7, TNK1, TOB1, TP53, TP53BP2, TP5313, TP73, TPBG, TPT1, TRADD, TRAM1, TRRAP, TSG101, TUFM, TXNRD1, TYRO3, UBC, UBE2L6, UCHL1, USP7, VDAC1, VEGF, VHL, VIL2, WEE1, WNT1, WNT2, WNT2B, WNT3, WNT5A, WT1, XRCC1, YES1, YWHAB, YWHAZ, ZAP70, e ZNF9.

[00114] Um vetor rAAV pode compreender como um transgene, um ácido nucleico que codifica uma proteína ou RNA funcional que modula a apoptose. O que se segue é uma lista não limitativa de genes associados à apoptose e ácidos nucleicos que codificam os produtos desses genes e seus homólogos e que codificam pequenos ácidos nucleicos interferentes (por exemplo, shRNAs, miRNAs) que inibem a expressão desses genes e seus homólogos são úteis como transgenes em certas modalidades da invenção: RPS27A, ABL1, AKT1, APAF1, BAD, BAG1, BAG3, BAG4, BAK1, BAX, BCL10, BCL2, BCL2A1, BCL2L1, BCL2L10, BCL2L11, BCL2L12, BCL2L13, BCL2L2, BCLAF1, BFAR, BID, BIK, NAIP, BIRC2, BIRC3, XIAP, BIRC5, BIRC6, BIRC7, BIRC8, BNIP1, BNIP2, BNIP3, BNIP3L, BOK, BRAF, CARD10, CARD11, NLRC4, CARD14, NOD2, NOD1, CARD6, CARDS, CARDS, CASP1, CASP10, CASP14, CASP2, CASP3, CASP4, CASP5, CASP6, CASP7, CASP8, CASP9, CFLAR, CIDEA, CIDEB, CRADD, DAPK1, DAPK2, DFFA,

DFFB, FADD, GADD45A, GDNF, HRK, IGF1R, LTA, LTBR, MCL1, NOL3, PYCARD, RIPK1, RIPK2, TNF, TNFRSF10A, TNFRSF10B, TNFRSF10C, TNFRSF10D, TNFRSF11B, TNFRSF12A, TNFRSF14, TNFRSF19, TNFRSF1A, TNFRSF1B, TNFRSF21, TNFRSF25, CD40, FAS, TNFRSF6B, CD27, TNFRSF9, TNFSF10, TNFSF14, TNFSF18, CD40LG, FASLG, CD70, TNFSF8, TNFSF9, TP53, TP53BP2, TP73, TP63, TRADD, TRAF1, TRAF2, TRAF3, TRAF4 e TRAF5.

[00115] Produtos transgênicos úteis também incluem miRNAs. miRNAs e outros pequenos ácidos nucleicos interferentes regulam a expressão de gene por meio da clivagem/degradação do transcrito do RNA alvo ou repressão translacional do RNA mensageiro alvo (mRNA). Os miRNAs são expressos nativamente, normalmente como produtos finais de RNA não traduzidos de 19-25. Os miRNAs exibem sua atividade por meio de interações específicas de sequência com as regiões 3' não traduzidas (UTR) dos mRNAs alvo. Esses miRNAs expressos endogenamente formam precursores em hairpin que são subsequentemente processados em um duplex de miRNA e, posteriormente, em uma molécula de miRNA de fita simples "madura". Este miRNA maduro guia um complexo multiproteico, miRISC, que identifica o sítio alvo, por exemplo, nas regiões 3' UTR, de mRNAs alvo com base em sua complementaridade com o miRNA maduro.

[00116] O que segue é uma lista não limitativa de genes de miRNA e seus homólogos, são úteis como transgenes ou como alvos para pequenos ácidos nucleicos interferentes codificados por transgenes (por exemplo, esponjas de miRNA, oligonucleotídeos antissentido, RNAs TuD) em certas modalidades dos métodos: hsa-let-7a, hsa-let-7a\*, hsa-let-7b, hsa-let-7b\*, hsa-let-7c, hsa-let-7c\*, hsa-let-7d, hsa-let-7d\*, hsa-let-7e, hsa-let-7e\*, hsa-let-7f, hsa-let-7f-1\*, hsa-let-7f-2\*, hsa-let-7g, hsa-let-7g\*, hsa-let-71, hsa-let-71\*, hsa-miR-1, hsa-miR-100, hsa-miR-100\*, hsa-miR-101, hsa-miR-101\*, hsa-miR-103, hsa-miR-105, hsa-

miR-105\*, hsa-miR-106a, hsa-miR-106a\*, hsa-miR-106b, hsa-miR-106b\*, hsa-miR-107, hsa-miR-10a, hsa-miR-10a\*, hsa-miR-10b, hsa-miR-10b\*, hsa-miR-1178, hsa-miR-1179, hsa-miR-1180, hsa-miR-1181, hsa-miR-1182, hsa-miR-1183, hsa-miR-1184, hsa-miR-1185, hsa-miR-1197, hsa-miR-1200, hsa-miR-1201, hsa-miR-1202, hsa-miR-1203, hsa-miR-1204, hsa-miR-1205, hsa-miR-1206, hsa-miR-1207-3p, hsa-miR-1207-5p, hsa-miR-1208, hsa-miR-122, hsa-miR-122\*, hsa-miR-1224-3p, hsa-miR-1224-5p, hsa-miR-1225-3p, hsa-miR-1225-5p, hsa-miR-1226, hsa-miR-1226\*, hsa-miR-1227, hsa-miR-1228, hsa-miR-1228\*, hsa-miR-1229, hsa-miR-1231, hsa-miR-1233, hsa-miR-1234, hsa-miR-1236, hsa-miR-1237, hsa-miR-1238, hsa-miR-124, hsa-miR-124\*, hsa-miR-1243, hsa-miR-1244, hsa-miR-1245, hsa-miR-1246, hsa-miR-1247, hsa-miR-1248, hsa-miR-1249, hsa-miR-1250, hsa-miR-1251, hsa-miR-1252, hsa-miR-1253, hsa-miR-1254, hsa-miR-1255a, hsa-miR-1255b, hsa-miR-1256, hsa-miR-1257, hsa-miR-1258, hsa-miR-1259, hsa-miR-125a-3p, hsa-miR-125a-5p, hsa-miR-125b, hsa-miR-125b-1\*, hsa-miR-125b-2\*, hsa-miR-126, hsa-miR-126\*, hsa-miR-1260, hsa-miR-1261, hsa-miR-1262, hsa-miR-1263, hsa-miR-1264, hsa-miR-1265, hsa-miR-1266, hsa-miR-1267, hsa-miR-1268, hsa-miR-1269, hsa-miR-1270, hsa-miR-1271, hsa-miR-1272, hsa-miR-1273, hsa-miR-127-3p, hsa-miR-1274a, hsa-miR-1274b, hsa-miR-1275, hsa-miR-127-5p, hsa-miR-1276, hsa-miR-1277, hsa-miR-1278, hsa-miR-1279, hsa-miR-128, hsa-miR-1280, hsa-miR-1281, hsa-miR-1282, hsa-miR-1283, hsa-miR-1284, hsa-miR-1285, hsa-miR-1286, hsa-miR-1287, hsa-miR-1288, hsa-miR-1289, hsa-miR-129\*, hsa-miR-1290, hsa-miR-1291, hsa-miR-1292, hsa-miR-1293, hsa-miR-129-3p, hsa-miR-1294, hsa-miR-1295, hsa-miR-129-5p, hsa-miR-1296, hsa-miR-1297, hsa-miR-1298, hsa-miR-1299, hsa-miR-1300, hsa-miR-1301, hsa-miR-1302, hsa-miR-1303, hsa-miR-1304, hsa-miR-1305, hsa-miR-1306, hsa-miR-1307, hsa-miR-1308, hsa-miR-130a, hsa-miR-130a\*, hsa-miR-



130b, hsa-miR-130b\*, hsa-miR-132, hsa-miR-132\*, hsa-miR-1321, hsa-miR-1322, hsa-miR-1323, hsa-miR-1324, hsa-miR-133a, hsa-miR-133b, hsa-miR-134, hsa-miR-135a, hsa-miR-135a\*, hsa-miR-135b, hsa-miR-135b\*, hsa-miR-136, hsa-miR-136\*, hsa-miR-137, hsa-miR-138, hsa-miR-138-1\*, hsa-miR-138-2\*, hsa-miR-139-3p, hsa-miR-139-5p, hsa-miR-140-3p, hsa-miR-140-5p, hsa-miR-141, hsa-miR-141\*, hsa-miR-142-3p, hsa-miR-142-5p, hsa-miR-143, hsa-miR-143\*, hsa-miR-144, hsa-miR-144\*, hsa-miR-145, hsa-miR-145\*, hsa-miR-146a, hsa-miR-146a\*, hsa-miR-146b-3p, hsa-miR-146b-5p, hsa-miR-147, hsa-miR-147b, hsa-miR-148a, hsa-miR-148a\*, hsa-miR-148b, hsa-miR-148b\*, hsa-miR-149, hsa-miR-149\*, hsa-miR-150, hsa-miR-150\*, hsa-miR-151-3p, hsa-miR-151-5p, hsa-miR-152, hsa-miR-153, hsa-miR-154, hsa-miR-154\*, hsa-miR-155, hsa-miR-155\*, hsa-miR-15a, hsa-miR-15a\*, hsa-miR-15b, hsa-miR-15b\*, hsa-miR-16, hsa-miR-16-1\*, hsa-miR-16-2\*, hsa-miR-17, hsa-miR-17\*, hsa-miR-181a, hsa-miR-181a\*, hsa-miR-181a-2\*, hsa-miR-181b, hsa-miR-181c, hsa-miR-181c\*, hsa-miR-181d, hsa-miR-182, hsa-miR-182\*, hsa-miR-1825, hsa-miR-1826, hsa-miR-1827, hsa-miR-183, hsa-miR-183\*, hsa-miR-184, hsa-miR-185, hsa-miR-185\*, hsa-miR-186, hsa-miR-186\*, hsa-miR-187, hsa-miR-187\*, hsa-miR-188-3p, hsa-miR-188-5p, hsa-miR-18a, hsa-miR-18a\*, hsa-miR-18b, hsa-miR-18b\*, hsa-miR-190, hsa-miR-190b, hsa-miR-191, hsa-miR-191\*, hsa-miR-192, hsa-miR-192\*, hsa-miR-193a-3p, hsa-miR-193a-5p, hsa-miR-193b, hsa-miR-193b\*, hsa-miR-194, hsa-miR-194\*, hsa-miR-195, hsa-miR-195\*, hsa-miR-196a, hsa-miR-196a\*, hsa-miR-196b, hsa-miR-197, hsa-miR-198, hsa-miR-199a-3p, hsa-miR-199a-5p, hsa-miR-199b-5p, hsa-miR-19a, hsa-miR-19a\*, hsa-miR-19b, hsa-miR-19b-1\*, hsa-miR-19b-2\*, hsa-miR-200a, hsa-miR-200a\*, hsa-miR-200b, hsa-miR-200b\*, hsa-miR-200c, hsa-miR-200c\*, hsa-miR-202, hsa-miR-202\*, hsa-miR-203, hsa-miR-204, hsa-miR-205, hsa-miR-206, hsa-miR-208a, hsa-miR-208b, hsa-miR-20a,

hsa-miR-20a\*, hsa-miR-20b, hsa-miR-20b\*, hsa-miR-21, hsa-miR-21\*, hsa-miR-210, hsa-miR-211, hsa-miR-212, hsa-miR-214, hsa-miR-214\*, hsa-miR-215, hsa-miR-216a, hsa-miR-216b, hsa-miR-217, hsa-miR-218, hsa-miR-218-1\*, hsa-miR-218-2\*, hsa-miR-219-1-3p, hsa-miR-219-2-3p, hsa-miR-219-5p, hsa-miR-22, hsa-miR-22\*, hsa-miR-220a, hsa-miR-220b, hsa-miR-220c, hsa-miR-221, hsa-miR-221\*, hsa-miR-222, hsa-miR-222\*, hsa-miR-223, hsa-miR-223\*, hsa-miR-224, hsa-miR-23a, hsa-miR-23a\*, hsa-miR-23b, hsa-miR-23b\*, hsa-miR-24, hsa-miR-24-1\*, hsa-miR-24-2\*, hsa-miR-25, hsa-miR-25\*, hsa-miR-26a, hsa-miR-26a-1\*, hsa-miR-26a-2\*, hsa-miR-26b, hsa-miR-26b\*, hsa-miR-27a, hsa-miR-27a\*, hsa-miR-27b, hsa-miR-27b\*, hsa-miR-28-3p, hsa-miR-28-5p, hsa-miR-296-3p, hsa-miR-296-5p, hsa-miR-297, hsa-miR-298, hsa-miR-299-3p, hsa-miR-299-5p, hsa-miR-29a, hsa-miR-29a\*, hsa-miR-29b, hsa-miR-296-1\*, hsa-miR-296-2\*, hsa-miR-29c, hsa-miR-29c\*, hsa-miR-300, hsa-miR-301a, hsa-miR-301b, hsa-miR-302a, hsa-miR-302a\*, hsa-miR-302b, hsa-miR-302b\*, hsa-miR-302c, hsa-miR-302c\*, hsa-miR-302d, hsa-miR-302d\*, hsa-miR-302e, hsa-miR-302f, hsa-miR-30a, hsa-miR-30a\*, hsa-miR-30b, hsa-miR-30b\*, hsa-miR-30c, hsa-miR-30c-1\*, hsa-miR-30c-2\*, hsa-miR-30d, hsa-miR-30d\*, hsa-miR-30e, hsa-miR-30e\*, hsa-miR-31, hsa-miR-31\*, hsa-miR-32, hsa-miR-32\*, hsa-miR-320a, hsa-miR-320b, hsa-miR-320c, hsa-miR-320d, hsa-miR-323-3p, hsa-miR-323-5p, hsa-miR-324-3p, hsa-miR-324-5p, hsa-miR-325, hsa-miR-326, hsa-miR-328, hsa-miR-329, hsa-miR-330-3p, hsa-miR-330-5p, hsa-miR-331-3p, hsa-miR-331-5p, hsa-miR-335, hsa-miR-335\*, hsa-miR-337-3p, hsa-miR-337-5p, hsa-miR-338-3p, hsa-miR-338-5p, hsa-miR-339-3p, hsa-miR-339-5p, hsa-miR-33a, hsa-miR-33a\*, hsa-miR-33b, hsa-miR-33b\*, hsa-miR-340, hsa-miR-340\*, hsa-miR-342-3p, hsa-miR-342-5p, hsa-miR-345, hsa-miR-346, hsa-miR-34a, hsa-miR-34a\*, hsa-miR-34b, hsa-miR-34b\*, hsa-miR-34c-3p, hsa-miR-34c-5p, hsa-miR-361-3p, hsa-miR-361-5p,

hsa-miR-362-3p, hsa-miR-362-5p, hsa-miR-363, hsa-miR-363\*, hsa-miR-365, hsa-miR-367, hsa-miR-367\*, hsa-miR-369-3p, hsa-miR-369-5p, hsa-miR-370, hsa-miR-371-3p, hsa-miR-371-5p, hsa-miR-372, hsa-miR-373, hsa-miR-373\*, hsa-miR-374a, hsa-miR-374a\*, hsa-miR-374b, hsa-miR-374b\*, hsa-miR-375, hsa-miR-376a, hsa-miR-376a\*, hsa-miR-376b, hsa-miR-376c, hsa-miR-377, hsa-miR-377\*, hsa-miR-378, hsa-miR-378\*, hsa-miR-379, hsa-miR-379\*, hsa-miR-380, hsa-miR-380\*, hsa-miR-381, hsa-miR-382, hsa-miR-383, hsa-miR-384, hsa-miR-409-3p, hsa-miR-409-5p, hsa-miR-410, hsa-miR-411, hsa-miR-411\*, hsa-miR-412, hsa-miR-421, hsa-miR-422a, hsa-miR-423-3p, hsa-miR-423-5p, hsa-miR-424, hsa-miR-424\*, hsa-miR-425, hsa-miR-425\*, hsa-miR-429, hsa-miR-431, hsa-miR-431\*, hsa-miR-432, hsa-miR-432\*, hsa-miR-433, hsa-miR-448, hsa-miR-449a, hsa-miR-449b, hsa-miR-450a, hsa-miR-450b-3p, hsa-miR-450b-5p, hsa-miR-451, hsa-miR-452, hsa-miR-452\*, hsa-miR-453, hsa-miR-454, hsa-miR-454\*, hsa-miR-455-3p, hsa-miR-455-5p, hsa-miR-483-3p, hsa-miR-483-5p, hsa-miR-484, hsa-miR-485-3p, hsa-miR-485-5p, hsa-miR-486-3p, hsa-miR-486-5p, hsa-miR-487a, hsa-miR-487b, hsa-miR-488, hsa-miR-488\*, hsa-miR-489, hsa-miR-490-3p, hsa-miR-490-5p, hsa-miR-491-3p, hsa-miR-491-5p, hsa-miR-492, hsa-miR-493, hsa-miR-493\*, hsa-miR-494, hsa-miR-495, hsa-miR-496, hsa-miR-497, hsa-miR-497\*, hsa-miR-498, hsa-miR-499-3p, hsa-miR-499-5p, hsa-miR-500, hsa-miR-500\*, hsa-miR-501-3p, hsa-miR-501-5p, hsa-miR-502-3p, hsa-miR-502-5p, hsa-miR-503, hsa-miR-504, hsa-miR-505, hsa-miR-505\*, hsa-miR-506, hsa-miR-507, hsa-miR-508-3p, hsa-miR-508-5p, hsa-miR-509-3-5p, hsa-miR-509-3p, hsa-miR-509-5p, hsa-miR-510, hsa-miR-511, hsa-miR-512-3p, hsa-miR-512-5p, hsa-miR-513a-3p, hsa-miR-513a-5p, hsa-miR-513b, hsa-miR-513c, hsa-miR-514, hsa-miR-515-3p, hsa-miR-515-5p, hsa-miR-516a-3p, hsa-miR-516a-5p, hsa-miR-516b, hsa-miR-517\*, hsa-miR-517a, hsa-miR-517b, hsa-miR-517c, hsa-miR-518a-3p, hsa-miR-518a-5p,

hsa-miR-518b, hsa-miR-518c, hsa-miR-518c\*, hsa-miR-518d-3p, hsa-miR-518d-5p, hsa-miR-518e, hsa-miR-518e\*, hsa-miR-518f, hsa-miR-518f\*, hsa-miR-519a, hsa-miR-519b-3p, hsa-miR-519c-3p, hsa-miR-519d, hsa-miR-519e, hsa-miR-519e\*, hsa-miR-520a-3p, hsa-miR-520a-5p, hsa-miR-520b, hsa-miR-520c-3p, hsa-miR-520d-3p, hsa-miR-520d-5p, hsa-miR-520e, hsa-miR-520f, hsa-miR-520g, hsa-miR-520h, hsa-miR-521, hsa-miR-522, hsa-miR-523, hsa-miR-524-3p, hsa-miR-524-5p, hsa-miR-525-3p, hsa-miR-525-5p, hsa-miR-526b, hsa-miR-526b\*, hsa-miR-532-3p, hsa-miR-532-5p, hsa-miR-539, hsa-miR-541, hsa-miR-541\*, hsa-miR-542-3p, hsa-miR-542-5p, hsa-miR-543, hsa-miR-544, hsa-miR-545, hsa-miR-545\*, hsa-miR-548a-3p, hsa-miR-548a-5p, hsa-miR-548b-3p, hsa-miR-548b-5p, hsa-miR-548c-3p, hsa-miR-548c-5p, hsa-miR-548d-3p, hsa-miR-548d-5p, hsa-miR-548e, hsa-miR-548f, hsa-miR-548g, hsa-miR-548h, hsa-miR-548i, hsa-miR-548j, hsa-miR-548k, hsa-miR-548l, hsa-miR-548m, hsa-miR-548n, hsa-miR-548o, hsa-miR-548p, hsa-miR-549, hsa-miR-550, hsa-miR-550\*, hsa-miR-551a, hsa-miR-551b, hsa-miR-551b\*, hsa-miR-552, hsa-miR-553, hsa-miR-554, hsa-miR-555, hsa-miR-556-3p, hsa-miR-556-5p, hsa-miR-557, hsa-miR-558, hsa-miR-559, hsa-miR-561, hsa-miR-562, hsa-miR-563, hsa-miR-564, hsa-miR-566, hsa-miR-567, hsa-miR-568, hsa-miR-569, hsa-miR-570, hsa-miR-571, hsa-miR-572, hsa-miR-573, hsa-miR-574-3p, hsa-miR-574-5p, hsa-miR-575, hsa-miR-576-3p, hsa-miR-576-5p, hsa-miR-577, hsa-miR-578, hsa-miR-579, hsa-miR-580, hsa-miR-581, hsa-miR-582-3p, hsa-miR-582-5p, hsa-miR-583, hsa-miR-584, hsa-miR-585, hsa-miR-586, hsa-miR-587, hsa-miR-588, hsa-miR-589, hsa-miR-589\*, hsa-miR-590-3p, hsa-miR-590-5p, hsa-miR-591, hsa-miR-592, hsa-miR-593, hsa-miR-593\*, hsa-miR-595, hsa-miR-596, hsa-miR-597, hsa-miR-598, hsa-miR-599, hsa-miR-600, hsa-miR-601, hsa-miR-602, hsa-miR-603, hsa-miR-604, hsa-miR-605, hsa-miR-606, hsa-miR-607, hsa-miR-608, hsa-miR-609, hsa-miR-610, hsa-miR-611, hsa-

miR-612, hsa-miR-613, hsa-miR-614, hsa-miR-615-3p, hsa-miR-615-5p, hsa-miR-616, hsa-miR-616\*, hsa-miR-617, hsa-miR-618, hsa-miR-619, hsa-miR-620, hsa-miR-621, hsa-miR-622, hsa-miR-623, hsa-miR-624, hsa-miR-624\*, hsa-miR-625, hsa-miR-625\*, hsa-miR-626, hsa-miR-627, hsa-miR-628-3p, hsa-miR-628-5p, hsa-miR-629, hsa-miR-629\*, hsa-miR-630, hsa-miR-631, hsa-miR-632, hsa-miR-633, hsa-miR-634, hsa-miR-635, hsa-miR-636, hsa-miR-637, hsa-miR-638, hsa-miR-639, hsa-miR-640, hsa-miR-641, hsa-miR-642, hsa-miR-643, hsa-miR-644, hsa-miR-645, hsa-miR-646, hsa-miR-647, hsa-miR-648, hsa-miR-649, hsa-miR-650, hsa-miR-651, hsa-miR-652, hsa-miR-653, hsa-miR-654-3p, hsa-miR-654-5p, hsa-miR-655, hsa-miR-656, hsa-miR-657, hsa-miR-658, hsa-miR-659, hsa-miR-660, hsa-miR-661, hsa-miR-662, hsa-miR-663, hsa-miR-663b, hsa-miR-664, hsa-miR-664\*, hsa-miR-665, hsa-miR-668, hsa-miR-671-3p, hsa-miR-671-5p, hsa-miR-675, hsa-miR-7, hsa-miR-708, hsa-miR-708\*, hsa-miR-7-1\*, hsa-miR-7-2\*, hsa-miR-720, hsa-miR-744, hsa-miR-744\*, hsa-miR-758, hsa-miR-760, hsa-miR-765, hsa-miR-766, hsa-miR-767-3p, hsa-miR-767-5p, hsa-miR-768-3p, hsa-miR-768-5p, hsa-miR-769-3p, hsa-miR-769-5p, hsa-miR-770-5p, hsa-miR-802, hsa-miR-873, hsa-miR-874, hsa-miR-875-3p, hsa-miR-875-5p, hsa-miR-876-3p, hsa-miR-876-5p, hsa-miR-877, hsa-miR-877\*, hsa-miR-885-3p, hsa-miR-885-5p, hsa-miR-886-3p, hsa-miR-886-5p, hsa-miR-887, hsa-miR-888, hsa-miR-888\*, hsa-miR-889, hsa-miR-890, hsa-miR-891a, hsa-miR-891b, hsa-miR-892a, hsa-miR-892b, hsa-miR-9, hsa-miR-9\*, hsa-miR-920, hsa-miR-921, hsa-miR-922, hsa-miR-923, hsa-miR-924, hsa-miR-92a, hsa-miR-92a-1\*, hsa-miR-92a-2\*, hsa-miR-92b, hsa-miR-92b\*, hsa-miR-93, hsa-miR-93\*, hsa-miR-933, hsa-miR-934, hsa-miR-935, hsa-miR-936, hsa-miR-937, hsa-miR-938, hsa-miR-939, hsa-miR-940, hsa-miR-941, hsa-miR-942, hsa-miR-943, hsa-miR-944, hsa-miR-95, hsa-miR-96, hsa-miR-96\*, hsa-miR-98, hsa-miR-99a, hsa-miR-99a\*, hsa-miR-99b, e hsa-miR-

99b\*. Por exemplo, miRNA direcionado à estrutura de leitura aberta 72 do cromossomo 8 (C9orf72) que expressa superóxido dismutase (SOD1), associado com esclerose lateral amiotrófica (ALS) pode ser de interesse.

[00117] Um miRNA inibe a função dos mRNAs que ele direciona e, como resultado, inibe a expressão dos polipeptídeos codificados pelos mRNAs. Assim, o bloqueio (parcial ou total) da atividade do miRNA (por exemplo, silenciamento do miRNA) pode efetivamente induzir, ou restaurar, a expressão de um polipeptídeo cuja expressão é inibida (desreprimir o polipeptídeo). Em uma modalidade, a desrepressão de polipeptídeos codificados por alvos de mRNA de um miRNA é realizada inibindo a atividade de miRNA em células através de qualquer um de uma variedade de métodos. Por exemplo, o bloqueio da atividade de um miRNA pode ser realizado por hibridação com um pequeno ácido nucleico de interferência (por exemplo, oligonucleotídeo antissenso, esponja de miRNA, RNA de TuD) que é complementar ou substancialmente complementar ao miRNA, bloqueando assim a interação do miRNA com seu mRNA alvo. Conforme usado neste documento, um pequeno ácido nucleico interferente que é substancialmente complementar a um miRNA é aquele que é capaz de hibridar com um miRNA e bloquear a atividade do miRNA. Em algumas modalidades, um pequeno ácido nucleico interferente que é substancialmente complementar a um miRNA é um pequeno ácido nucleico interferente que é complementar ao miRNA em tudo, exceto 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 ou 18 bases. Um “inibidor de miRNA” é um agente que bloqueia a função, expressão e/ou processamento de miRNA. Por exemplo, essas moléculas incluem, mas não estão limitadas a antissenso específico de microRNA, esponjas de microRNA, RNAs decoy resistentes (RNAs TuD) e oligonucleotídeos de microRNA (fita dupla, grampo de cabelo, oligonucleotídeos curtos) que inibem a interação de miRNA com um complexo

Drosha.

[00118] Ainda outros transgenes úteis podem incluir aqueles que codificam imunoglobulinas que conferem imunidade passiva a um patógeno. Uma "molécula de imunoglobulina" é uma proteína que contém as porções imunologicamente ativas de uma cadeia pesada de imunoglobulina e uma cadeia leve de imunoglobulina acoplada covalentemente e capaz de se combinar especificamente com o antígeno. As moléculas de imunoglobulina são de qualquer tipo (*por exemplo*, classe IgG, IgE, IgM, IgD, IgA e IgY), (*por exemplo*, IgG1, IgG2, IgG3, IgG4, IgA1 e IgA2) ou subclasses. Os termos "anticorpo" e "imunoglobulina" podem ser usados de forma intercambiável neste documento.

[00119] Uma "cadeia pesada de imunoglobulina" é um polipeptídeo que contém pelo menos uma porção do domínio de ligação ao antígeno de uma imunoglobulina e pelo menos uma porção de uma região variável de uma cadeia pesada de imunoglobulina ou pelo menos uma porção de uma região constante de uma cadeia pesada de imunoglobulina. Assim, a cadeia pesada derivada de imunoglobulina tem regiões significativas de homologia de sequência de aminoácidos com um membro da superfamília de genes de imunoglobulina. Por exemplo, a cadeia pesada em um fragmento Fab é uma cadeia pesada derivada de imunoglobulina.

[00120] Uma "cadeia leve de imunoglobulina" é um polipeptídeo que contém pelo menos uma porção do domínio de ligação ao antígeno de uma imunoglobulina e pelo menos uma porção da região variável ou pelo menos uma porção de uma região constante de uma cadeia leve de imunoglobulina. Assim, a cadeia leve derivada de imunoglobulina tem regiões significativas de homologia de aminoácidos com um membro da superfamília de genes de imunoglobulina.

[00121] Uma "imunoadesina" é uma molécula quimérica semelhante a um anticorpo que combina o domínio funcional de uma proteína de

ligação, geralmente um receptor, ligando ou molécula de adesão celular, com domínios constantes de imunoglobulina, geralmente incluindo as regiões de dobradiça e Fc.

[00122] Um "fragmento de ligação ao antígeno" (fragmento Fab)" é uma região em um anticorpo que se liga a antígenos. É composta por um domínio constante e um domínio variável de cada uma das cadeias pesada e leve.

[00123] O construto antipatógeno é selecionado com base no agente causador (patógeno) para a doença contra a qual se busca proteção. Esses patógenos podem ser de origem viral, bacteriana ou fúngica e podem ser usados para prevenir a infecção em humanos contra doenças humanas, ou em mamíferos não humanos ou outros animais para prevenir doenças veterinárias.

[00124] O rAAV pode incluir genes que codificam anticorpos e, particularmente, anticorpos neutralizantes contra um patógeno viral. Esses anticorpos antivirais podem incluir anticorpos anti-influenza dirigidos contra um ou mais de Influenza A, Influenza B e Influenza C. Os vírus do tipo A são os patógenos humanos mais virulentos. Os sorotipos de influenza A que têm sido associados a pandemias incluem o H1N1, que causou a gripe espanhola em 1918, e a gripe suína em 2009; H2N2, que causou a gripe asiática em 1957; H3N2, que causou a gripe de Hong Kong em 1968; H5N1, que causou a gripe aviária em 2004; H7N7; H1N2; H9N2; H7N2; H7N3; e H10N7. Outros vírus patogênicos alvo incluem arenavírus (incluindo funin, machupo e Lassa), filovírus (incluindo Marburg e Ebola), hantavírus, picornoviridae (incluindo rinovírus, ecovírus), coronavírus, paramixovírus, morbilivírus, vírus sincicial respiratório, togavírus, cox vírus, parvovírus B19, parainfluenza, adenovírus, reovírus, varíola (*Variola major* (Varíola)) e *Vaccinia* (Cowpox) da família dos poxvírus e varicela-zoster (pseudorabies). As febres hemorrágicas virais são causadas por membros da família dos arenavírus (febre de



Lassa) (cuja família também está associada à coriomeningite linfocítica (LCM)), filovírus (vírus do ebola) e hantavírus (puremala). Os membros do picornavírus (uma subfamília de rinovírus) estão associados ao resfriado comum em humanos. A família do coronavírus, que inclui uma série de vírus não humanos, como o vírus da bronquite infecciosa (aves), vírus gastroentérico transmissível porcino (porco), vírus da encefalomielite hemaglutinatina suína (porco), vírus da peritonite infecciosa felina (gato), coronavírus entérico felino (gato), coronavírus canino (cão). Os coronavírus respiratórios humanos foram supostamente associados ao resfriado comum, hepatite não A, B ou C e síndrome respiratória aguda súbita (SARS). A família do paramixovírus inclui vírus parainfluenza tipo 1, vírus parainfluenza tipo 3, vírus parainfluenza bovino tipo 3, rubulavírus (vírus da caxumba, vírus parainfluenza tipo 2, vírus parainfluenza tipo 4, vírus da doença de Newcastle (galinhas), peste bovina, morbilivírus, que inclui sarampo e cinomose canina e pneumovírus, que inclui o vírus sincicial respiratório (RSV). A família dos parvovírus inclui o parvovírus felino (enterite felina), o vírus da panleucopenia felina, o parvovírus canino e o parvovírus porcino. A família de adenovírus inclui vírus (EX, AD7, ARD, OB) que causam doenças respiratórias. Assim, em certas modalidades, um vetor rAAV como aqui descrito pode ser projetado para expressar um anticorpo anti-ebola, por exemplo, 2G4, 4G7, 13C6, um anticorpo anti-influenza, por exemplo, FI6, CR8033 e anticorpo anti-RSV, por exemplo, palivizumabe, motavizumabe. Um construto de anticorpo neutralizante contra um patógeno bacteriano também pode ser selecionado para uso na presente invenção. Em uma modalidade, o construto de anticorpo neutralizante é direcionado contra a própria bactéria. Em outra modalidade, o construto de anticorpo neutralizante é direcionado contra uma toxina produzida pela bactéria. Exemplos de patógenos bacterianos transportados pelo ar incluem, *por exemplo*, *Neisseria meningitidis* (meningite), *Klebsiella pneumonia*

(pneumonia), *Pseudomonas aeruginosa* (pneumonia), *Pseudomonas pseudomallei* (pneumonia), *Pseudomonas mallei* (pneumonia), *Acinetobacter* (pneumonia), *Moraxella catarrhalis*, *Moraxella lacunata*, *Alkaligenes*, *Cardiobacterium*, *Haemophilus influenzae* (gripe), *Haemophilusparainfluenzae*, *Bordetella pertussis* (tosse convulsa), *Francisellatularensis*(pneumonia/febre), *Legionella pneumonia* (Doença dos legionários), *Chlamydia psittaci*(pneumonia), *Chlamydia pneumoniae* (pneumonia), *Mycobacterium tuberculosis* (tuberculose (TB)), *Mycobacterium kansasii* (TB), *Mycobacterium avium* (pneumonia), *Nocardia asteroides*(pneumonia), *Bacillus anthracis* (antraz), *Staphylococcus aureus* (pneumonia), *Streptococcus pyogenes* (escarlatina), *Streptococcus pneumoniae* (pneumonia), *Corynebacteria diphtheria* (difteria), *Mycoplasma pneumoniae* (pneumonia).

[00125] O rAAV pode incluir genes que codificam anticorpos e, particularmente, anticorpos neutralizantes contra um patógeno bacteriano, tal como o agente causador do antraz, uma toxina produzida por *Bacillus anthracis*. Foram descritos anticorpos neutralizantes contra o agente protetor (PA), um dos três peptídeos que formam o toxoide. Os outros dois polipeptídeos consistem em fator letal (LF) e fator de edema (FE). Os anticorpos neutralizantes anti-PA foram descritos como eficazes na imunização passiva contra o antraz. Ver, por exemplo, Patente U.S. 7.442.373; R. Sawada-Hirai et al, J Immune Based Ther Vaccines. 2004; 2: 5. (on-line 2004 May 12). Ainda outros anticorpos neutralizantes da toxina anti-antraz foram descritos e/ou podem ser gerados. Da mesma forma, anticorpos neutralizantes contra outras bactérias e/ou toxinas bacterianas podem ser usados para gerar um construto antipatógeno distribuído por AAV como aqui descrito.

[00126] Os anticorpos contra doenças infecciosas podem ser causados por parasitas ou por fungos, incluindo, por exemplo, espécies *Aspergillus*, *Absidiacorymbifera*, *Rhizopusstolonifer*, *Mucor plumbeus*,

*Cryptococcus neoformans*, *Histoplasma capsulatum*, *Blastomyces dermatitidis*, *Coccidioides immitis*, espécies *Penicillium*, *Micropolysporafeni*, *Thermoactinomyces vulgaris*, *Alternaria alternate*, espécies *Cladosporium*, *Helminthosporium*, e espécies *Stachybotrys*.

[00127] O rAAV pode incluir genes que codificam anticorpos e, particularmente, anticorpos neutralizantes, contra fatores patogênicos de doenças como a doença de Alzheimer (AD), doença de Parkinson (PD), doença de Parkinson associada à GBA (GBA - PD), artrite reumatoide (RA), Síndrome do intestino irritável (IBS), doença pulmonar obstrutiva crônica (COPD), cânceres, tumores, esclerose sistêmica, asma e outras doenças. Tais anticorpos podem ser, sem limitação, por exemplo, alfa-sinucleína, fator de crescimento endotelial antivascolar (VEGF) (anti-VEGF), anti-VEGFA, anti-PD-1, anti-PDL1, anti-CTLA-4, anti-TNF-alpha, anti-IL-17, anti-IL-23, anti-IL-21, anti-IL-6, anti-IL-6 receptor, anti-IL-5, anti-IL-7, anti-Factor XII, anti-IL-2, anti-HIV, anti-IgE, receptor-1 do fator de necrose antitumoral (TNFR1), anti-notch 2/3, anti-notch 1, anti-OX40, receptor anti-erb-b2 tirosina quinase 3 (ErbB3), anti-ErbB2, antígeno de maturação de células anti-beta, estimulador de linfócitos anti-B, anti-CD20, anti-HER2, fator estimulador de colônia de macrófagos antigranulócitos, antioncostatina M (OSM), proteína do gene 3 de ativação de antilinfócitos (LAG3), anti-CCL20, componente antiamilóide P sérico (SAP), inibidor anti-prolil hidroxilase, anti-CD38, antiglicoproteína IIb/IIIa, anti-CD52, anti-CD30, anti-IL-1beta, receptor do fator de crescimento anti-epidérmico, anti-CD25, ligando anti-RANK, proteína C5 do sistema anticomplemento, anti-CD11a, receptor anti-CD3, integrina anti-alfa-4 ( $\alpha 4$ ), proteína F anti-RSV e anti-integrina  $\alpha 4\beta 7$ . Ainda outros patógenos e doenças serão evidentes para um versado na técnica. Outros anticorpos adequados podem incluir aqueles úteis para o tratamento da doença de Alzheimer, tais como, por exemplo, anti-beta-amilóide (por exemplo, crenezumabe, solanezumabe, aducanumabe), fibrila anti-

beta-amiloide, placas anti-beta-amiloide, anti-tau, um bapineuzamabe, entre outros. Outros anticorpos adequados para o tratamento de uma variedade de indicações incluem aqueles descritos, por exemplo, em PCT/US2016/058968, depositado em 27 de outubro de 2016, publicado como WO 2017/075119A1.

## II. Produção de vetor rAAV

[00128] Para uso na produção de um vetor viral AAV (*por exemplo*, um AAV recombinante (r)), os cassetes de expressão podem ser transportados em qualquer vetor adequado, *por exemplo*, um plasmídeo, que é distribuído a uma célula hospedeira de empacotamento. Os plasmídeos úteis nesta invenção podem ser engenheirados de modo que sejam adequados para replicação e empacotamento *in vitro* em células procarióticas, células de inseto, células de mamífero, entre outras. As técnicas de transfecção adequadas e as células hospedeiras de empacotamento são conhecidas e/ou podem ser prontamente concebidas por um versado na técnica.

[00129] Os métodos para gerar e isolar AAVs adequados para uso como vetores são conhecidos na técnica. *Ver em geral, por exemplo*, Grieger&Samulski, 2005, "Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications," *Adv. Biochem. Engin/Biotechnol.* 99: 119-145; Buninget *al.*, 2008, "Recent developments in adeno-associated virus vector technology," *J. Gene Med.* 10:717-733; e as referências citadas abaixo, cada uma das quais é aqui incorporada por referência na sua totalidade. Para empacotar um transgene em vírions, as ITRs são as únicas componentes de AAV necessárias em *cis* no mesmo construto que a molécula de ácido nucleico que contém os cassetes de expressão. Os genes cap e rep podem ser fornecidos *in trans*.

[00130] Em uma modalidade, os cassetes de expressão descritos neste documento são engenheirados em um elemento genético (*por*

*exemplo*, um plasmídeo de transporte) que transfere as sequências de construto transgênicas transportadas nele para uma célula hospedeira de empacotamento para a produção de um vetor viral. Em uma modalidade, o elemento genético selecionado pode ser distribuído a uma célula de empacotamento de AAV por qualquer método adequado, incluindo transfecção, eletroporação, distribuição de lipossomos, técnicas de fusão de membranas, peletes revestidos com DNA de alta velocidade, infecção viral e fusão de protoplastos. Células de empacotamento de AAV estáveis também podem ser feitas. Alternativamente, os cassetes de expressão podem ser usados para gerar um vetor viral diferente de AAV. Os métodos utilizados para fazer tais construtos são conhecidos daqueles versados na manipulação de ácido nucleico e incluem engenharia genética, engenharia recombinante e técnicas sintéticas. *Ver, por exemplo*, Molecular Cloning: A Laboratory Manual, ed. Green and Sambrook, Cold Spring Harbor Press, Cold Spring Harbor, NY (2012).

[00131] O termo "intermediário de AAV" ou "intermediário de vetor de AAV" refere-se a um capsídeo de rAAV montado que não possui as sequências genômicas desejadas empacotadas nele. Eles também podem ser chamados de capsídeo "vazio". Tal capsídeo pode conter nenhuma sequência genômica detectável de um cassete de expressão, ou apenas sequências genômicas parcialmente empacotadas que são insuficientes para atingir a expressão do produto de gene. Esses capsídeos vazios não são funcionais para transferir o gene de interesse para uma célula hospedeira.

[00132] O vírus adenoassociado recombinante (AAV) aqui descrito pode ser gerado usando técnicas conhecidas. *Ver, por exemplo*, WO 2003/042397; WO 2005/033321, WO 2006/110689; US 7588772 B2. Tal método envolve a cultura de uma célula hospedeira que contém uma sequência de ácido nucleico que codifica uma proteína do capsídeo de AAV; um gene rep funcional; um cassete de expressão composto por,

no mínimo, repetições terminais invertidas (ITRs) de AAV e um transgene; e funções auxiliares suficientes para permitir o empacotamento do cassete de expressão na proteína do capsídeo de AAV. Métodos de geração do capsídeo, sequências de codificação, portanto, e métodos para produção de vetores virais de rAAV foram descritos. *Ver, por exemplo, Gao, et al, Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003) e US 2013/0045186A1.*

[00133] Em uma modalidade, é fornecida uma cultura de células de produção útil para a produção de um AAV recombinante. Tal cultura de células contém um ácido nucleico que expressa a proteína do capsídeo de AAV na célula hospedeira; uma molécula de ácido nucleico adequada para empacotamento no capsídeo de AAV, por exemplo, um genoma de vetor que contém ITRs de AAV e uma sequência de ácido nucleico não AAV que codifica um produto de gene operacionalmente ligado a sequências que dirigem a expressão do produto em uma célula hospedeira; e funções rep de AAV e funções auxiliares de adenovírus suficientes para permitir o empacotamento da molécula de ácido nucleico no capsídeo de AAV recombinante. Em uma modalidade, a cultura de células é composta por células de mamíferos (por exemplo, células 293 de rim embrionário humano, entre outras) ou células de inseto (por exemplo, baculovírus).

[00134] Opcionalmente, as funções rep são fornecidas por um AAV diferente do AAV que fornece o capsídeo. Por exemplo, o rep pode ser, mas não está limitado a, proteína rep AAV1, proteína rep AAV2, proteína rep AAV3, proteína rep AAV4, proteína rep AAV5, proteína rep AAV6, proteína rep AAV7, proteína rep AAV8; ou rep 78, rep 68, rep 52, rep 40, rep68/78 e rep40/52; ou um fragmento das mesmas; ou outra fonte. Opcionalmente, as sequências rep e cap estão no mesmo elemento genético na cultura de células. Pode haver um espaçador entre a sequência rep e o gene cap. Qualquer uma dessas sequências de capsídeo de

AAV ou AAV mutante pode estar sob o controle de sequências de controle regulatórias exógenas que direcionam sua expressão em uma célula hospedeira.

[00135] Em uma modalidade, as células são fabricadas em uma cultura de células adequadas (por exemplo, células HEK 293). Os métodos para fabricar os vetores de terapia genética aqui descritos incluem métodos bem conhecidos na técnica, tais como geração de DNA de plasmídeo utilizado para a produção dos vetores de terapia genética, geração vetorial e purificação vetorial. Em algumas modalidades, o vetor de terapia genética é um vetor AAV e os plasmídeos gerados são um plasmídeo cis de AAV que codifica o genoma do vetor incluindo o gene de interesse, um plasmídeo trans de AAV contendo genes rep e cap de AAV e um plasmídeo auxiliar de adenovírus. O processo de geração vetorial pode incluir etapas do método, como iniciação de cultura celular, passagem de células, semeadura de células, transfecção de células com o DNA de plasmídeo, permuta de meio pós-transfecção para meio livre de soro e colheita de células contendo vetor e meio cultural. As células e os meios de cultura contendo vetor colhidos são aqui referidos como colheita de células brutas. Em ainda outro sistema, os vetores de terapia genética são introduzidos em células de inseto por infecção com vetores baseados em baculovírus. Para revisões sobre estes sistemas de produção, ver geralmente, por exemplo, Zhang et al., 2009, "Adenovirus-adenovirus hybrid for large-scale recombinant adenovirus production," *Human Gene Therapy* 20:922-929, o conteúdo de cada um dos quais é aqui incorporado por referência na sua totalidade. Métodos de produção e utilização destes e outros sistemas de produção de AAV são também descritos nas seguintes Patentes U.S., o conteúdo de cada uma das quais é aqui incorporado por referência na sua totalidade: 5.139.941; 5.741.683; 6.057.152; 6.204.059; 6.268.213; 6.491.907; 6.660.514; 6.951.753; 7.094.604; 7.172.893;

7.201.898; 7.229.823; e 7.439.065.

[00136] A colheita de células brutas pode depois ser submetida a etapas do método, tais como concentração da colheita do vetor, diafiltração da colheita do vetor, microfluidização da colheita do vetor, digestão por nuclease da colheita do vetor, filtração do intermediário microfluidizado, purificação bruta por cromatografia, purificação bruta por ultracentrifugação, permuta de tampão por filtração de fluxo tangencial e/ou formulação e filtração para preparar o vetor em massa.

[00137] Uma purificação por cromatografia de afinidade em duas etapas em alta concentração de sal seguida por cromatografia em resina de troca aniônica é usada para purificar o produto de droga do vetor e remover os capsídeos vazios. Esses métodos são descritos em mais detalhes na Publicação de Patente Internacional WO 2017/160360, que é aqui incorporada por referência. Purification methods for AAV8 Publicação WO2017/100676 e rh10 Publicação de Patente Internacional WO 2017/100704 e para AAV1 Publicação de Patente Internacional WO 2017/100674, são todas incorporadas aqui por referência.

[00138] Para calcular o teor de partículas vazias e cheias, os volumes da banda VP3 para uma amostra selecionada (*por exemplo*, em exemplos aqui, uma preparação purificada por gradiente de iodixanol onde # de GC = # de partículas) é representada graficamente contra partículas de GC carregadas. A equação linear resultante ( $y = mx+c$ ) é usada para calcular o número de partículas nos volumes de banda dos picos do artigo de teste. O número de partículas (pt) por 20  $\mu$ L carregado é então multiplicado por 50 para dar partículas (pt)/mL. Pt/mL dividido por GC/mL fornece a razão de partículas para cópias do genoma (pt/GC). Pt/mL – GC/mL fornece pt/mL vazio. Pt/mL vazio dividido por pt/mL e x 100 fornece a porcentagem de partículas vazias.

[00139] Geralmente, os métodos para ensaiar capsídeos vazios e partículas de vetor AAV com genomas empacotados são conhecidos na



técnica. Ver, por exemplo, Grimm et al., *Gene Therapy* (1999) 6:1322-1330; Sommer et al., *Molec. Ther.* (2003) 7:122-128. Para testar o capsídeo desnaturado, os métodos incluem submeter o estoque de AAV tratado a eletroforese em gel de poliacrilamida-SDS, consistindo em qualquer gel capaz de separar as três proteínas do capsídeo, por exemplo, um gradiente de gel contendo Tris-acetato a 3-8% no tampão, depois levar o gel até o material da amostra ser separado, e transferir o gel para as membranas de nylon ou nitrocelulose, de preferência náilon. Os anticorpos anticapsídeo AAV são então utilizados como anticorpos primários que se ligam a proteínas de capsídeo desnaturadas, de preferência um anticorpo monoclonal anticapsídeo AAV, mais preferencialmente o anticorpo monoclonal anti-AAV-2 B1 (Wobus et al., *J. Virol.* (2000) 74:9281-9293). Um anticorpo secundário é, então, usado, um que se liga ao anticorpo primário e contém um meio para detectar ligação com o anticorpo primário, mais preferencialmente, um anticorpo anti-IgG contendo uma molécula de detecção ligada covalentemente a este, mais preferencialmente um anticorpo de ovelha anti-IgG de camundongo ligado covalentemente a peroxidase de rábano. Um método para detectar a ligação é utilizado para determinar semiquantitativamente a ligação entre os anticorpos primário e secundário, de preferência um método de detecção capaz de detectar emissões de isótopos radioativos, radiação eletromagnética ou alterações colorimétricas, mais preferencialmente um kit de detecção de quimiluminescência. Por exemplo, para SDS-PAGE, amostras das frações da coluna podem ser tomadas e aquecidas em tampão de carga de SDS-PAGE contendo agente de redução (*por exemplo*, DTT) e proteínas do capsídeo foram resolvidas em géis de poliacrilamida pré-moldados (*por exemplo*, Novex). A coloração com prata pode ser realizada usando o SilverXpress (Invitrogen, CA) de acordo com as instruções do fabricante ou outro método de coloração adequado, por exemplo, manchas SYPRO de rubi ou

coomassie. Numa modalidade, a concentração de genomas vetoriais de AAV (vg) em frações da coluna pode ser medida por PCR quantitativa em tempo real (Q-PCR). As amostras são diluídas e digeridas com DNase I (ou outra nuclease adequada) para remover o DNA exógeno. Após a inativação da nuclease, as amostras são posteriormente diluídas e amplificadas utilizando iniciadores e uma sonda fluorogênica TaqMan™ específica para a sequência de DNA entre os iniciadores. O número de ciclos necessários para atingir um nível definido de fluorescência (ciclo de limiar, Ct) é medido para cada amostra num Sistema de Detecção de Sequências Applied Biosystems Prism 7700. DNA plasmídico contendo sequências idênticas às contidas no vetor AAV é empregado para gerar uma curva-padrão na reação Q-PCR. Os valores do limiar do ciclo (Ct) obtidos a partir das amostras são utilizados para determinar o título do genoma vetorial normalizando-o ao valor Ct da curva padrão do plasmídeo. Ensaio de ponto final baseados na PCR digital também podem ser usados.

[00140] Num aspecto, é utilizado um método de q-PCR otimizado que utiliza uma serina protease de largo espectro, por exemplo, proteinase K (tal como está comercialmente disponível na Qiagen). Mais particularmente, o ensaio de título de genoma qPCR otimizado é semelhante a um ensaio padrão, exceto que após a digestão com DNase I, as amostras são diluídas com tampão de proteinase K e tratadas com proteinase K seguida de inativação por calor. Adequadamente, as amostras são diluídas com tampão de proteinase K numa quantidade igual ao tamanho da amostra. O tampão de proteinase K pode ser concentrado até 2 vezes ou mais. Normalmente, o tratamento com proteinase K é de cerca de 0,2 mg/mL, mas pode variar de 0,1 mg/mL a cerca de 1 mg/mL. A etapa de tratamento é geralmente conduzida a cerca de 55°C por cerca de 15 minutos, mas pode ser realizada a uma temperatura mais baixa (por exemplo, cerca de 37°C a cerca de 50°C) por um

período de tempo mais longo (por exemplo, cerca de 20 minutos a cerca de 30 minutos), ou uma temperatura mais alta (por exemplo, até cerca de 60°C) por um período de tempo mais curto (por exemplo, cerca de 5 a 10 minutos). Da mesma forma, a inativação do calor geralmente, a cerca de 95°C durante cerca de 15 minutos, mas a temperatura pode ser diminuída (por exemplo, cerca de 70 a cerca de 90°C) e o tempo prolongado (por exemplo, cerca de 20 minutos a cerca de 30 minutos). As amostras são então diluídas (por exemplo, 1000 vezes) e submetidas à análise TaqMan como descrito no ensaio padrão.

[00141] Além disso, ou alternativamente, a PCR digital por gota (ddPCR) pode ser usada. Por exemplo, foram descritos métodos para determinar títulos do genoma de vetores AAV de fita simples e autocomplementares por ddPCR. Ver, por exemplo, M. Lock et al, Hum Gene Ther Methods. 2014 Apr;25(2):115-25. doi: 10.1089/hgtb.2013.131. Epub 2014 Feb 14.

[00142] Em resumo, o método para separar partículas de rAAV com sequências genômicas embaladas de intermediários de AAV deficientes em genoma envolve submeter uma suspensão compreendendo partículas virais AAV recombinantes e intermediários de capsídeo AAV à cromatografia líquida de rápido desempenho, em que as partículas virais de AAV e os intermediários de AAV estão ligados a uma resina de permuta aniônica forte equilibrada a um pH alto e submetidos a um gradiente de sal enquanto monitoram o eluato quanto à absorvância ultravioleta em cerca de 260 e cerca de 280. O pH pode ser ajustado dependendo do AAV selecionado. Ver, por exemplo, WO2017/160360 (AAV9), WO2017/100704 (AAVrh10), WO 2017/100676 (por exemplo, AAV8) e WO 2017/100674 (AAV1)] que são aqui incorporados por referência. Neste método, os capsídeos completos de AAV são coletados a partir de uma fração que é eluída quando a razão de A260/A280 atinge um ponto de inflexão. Em um exemplo, para a etapa de Cromatografia de

Afinidade, o produto diafiltrado pode ser aplicado a uma resina de afinidade Capture Select™ Poros- AAV2/9 (Life Technologies) que captura com eficiência o sorotipo AAV2. Sob essas condições iônicas, uma porcentagem significativa de DNA e proteínas celulares residuais flui através da coluna, enquanto as partículas do AAV são capturadas com eficiência.

### III. Composições e Usos

[00143] São fornecidas aqui composições contendo pelo menos um estoque de rAAV (por exemplo, um estoque de rAAV ou um estoque de rAAV mutante) e um transportador, excipiente e/ou conservante opcional. Um estoque de rAAV refere-se a uma pluralidade de vetores de rAAV que são iguais, por exemplo, tais como nas quantidades descritas abaixo na discussão de concentrações e unidades de dosagem.

[00144] Como usado neste documento, "transportador" inclui todo e qualquer solvente, meio de dispersão, veículos, revestimentos, diluentes, agentes antibacterianos e antifúngicos, agentes isotônicos e de retardamento da absorção, tampões, soluções transportadoras, suspensões, coloides e semelhantes. A utilização de tais meios e agentes para substâncias farmacologicamente ativas é bem conhecida na técnica. Ingredientes ativos suplementares também podem ser incorporados nas composições. A frase "farmacologicamente aceitável" refere-se a entidades e composições moleculares que não produzem uma reação indesejável alérgica ou semelhante quando administradas a um hospedeiro. Os veículos de distribuição, tal como lipossomas, nanocápsulas, micropartículas, microesferas, partículas lipídicas, vesículas e semelhantes, podem ser utilizados para a introdução das composições da presente invenção em células hospedeiras adequadas. Em particular, os transgenes distribuídos por vetor rAAV pode ser formulado para distribuição ou encapsulado em uma partícula lipídica, um lipossoma, uma vesícula, uma nanoesfera ou uma nanopartícula ou semelhante.

[00145] Em uma modalidade, uma composição inclui uma formulação final adequada para distribuição a um sujeito, por exemplo, é uma suspensão líquida aquosa tamponada a um pH fisiologicamente compatível e concentração de sal. Opcionalmente, um ou mais tensoativos estão presentes na formulação. Em outra modalidade, a composição pode ser transportada como um concentrado que é diluído para administração a um sujeito. Em outras modalidades, a composição pode ser liofilizada e reconstituída no momento da administração.

[00146] Um tensoativo adequado, ou combinação de tensoativos, pode ser selecionado de tensoativos não iônicos que não sejam tóxicos. Em uma modalidade, um tensoativo copolímero de bloco difuncional que termina em grupos hidroxil primários é selecionado, por exemplo, como Pluronic® F68 [BASF], também conhecido como Poloxamer 188, que tem um pH neutro, tem um peso molecular médio de 8400. Outros tensoativos e outros poloxâmeros podem ser selecionados, ou seja, copolímeros triblocos não iônicos compostos de uma cadeia hidrofóbica central de polioxipropileno (poli(óxido de propileno)) flanqueada por duas cadeias hidrofílicas de polioxietileno (poli(óxido de etileno)), SOLUTOL HS 15 (Macrogol-15 Hidroxiestearato), LABRASOL (Glicerídeo polioxi caprílico), éter oleílico de polioxi 10, TWEEN (ésteres de ácidos graxos de polioxietileno sorbitano), etanol e polietileno glicol. Numa modalidade, a formulação contém um poloxâmero. Estes copolímeros são comumente nomeados com a letra "P" (para poloxâmero) seguidos de três dígitos: os dois primeiros dígitos x 100 dão a massa molecular aproximada do núcleo de polioxipropileno, e o último dígito x 10 indica o percentual de polioxietileno. Em uma modalidade, o Poloxamer 188 é selecionado. O tensoativo pode estar presente numa quantidade até cerca de 0,0005% a cerca de 0,001% da suspensão.

[00147] Os vetores são administrados em quantidades suficientes para transfectar as células e fornecer níveis suficientes de transferência

e expressão de gene para fornecer um benefício terapêutico sem efeitos adversos indevidos, ou com efeitos fisiológicos medicamente aceitáveis, que podem ser determinados por aqueles versados nas técnicas médicas. As vias convencionais e farmacologicamente aceitáveis de administração incluem, mas não estão limitadas a, distribuição direta a um órgão desejado (*por exemplo*, o fígado (opcionalmente através da artéria hepática), pulmão, coração, olho, rim,), oral, inalação, intranasal, administração intratecal, intratraqueal, intra-arterial, direta ao olho (opcionalmente via distribuição ocular, injeção subretiniana, injeção intrarretinal, intravítrea, tópica), intravenosa, intramuscular, subcutânea, intradérmica e outras vias de administração parentais. Em uma modalidade, a via de administração é a injeção subretinal ou intravítrea. As vias de administração podem ser combinadas, se desejado.

[00148] As dosagens do vetor viral dependerão principalmente de fatores como a condição a ser tratada, a idade, o peso e a saúde do paciente e, portanto, podem variar entre os pacientes. Por exemplo, uma dosagem humana terapêuticamente eficaz do vetor viral está geralmente na faixa de cerca de 25 a cerca de 1000 microlitros a cerca de 100 mL de solução contendo concentrações de cerca de  $1 \times 10^9$  a  $1 \times 10^{16}$  genomas de vetor de vírus. A dosagem será ajustada para equilibrar o benefício terapêutico contra quaisquer efeitos colaterais e essas dosagens podem variar dependendo da aplicação terapêutica para a qual o vetor recombinante é empregado. Os níveis de expressão do transgene podem ser monitorados para determinar a frequência de dosagem resultando em vetores virais, preferencialmente vetores AAV contendo o minigene. Opcionalmente, regimes de dosagem semelhantes aos descritos para fins terapêuticos podem ser utilizados para imunização usando as composições da invenção.

[00149] As composições de vírus com defeito de replicação podem ser formuladas em unidades de dosagem para conter uma quantidade de

vírus com defeito de replicação que está na faixa de cerca de  $1,0 \times 10^9$  GC a cerca de  $1,0 \times 10^{16}$  GC (para tratar um sujeito médio de 70 kg em peso corporal) incluindo todos os números inteiros ou quantidades fracionárias dentro do intervalo e, de preferência  $1,0 \times 10^{12}$  GC a  $1,0 \times 10^{14}$  GC para um paciente humano. Em uma modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^9$ ,  $2 \times 10^9$ ,  $3 \times 10^9$ ,  $4 \times 10^9$ ,  $5 \times 10^9$ ,  $6 \times 10^9$ ,  $7 \times 10^9$ ,  $8 \times 10^9$ , ou  $9 \times 10^9$  GC por dose, incluindo todos os números inteiros ou quantidades fracionárias dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{10}$ ,  $2 \times 10^{10}$ ,  $3 \times 10^{10}$ ,  $4 \times 10^{10}$ ,  $5 \times 10^{10}$ ,  $6 \times 10^{10}$ ,  $7 \times 10^{10}$ ,  $8 \times 10^{10}$ , ou  $9 \times 10^{10}$  GC por dose, incluindo todos os números inteiros ou quantidades fracionárias dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{11}$ ,  $2 \times 10^{11}$ ,  $3 \times 10^{11}$ ,  $4 \times 10^{11}$ ,  $5 \times 10^{11}$ ,  $6 \times 10^{11}$ ,  $7 \times 10^{11}$ ,  $8 \times 10^{11}$ , ou  $9 \times 10^{11}$  GC por dose, incluindo todos os números inteiros ou quantidades fracionárias dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{12}$ ,  $2 \times 10^{12}$ ,  $3 \times 10^{12}$ ,  $4 \times 10^{12}$ ,  $5 \times 10^{12}$ ,  $6 \times 10^{12}$ ,  $7 \times 10^{12}$ ,  $8 \times 10^{12}$ , ou  $9 \times 10^{12}$  GC por dose, incluindo todos os números inteiros ou frações dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{13}$ ,  $2 \times 10^{13}$ ,  $3 \times 10^{13}$ ,  $4 \times 10^{13}$ ,  $5 \times 10^{13}$ ,  $6 \times 10^{13}$ ,  $7 \times 10^{13}$ ,  $8 \times 10^{13}$ , ou  $9 \times 10^{13}$  GC por dose, incluindo todos os números inteiros ou frações dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{14}$ ,  $2 \times 10^{14}$ ,  $3 \times 10^{14}$ ,  $4 \times 10^{14}$ ,  $5 \times 10^{14}$ ,  $6 \times 10^{14}$ ,  $7 \times 10^{14}$ ,  $8 \times 10^{14}$ , ou  $9 \times 10^{14}$  GC por dose, incluindo todos os números inteiros ou frações dentro do intervalo. Em outra modalidade, as composições são formuladas para conter pelo menos  $1 \times 10^{15}$ ,  $2 \times 10^{15}$ ,  $3 \times 10^{15}$ ,  $4 \times 10^{15}$ ,  $5 \times 10^{15}$ ,  $6 \times 10^{15}$ ,  $7 \times 10^{15}$ ,  $8 \times 10^{15}$ , ou  $9 \times 10^{15}$  GC incluindo todos os números inteiros ou quantidades fracionárias dentro do intervalo. Numa modalidade, para a aplicação humana, a dose pode variar de  $1 \times 10^{10}$  a cerca de  $1 \times 10^{12}$  GC por dose, incluindo todos os

números inteiros ou frações dentro do intervalo.

[00150] Estas doses acima podem ser administradas numa variedade de volumes de formulação de transportador, excipiente ou tampão, variando de cerca de 25 a cerca de 1000 microlitros, ou volumes maiores incluindo todos os números dentro do intervalo, dependendo do tamanho da área a ser tratada, o título viral usado, a via de administração e o efeito desejado do método. Numa modalidade, o volume de transportador, excipiente ou tampão é pelo menos cerca de 25 µL. Numa modalidade, o volume é de cerca de 50 µL. Numa outra modalidade, o volume é de cerca de 75 µL. Numa outra modalidade, o volume é de cerca de 100 µL. Numa outra modalidade, o volume é de cerca de 125 µL. Numa outra modalidade, o volume é de cerca de 150 µL. Numa outra modalidade, o volume é de cerca de 175 µL. Ainda numa outra modalidade, o volume é de cerca de 200 µL. Numa outra modalidade, o volume é de cerca de 225 µL. Ainda numa outra modalidade, o volume é de cerca de 250 µL. Ainda numa outra modalidade, o volume é de cerca de 275 µL. Ainda numa outra modalidade, o volume é de cerca de 300 µL. Ainda numa outra modalidade, o volume é de cerca de 325 µL. Numa outra modalidade, o volume é de cerca de 350 µL. Numa outra modalidade, o volume é de cerca de 375 µL. Numa outra modalidade, o volume é de cerca de 400 µL. Numa outra modalidade, o volume é de cerca de 450 µL. Numa outra modalidade, o volume é de cerca de 500 µL. Numa outra modalidade, o volume é de cerca de 550 µL. Numa outra modalidade, o volume é de cerca de 600 µL. Numa outra modalidade, o volume é de cerca de 650 µL. Numa outra modalidade, o volume é de cerca de 700 µL. Numa outra modalidade, o volume está entre cerca de 700 e 1000 µL.

[00151] Em certas modalidades, a dose pode estar no intervalo de cerca de  $1 \times 10^9$  GC/g de massa cerebral a cerca de  $1 \times 10^{12}$  GC/g de massa cerebral. Em certas modalidades, a dose pode estar no intervalo de



cerca de  $3 \times 10^{10}$  GC/g de massa cerebral a cerca de  $3 \times 10^{11}$  GC/g de massa cerebral. Em certas modalidades, a dose pode estar no intervalo de cerca de  $5 \times 10^{10}$  GC/g de massa cerebral a cerca de  $1,85 \times 10^{11}$  GC/g de massa cerebral.

[00152] Em outro aspecto, é fornecida uma suspensão aquosa adequada para administração a um sujeito. Em uma modalidade, a suspensão inclui um líquido de suspensão aquoso e cerca de  $1 \times 10^9$  partículas virais a cerca de  $1 \times 10^{13}$  GC ou partículas virais por olho de um vírus adenoassociado recombinante (rAAV), conforme descrito neste documento, útil como um agente terapêutico para o tratamento ou prevenção de doenças oculares. Em uma modalidade, a suspensão é adequada para injeção subretiniana ou intravítrea.

[00153] Numa modalidade, os construtos virais podem ser distribuídos em doses de pelo menos cerca de  $1 \times 10^9$  GCs a cerca de  $1 \times 10^{15}$  GC, ou cerca de  $1 \times 10^{11}$  GC a  $5 \times 10^{13}$  GC. Os volumes adequados para distribuição destas doses e concentrações podem ser determinados por um versado na técnica. Por exemplo, volumes de cerca de  $1 \mu\text{L}$  a  $150 \text{ mL}$  podem ser selecionados, com os volumes mais altos sendo selecionados para adultos. Tipicamente, para recém-nascidos, um volume adequado é de cerca de  $0,5 \text{ mL}$  a cerca de  $10 \text{ mL}$ , para bebês mais velhos, podem ser selecionados cerca de  $0,5 \text{ mL}$  a cerca de  $15 \text{ mL}$ . Para crianças pequenas, um volume de cerca de  $0,5 \text{ mL}$  a cerca de  $20 \text{ mL}$  pode ser selecionado. Para crianças, volumes de até  $30 \text{ mL}$  podem ser selecionados. Para pré-adolescentes e adolescentes, volumes de até  $50 \text{ mL}$  podem ser selecionados. Em ainda outras modalidades, um paciente pode receber uma administração intratecal em um volume de cerca de  $5 \text{ mL}$  a cerca de  $15 \text{ mL}$  são selecionados, ou cerca de  $7,5 \text{ mL}$  a cerca de  $10 \text{ mL}$ . Outros volumes e dosagens adequados podem ser determinados. A dosagem será ajustada para equilibrar o benefício terapêutico

contra quaisquer efeitos colaterais e essas dosagens podem variar dependendo da aplicação terapêutica para a qual o vetor recombinante é empregado.

[00154] Os vetores recombinantes descritos acima podem ser distribuídos às células hospedeiras de acordo com os métodos publicados. O rAAV, de preferência suspenso num transportador fisiologicamente compatível, pode ser administrado a um paciente mamífero humano ou não humano. Em certas modalidades, para administração a um paciente humano, o rAAV é adequadamente suspenso em uma solução aquosa contendo solução salina, um tensoativo e um sal ou mistura de sais fisiologicamente compatível. Adequadamente, a formulação é ajustada para um pH fisiologicamente aceitável, por exemplo, na faixa de pH 6 a 9, ou pH 6,5 a 7,5, pH 7,0 a 7,7 ou pH 7,2 a 7,8. Como o pH do líquido cefalorraquidiano é de cerca de 7,28 a cerca de 7,32, para distribuição intratecal, um pH dentro dessa faixa pode ser desejado; enquanto que para distribuição intravenosa, subretinal e intravítrea, um pH de cerca de 6,8 a cerca de 7,2 pode ser desejado. No entanto, outros pHs dentro das faixas mais amplas e esses subintervalos podem ser selecionados para outra via de distribuição.

[00155] Numa outra modalidade, a composição inclui um transportador, diluente, excipiente e/ou adjuvante. Os transportadores adequados podem ser facilmente selecionados por um versado na técnica, tendo em vista a indicação para a qual o vírus de transferência é direcionado. Por exemplo, um transportador adequado inclui solução salina, que pode ser formulada com uma variedade de soluções tampão (*por exemplo*, solução salina tamponada com fosfato). Outros transportadores exemplificativos incluem solução salina estéril, lactose, sacarose, fosfato de cálcio, gelatina, dextrano, ágar, pectina, óleo de amendoim, óleo de gergelim e água. O tampão/transportador deve incluir um componente que impeça a aderência do rAAV ao tubo de infusão, mas não

interfira com a atividade de ligação do rAAV *in vivo*. Um tensoativo adequado, ou combinação de tensoativos, pode ser selecionado de tensoativos não iônicos que não sejam tóxicos. Em uma modalidade, um tensoativo copolímero de bloco difuncional que termina em grupos hidroxil primários é selecionado, por exemplo, como Pluronic® F68 [BASF], também conhecido como Poloxamer 188, que tem um pH neutro, tem um peso molecular médio de 8400. Outros tensoativos e outros poloxâmeros podem ser selecionados, ou seja, copolímeros triblocos não iônicos compostos de uma cadeia hidrofóbica central de polioxipropileno (poli(óxido de propileno)) flanqueada por duas cadeias hidrofílicas de polioxietileno (poli(óxido de etileno)), SOLUTOL HS 15 (Macrogol-15 Hidroxiestearato), LABRASOL (Glicerídeo polioxi caprílico), éter oleílico de polióxi, TWEEN (ésteres de ácidos graxos de polioxietileno sorbitano), etanol e polietileno glicol. Numa modalidade, a formulação contém um poloxâmero. Estes copolímeros são comumente nomeados com a letra "P" (para poloxâmero) seguidos de três dígitos: os dois primeiros dígitos x 100 dão a massa molecular aproximada do núcleo de polioxipropileno, e o último dígito x 10 indica o percentual de polioxietileno. Em uma modalidade, o Poloxamer 188 é selecionado. O tensoativo pode estar presente numa quantidade até cerca de 0,0005% a cerca de 0,001% da suspensão. Em um exemplo, a formulação pode conter, por exemplo, solução salina tamponada compreendendo um ou mais de cloreto de sódio, bicarbonato de sódio, dextrose, sulfato de magnésio (por exemplo, sulfato de magnésio ·7H<sub>2</sub>O), cloreto de potássio, cloreto de cálcio (por exemplo, cloreto de cálcio ·2H<sub>2</sub>O), fosfato de sódio dibásico e misturas dos mesmos, em água. Adequadamente, para distribuição intratecal, a osmolaridade está dentro de um intervalo compatível com o líquido cefalorraquidiano (por exemplo, cerca de 275 a cerca de 290); ver, por exemplo, [emedicine.medscape.com/article/2093316-overview](http://emedicine.medscape.com/article/2093316-overview).

Opcionalmente, para distribuição intratecal, um diluente disponível comercialmente pode ser usado como um agente de suspensão ou em combinação com outro agente de suspensão e outros excipientes opcionais. Ver, por exemplo, a solução Elliotts B® [Lukare Medical]. Em outras modalidades, a formulação pode conter um ou mais intensificadores de permeação. Exemplos de intensificadores de permeação adequados podem incluir, por exemplo, manitol, glicocolato de sódio, taurocolato de sódio, desoxicolato de sódio, salicilato de sódio, caprilato de sódio, caprato de sódio, lauril sulfato de sódio, éter polioxietileno-9-laurel ou EDTA.

[00156] Opcionalmente, as composições da invenção podem conter, além do rAAV e do(s) transportador(es), outros ingredientes farmacêuticos convencionais, tais como conservantes ou estabilizadores químicos. Conservantes exemplificativos adequados incluem clorobutanol, sorbato de potássio, ácido sórbico, dióxido de enxofre, propil galato, parabenos, etil vanilina, glicerina, fenol e paraclorofenol. Estabilizadores químicos adequados incluem gelatina e albumina.

[00157] As composições de acordo com a presente invenção podem compreender um transportador farmacologicamente aceitável, tal como definido acima. Adequadamente, as composições aqui descritas compreendem uma quantidade eficaz de um ou mais AAV suspensos em um transportador farmacologicamente adequado e/ou misturados com excipientes adequados concebidos para distribuição ao sujeito por injeção, bomba osmótica, cateter intratecal ou para distribuição por outro dispositivo ou via. Em um exemplo, a composição é formulada para distribuição intratecal.

[00158] Como usado neste documento, os termos "distribuição intratecal" ou "administração intratecal" referem-se a uma via de administração de drogas por injeção no canal medular, mais especificamente no espaço subaracnoide, de modo a atingir o líquido cefalorraquidiano

(CSF). A distribuição intratecal pode incluir punção lombar, intraventricular (incluindo intracerebroventricular (ICV)), suboccipital/intracisternal e/ou C 1-2. Por exemplo, o material pode ser introduzido para difusão através do espaço subaracnoide por meio de punção lombar. Em outro exemplo, a injeção pode estar na cisterna magna.

[00159] Como usado neste documento, os termos "distribuição intracisternal" ou "administração intracisternal" se referem a uma via de administração de drogas diretamente no líquido cefalorraquidiano da cisterna magna cerebellomedularis, mais especificamente por meio de uma punção suboccipital ou por injeção direta em cisterna magna ou via tubo permanentemente posicionado.

[00160] Em um aspecto, os vetores fornecidos neste documento podem ser administrados por via intratecal por meio do método e/ou dispositivo. Ver, por exemplo, WO 2017/181113, que é aqui incorporado por referência. Alternativamente, outros dispositivos e métodos podem ser selecionados. O método compreende as etapas de avançar uma agulha espinhal na cisterna magna de um paciente, conectar um comprimento de tubo flexível a um cubo proximal da agulha espinhal e uma porta de saída de uma válvula a uma extremidade proximal do tubo flexível e depois das referidas etapas de avanço e conexão e depois de permitir que o tubo seja autoiniciado com o líquido cefalorraquidiano do paciente, conectar um primeiro recipiente contendo uma quantidade de solução isotônica a uma porta de entrada de descarga da válvula e, posteriormente, conectar um segundo recipiente contendo uma quantidade de uma composição farmacêutica para uma porta de entrada de vetor da válvula. Após conectar o primeiro e o segundo recipientes à válvula, é aberto um caminho para o fluxo de fluido entre a porta de entrada do vetor e a porta de saída da válvula e a composição farmacêutica é injetada no paciente através da agulha espinhal e após a injeção da com-

posição farmacêutica, é aberto um caminho para o fluxo de fluido através da porta de entrada de descarga e da porta de saída da válvula e a solução isotônica é injetada na agulha espinhal para liberar a composição farmacêutica no paciente.

[00161] Este método e este dispositivo podem cada um ser opcionalmente utilizados para distribuição intratecal das composições aqui fornecidas. Alternativamente, outros métodos e dispositivos podem ser utilizados para essa distribuição intratecal.

[00162] É de notar que o termo "um" ou "uma" se refere a um ou mais. Como tal, os termos "um" (ou "uma"), "um ou mais" e "pelo menos um" são aqui usados indistintamente.

[00163] As palavras "compreendem", "compreende" e "compreendendo" devem ser interpretadas inclusivamente em vez de exclusivamente. As palavras "consiste", "consistir" e suas variantes devem ser interpretadas exclusivamente, e não de forma inclusiva. Embora várias modalidades no relatório descritivo sejam apresentadas utilizando linguagem "compreensiva", sob outras circunstâncias, uma modalidade relacionada também se destina a ser interpretada e descrita usando a linguagem "consistindo em" ou "consistindo essencialmente em".

[00164] Como usado neste documento, o termo "cerca de" significa uma variabilidade de 10% ( $\pm 10\%$ ) da referência dada, salvo indicação em contrário.

[00165] Como usado neste documento, "doença", "distúrbio" e "condição" são usados de forma intercambiável, para indicar um estado anormal em um sujeito.

[00166] Salvo definido em contrário neste pedido, os termos técnicos e científicos aqui utilizados têm o mesmo significado como vulgarmente entendido por um versado na técnica e por referência a textos publicados, que proporcionam aos versados na técnica um guia geral para muitos dos termos utilizados no presente pedido.

[00167] O termo "expressão" é aqui utilizado no seu significado mais amplo e compreende a produção de RNA ou de RNA e proteína. Com relação ao RNA, o termo "expressão" ou "tradução" se refere, em particular, à produção de peptídeos ou proteínas. A expressão pode ser transitória ou estável.

[00168] Como usado neste documento, o termo "título de NAb" mede a quantidade de anticorpo neutralizante (por exemplo, NAB anti-AAV) que neutraliza o efeito fisiológico do seu epítipo alvo (por exemplo, um AAV). Os títulos de NAb anti-AAV podem ser medidos como descrito em, por exemplo, Calcedo, R., et al., Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. *Journal of Infectious Diseases*, 2009. 199(3): p. 381-390, que é aqui incorporado por referência.

[00169] Como usado neste documento, um "cassete de expressão" refere-se a uma molécula de ácido nucleico que compreende uma sequência de codificação, promotor e pode incluir outras sequências regulatórias para esse fim, cujo cassete pode ser distribuído através de um elemento genético (*por exemplo*, um plasmídeo) para uma célula hospedeira de empacotamento e empacotado no capsídeo de um vetor viral (*por exemplo*, uma partícula viral). Tipicamente, tal cassete de expressão para gerar um vetor viral contém a sequência de codificação para o produto de gene aqui descrito flanqueada por sinais de empacotamento do genoma viral e outras sequências de controle de expressão, tais como aquelas aqui descritas.

[00170] A abreviatura "sc" refere-se a autocomplementaridade. "AAV autocomplementar" refere-se a um construto no qual uma região de codificação transportada por uma sequência de ácido nucleico de AAV recombinante foi concebida de modo a formar um modelo de DNA de fita dupla intramolecular. Após a infecção, em vez de esperar pela síntese mediada por células da segunda fita, as duas metades complementares

de scAAV irão associar-se para formar uma unidade de DNA de fita dupla (dsADN) que está pronta para replicação e transcrição imediatas. *Ver, por exemplo*, D M McCarty *et al*, “Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis”, *Gene Therapy*, (August 2001), Vol 8, Number 16, Páginas 1248-1254. Os AAVs autocomplementares são descritos, *por exemplo*, em Patentes U.S. 6.596.535; 7.125.717; e 7.456.683, cada um dos quais é aqui incorporado por referência na sua totalidade.

[00171] Como usado neste documento, o termo “operacionalmente ligado” refere-se a ambas as sequências de controle de expressão que são contíguas com o gene de interesse e sequências de controle de expressão que atuam em trans ou a uma distância para controlar o gene de interesse.

[00172] O termo "heterólogo", quando usado com referência a uma proteína ou um ácido nucleico, indica que a proteína ou o ácido nucleico compreende duas ou mais sequências ou subsequências que não são encontradas na mesma relação entre si na natureza. Por exemplo, o ácido nucleico é tipicamente produzido de forma recombinante, tendo duas ou mais sequências de genes não relacionados dispostos para fazer um novo ácido nucleico funcional. Por exemplo, em uma modalidade, o ácido nucleico tem um promotor de um gene arranjado para dirigir a expressão de uma sequência de codificação de um gene diferente. Assim, com referência à sequência de codificação, o promotor é heterólogo.

[00173] Um “vírus com replicação defeituosa” ou “vetor viral” refere-se a uma partícula viral, sintética ou artificial na qual um cassete de expressão contendo um gene de interesse é empacotado em um capsídeo ou envelope viral, onde qualquer sequência genômica viral também é



empacotada dentro do capsídeo ou envelope viral é deficiente na replicação; isto é, eles não podem gerar vírions de progênes, mas retêm a capacidade de infectar células-alvo. Numa modalidade, o genoma do vetor viral não inclui genes que codificam as enzimas necessárias para replicar (o genoma pode ser engenheirado para ser "sem reforço" - contendo apenas o transgene de interesse flanqueado pelos sinais necessários para amplificação e embalagem do genoma artificial), mas esses genes podem ser fornecidos durante a produção. Por conseguinte, considera-se seguro para uso em terapia genética, uma vez que a replicação e a infecção por vírions da progênie não podem ocorrer exceto na presença da enzima viral necessária para a replicação.

[00174] Em muitos casos, as partículas de rAAV são referidas como resistentes a DNase. No entanto, além desta endonuclease (DNase), outras endo- e exonucleases também podem ser usadas nas etapas de purificação aqui descritas, para remover ácidos nucleicos contaminantes. Tais nucleases podem ser selecionadas para degradar DNA de fita simples e/ou DNA de fita dupla e RNA. Tais etapas podem conter uma única nuclease ou misturas de nucleases direcionadas a diferentes alvos e podem ser endonucleases ou exonucleases.

[00175] O termo "resistente à nuclease" indica que o capsídeo de AAV foi totalmente montado em torno do cassete de expressão que é projetado para entregar um transgene a uma célula hospedeira e protege essas sequências genômicas empacotadas da degradação (digestão) durante as etapas de incubação de nuclease projetadas para remover ácidos nucleicos contaminantes que podem estar presentes no processo de produção.

[00176] O termo "tradução" no contexto da presente invenção refere-se a um processo no ribossomo, em que uma fita de mRNA controla a montagem de uma sequência de aminoácidos para gerar uma proteína ou um peptídeo.

[00177] Conforme usado ao longo deste relatório descritivo e das reivindicações, os termos "compreendendo" e "incluindo" incluem outros componentes, elementos, números inteiros, etapas e semelhantes. Por outro lado, o termo "consistindo" e suas variantes são exclusivos de outros componentes, elementos, inteiros, etapas e semelhantes.

[00178] Conforme descrito acima, o termo "cerca de" quando usado para modificar um valor numérico significa uma variação de  $\pm 10\%$ , a menos que especificado de outra forma.

[00179] Os exemplos seguintes são apenas ilustrativos e não pretendem limitar a presente invenção.

#### EXEMPLOS:

[00180] Os exemplos a seguir relatam a extensa desamidação de AAV8 e diversos sorotipos de AAV adicionais, com evidências de apoio de abordagens estruturais, bioquímicas e de espectrometria de massa. A extensão da desamidação em cada local era dependente da idade do vetor e de múltiplos fatores estruturais e de sequência primária 3D, mas era amplamente independente das condições de recuperação e purificação do vetor. Demonstramos o potencial de desamidação para impactar a atividade de transdução do vetor e correlacionar uma perda de ponto de tempo precoce na atividade do vetor com a desamidação espontânea de progresso rápido em várias asparaginas AAV8. Exploramos estratégias mutacionais que estabilizam as amidas da cadeia lateral, melhorando a transdução do vetor e reduzindo a variabilidade molecular lote a lote, que é uma preocupação fundamental na fabricação de produtos biológicos. Este estudo ilustra um aspecto até então desconhecido da heterogeneidade do capsídeo de AAV e destaca sua importância no desenvolvimento desses vetores para terapia genética.

[00181] No Exemplo 1, a caracterização de modificações pós-tradução no capsídeo do vetor AAV8 por eletroforese em gel uni e bidimensional, espectrometria de massa e modelagem estrutural de novo. Após

a identificação de um número de sítios de desamidação putativos na superfície do capsídeo, avaliamos seu impacto na estrutura e função do capsídeo *in vitro* e *in vivo*. O Exemplo 1 estende ainda mais esta análise para AAV9 para determinar se este fenômeno se aplica a sorotipos diferentes de AAV8, confirmando que a desamidação do capsídeo de AAV não é específica para sorotipos. Os Exemplos 2-5 mostram a desamidação em AAVs distintos.

EXEMPLO 1: Desamidação de aminoácidos na superfície de capsídeos de vírus adenoassociados

A. Materiais e Métodos

1. Eletroforese em gel 1D e 2D

[00182] Para a análise de eletroforese em gel de poliacrilamida 1D SDS (SDS-PAGE), primeiro desnaturamos os vetores AAV a 80°C por 20 minutos na presença de dodecil sulfato de lítio e agente redutor. Em seguida, os aplicamos em um gel de Bis-Tris 4-12% por 90 minutos a 200 V e corados com azul de Coomassie. Para os dados da FIG. 1A - FIG. 1D, Kendrick Laboratories, Inc. (Madison, WI) realizou a eletroforese em gel 2D. Para experimentos subsequentes, realizamos SDS-PAGE 2D internamente. Para isso, combinamos  $3 \times 10^{11}$  GCs do vetor AAV com marcador de turbonuclease 500U (Accelagen, San Diego, CA) em 150µL solução salina tamponada com fosfato (PBS) com 35mM NaCl e 1mM MgCl<sub>2</sub> e incubada em 37°C por dez minutos. Em seguida, adicionamos nove volumes de etanol absoluto, agitamos as amostras em vórtice e as incubamos a -80°C pelo menos duas horas, seguido de incubação em gelo por cinco minutos e centrifugação em velocidade máxima por 30 minutos a 15°C. Decantamos o sobrenadante e secamos ao ar o pelete, que então ressuspendemos em tampão de ressuspensão #1 [SDS 0,15%, ditionitrotol 50 mM (DTT), Tris 10 mM pH 7,5 e 1µL de pH6-9 anfólitos, ThermoFisher ZM0023, adicionado no dia de, em ddH<sub>2</sub>O] e incubamos sem perturbações à temperatura ambiente. Após

30 minutos, agitamos os tubos de amostra para misturá-los, adicionamos 1µg marcador de conalbumina de frango (Sigma Aldrich, St. Louis, MO), e incubamos as amostras em 37°C por 30 minutos, sacudindo para misturar em 15 minutos. As amostras foram então transferidas para 50°C por 15-20 minutos, agitadas em vórtice, incubadas em 95°C por 2,5 minutos e deixou-se esfriar antes de ser centrifugadas à velocidade máxima por um minuto e brevemente agitadas em vórtice. Em seguida, misturamos 10µL de cada amostra com 140µL tampão de ressuspensão # 2 (ureia 9,7 M, CHAPS 2%, azul de bromofenol 0,002% e anfólitos 0,05%, descrito acima, adicionado no dia de, em ddH<sub>2</sub>O) e incubamos à temperatura ambiente durante dez minutos. Em seguida, aplicamos as misturas em tiras de gradiente de pH imobilizado (IPG) de pH 6-10 (ThermoFisher Waltham, MA) e as executamos no sistema ZOOM IPGRunner de acordo com as instruções do fabricante. Usamos os seguintes parâmetros de focalização isoeétrica: 100-1.000 V por 120 minutos, 1.000-2.000 V por 120 minutos, 2.000 V por 120 minutos, limites de 0,1 W e 0,05 mA por tira de execução. As tiras de IPG foram então reduzidas e carregadas em um gel de Bis-Tris de 4-12% de poço único e executadas em 1D como descrito acima. Determinamos a migração relativa de VPs de AAV por comparação com proteínas de controle interno turbonuclease (Accelagen, 27kDa) e conalbumina de clara de ovo de frango (Sigma Aldrich, 76kDa, pI 6,0-6,6).

## 2. Produção vetorial

[00183] O University of Pennsylvania Vector Core produziu vetores AAV recombinantes para eletroforese em gel 1D e 2D e experimentos de espectrometria de massa e os purificou por cloreto de céσιο ou gradientes de iodixanol como descrito anteriormente. (Lock M, et al. *Hum Gene Ther* 2010; 21(10):1259-71; Gao GP, et al. *Proc Natl Acad Sci USA*. 2002; 99(18):11854-9). Produzimos os vetores purificados por afi-

nidade da seguinte forma: cultivamos células HEK293 em dez recipientes hiperestack de 36 camadas (Corning), cotransfectamos com uma mistura de plasmídeo de genoma de vetor (pAAV-LSP-IVS2.hFIXco-WPRE-bGH), trans plasmídeo contendo genes *AAV2 rep* e *AAV8 cap*, e plasmídeo auxiliar de adenovírus. Usamos PEIpro (PolyPlus) como o reagente de transfecção. Cinco dias após a transfecção, o sobrenadante foi colhido, clarificado através de filtros Sartoguard PES Midicap (SartoriusStedim) e tratado com benzonase (Millipore), após o qual adicionamos sal para trazê-lo a 0,6M. O material de colheita a granel clarificado foi concentrado dez vezes por filtração de fluxo tangencial (TFF) e, em seguida, diafiltrado contra quatro volumes de tampão de carregamento de coluna de afinidade. Capturamos os vetores em uma coluna de afinidade POROS CaptureSelect (ThermoFisher) e eluímos o pico do vetor em pH baixo diretamente no tampão de neutralização. Diluímos o eluato neutralizado em um tampão de ligação de alto pH e o carregamos em uma coluna de polimento de troca aniônica (Cimultus QA-8; Bia Separations), onde a preparação foi enriquecida para partículas contendo genoma (completas). As partículas do vetor completo foram eluídas com um gradiente de eluição de sal raso e neutralizadas imediatamente. Finalmente, submetemos o vetor a uma segunda rodada de TFF para concentração final e troca de tampão em tampão de formulação (PBS + 0,001% de F-68 plurônico).

[00184] Nós produzimos vetores mutantes para ensaios *in vitro* por transfecção tripla em pequena escala de células HEK293 em placas de seis poços. Misturamos 5,6 µL de uma solução de polietilenimina 1mg/mL em 90µL de meio sem soro com DNA de plasmídeo (0,091µg de plasmídeo cis, 0,91µg de plasmídeo trans, 1,82 µg de plasmídeo deltaF6 Ad-helper, em 90µL de meio sem soro), incubamos em temperatura ambiente por 15 minutos, e adicionamos às células e 0,8 mL adicionais de meio fresco livre de soro. No dia seguinte, substituímos 0,5mL

do meio superior por meio sérico completo. Colhemos o vetor três dias após a transfecção por três ciclos de congelamento/descongelamento seguido por centrifugação para remover os resíduos celulares e colheita do sobrenadante. O plasmídeo cis continha um cassete de transgene que codifica o transgene da luciferase do pirilampo sob o controle do promotor da beta-actina de frango (CB7) com o íntron quimérico Pro-mega e sinal de poliadenilação da beta-globina de coelho (RBG). O plasmídeo trans codificou o gene *cap wtAAV8*; para gerar variantes de *cap AAV8* mutantes, usamos o kit Quikchange Lightning Mutagenesis (Agilent Technologies, Wilmington, DE). O vetor foi titulado conforme descrito anteriormente. (Lock M, et al. *Hum Gene Ther* 2010; 21(10):1259-71).

[00185] Para experimentos de produção de vetor de curso de tempo, geramos vetor por transfecção tripla em média escala de células HEK293 em placas de cultura de tecidos de 15 cm. Por placa, misturaram-se 36µL de uma solução de polietilenimina 1mg/mL em meio sem soro 2mL com DNA de plasmídeo (0,6µg cis plasmídeo, 5,8µg trans plasmídeo, 11,6µg deltaF6 Ad-auxiliar plasmídeo), incubado em temperatura ambiente por 15 minutos, e adicionaram-se às células em aproximadamente 60% de confluência em placas renovadas com 14ml de meio sem soro. No dia seguinte, foram substituídos 8ml do meio superior por meio sérico completo e fresco. Colheu-se o vetor coletando todos os meios de comunicação, raspando as células do prato e congelando a -80<sup>0</sup>C. Recuperou-se o vetor bruto da mistura sobrenadante/células aplicando 3 ciclos de congelamento/descongelamento e clarificando o lisado por centrifugação. Purificou-se e concentrou-se o vetor para análise de espectrometria de massa adicionando benzonase, Tris 1M pH 7,5 e NaCl 5M ao lisado clarificado para concentrações finais de Tris 20 mM e NaCl 360 mM. Capturaram-se vetores em uma coluna de afinidade POROSCaptureSelect de 1 ml e eluiu-se o pico do vetor em

pH baixo diretamente no tampão de neutralização. As frações foram analisadas por absorvância a 280nm, e a fração mais concentrada foi submetida à análise de espectrometria de massa

[00186] Para experimentos *in vivo*, produzimos vetores conforme descrito anteriormente com um capsídeo wtAAV8 ou com um dos 6 mutantes de desamidação; o cassete do transgene incluiu um promotor CB7, íntron PI, transgene de luciferase de pirilampo e sinal de poliadenilação RBG (Lock M, et al. *Hum Gene Ther* 2010; 21(10):1259-71).

### 3. Execução/digestão/análise de especificação de massa

[00187] *Materiais:* Adquirimos bicarbonato de amônio, DTT, iodoacetamida (IAM) e água enriquecida com  $^{18}\text{O}$  (97,1% de pureza) da Sigma (St. Louis, MO); e acetonitrila, ácido fórmico, ácido trifluoroacético (TFA), cloridrato de guanidina 8 M (GndHCl) e tripsina da Thermo Fischer Scientific (Rockford, IL).

[00188] *Digestão com tripsina:* preparamos soluções de estoque de 1M DTT e 1,0M de iodoacetamida. As proteínas do capsídeo foram desnaturadas e reduzidas a 90°C por dez minutos na presença de 10mM DTT e 2M GndHCl. Permitiu-se que as amostras esfriassem até a temperatura ambiente e, em seguida, alquilamos com 30 mM IAM em temperatura ambiente por 30 minutos no escuro. Extinguiu-se a reação de alquilação com a adição de 1 mL de DTT. Adicionou-se bicarbonato de amônio 20 mM (pH 7,5-8) à solução de proteína desnaturada em um volume que diluiu a concentração final de GndHCl para 200 mM. Adicionou-se solução de tripsina para uma razão de 1:20 de tripsina para proteína e incubou-se a 37°C durante a noite. Após a digestão, adicionou-se TFA a uma concentração final de 0,5% para extinguir a reação de digestão.

[00189] Para experimentos com  $^{18}\text{O}$ -água, a amostra de capsídeo foi primeiro trocada por tampão em 100 mM de bicarbonato de amônio preparado em  $^{18}\text{O}$ -água usando colunas de dessalinização spin Zeba

(Thermo Scientific, Rockford, IL). Para garantir a remoção completa da água da amostra, realizou-se a troca do tampão duas vezes. Prepararam-se soluções de estoque de 1M DTT e 1M IAM em 18O-água. Seguiu-se as mesmas etapas de desnaturação, alquilação e digestão acima com reagentes e tampões 18O-água.

[00190] *Cromatografia líquida espectrometria de massa em tandem:* Foram realizados cromatografia online com uma coluna Acclaim PepMap (15 cm de comprimento, 300µm de diâmetro interno) e um sistema Thermo UltiMate 3000 RSLC (Thermo Fisher Scientific) acoplado a um Q Exactive HF com uma fonte NanoFlex (Thermo Fisher Scientific). Durante a análise online, a temperatura da coluna foi mantida a 35°C. Foram separados os peptídeos com gradiente de fase móvel A (água MilliQ com ácido fórmico 0,1%) e fase móvel B (acetonitrila com ácido fórmico 0,1%). Executou-se o gradiente de 4% B a 6% B ao longo de 15 minutos, a 10% B por 25 minutos (40 minutos no total) e, em seguida, a 30% B por 46 minutos (86 minutos no total). Foram carregadas as amostras diretamente na coluna. O tamanho da coluna foi de 75 cm x 15 um ID e foi embalado com meio C18 de 2 microns (Acclaim PepMap). Devido às etapas de carregamento, introdução e lavagem, o tempo total para cada execução de espectrometria de massa em tandem de cromatografia líquida foi de cerca de duas horas.

[00191] Foram adquiridos dados de espectrometria de massa usando um método top-20 dependente de dados no espectrômetro de massa Q Exactive HF, escolhendo dinamicamente os íons precursores ainda não sequenciados mais abundantes das varreduras de pesquisa (200–2000 m/z). Realizou-se o sequenciamento por meio de fragmentação de dissociação colisional de alta energia com um valor alvo de íons 1e5 determinado com controle de ganho automático preditivo; realizou-se o isolamento dos precursores com uma janela de 4m/z. Adquiriram-se varreduras de levantamento com uma resolução de 120.000 a



200m/z. Definiu-se a resolução dos espectros de HCD para 30.000 em m/z200 com um tempo máximo de injeção de íons de 50 ms e uma energia de colisão normalizada de 30. Ajustou-se o nível de RF da lente S em 50, o que deu uma transmissão ideal da região m/z ocupada pelos peptídeos de nosso digest. Foram excluídos íons precursores com estados de carga simples, não atribuídos ou seis e mais altos da seleção de fragmentação.

[00192] *Processamento de dados:* Utilizou-se o software BioPharma Finder 1.0 (Thermo Fisher Scientific) para analisar todos os dados adquiridos. Para o mapeamento de peptídeos, realizaram-se pesquisas usando um banco de dados de proteína de entrada única FASTA com carbamidometilação definida como uma modificação fixa e oxidação, desamidação e fosforilação definidas como modificações variáveis. Usou-se uma precisão de massa de 10 ppm, uma alta especificidade de protease e um nível de confiança de 0,8 para espectros de espectrometria de massa em tandem. A identificação por espectrometria de massa de peptídeos desamidados é relativamente simples, pois a desamidação adiciona à massa da molécula intacta +0,984 Da (a diferença de massa entre os grupos–OH e –NH<sub>2</sub>). Determinou-se a percentagem de desamidação de um determinado peptídeo dividindo a área de massa do peptídeo desamidado pela soma da área dos peptídeos desamidados e nativos. Considerando o número de sítios de desamidação possíveis, as espécies isobáricas que são desamidadas em diferentes sítios podem comigrar em um único pico. Consequentemente, fragmentos de íons originários de peptídeos com múltiplos sítios de desamidação potenciais podem ser usados para localizar ou diferenciar múltiplos sítios de desamidação. Nestes casos, as intensidades relativas dentro dos padrões de isótopos observados podem ser usadas para determinar especificamente a abundância relativa dos diferentes isômeros de peptídeo desamidado. Este método presume que a eficiência de fragmentação

para todas as espécies isoméricas é a mesma e independente do sítio de desamidação. Esta abordagem permite a definição dos sítios específicos envolvidos na desamidação e as combinações potenciais envolvidas na desamidação.

[00193] *Processamento de dados secundários:* A análise secundária de espectrometria de massa bruta foi realizada na Universidade de Maryland, Condado de Baltimore, usando o seguinte método. O software Peaks Studio v5.3 (Bioinformatics Solutions Inc.) foi usado para todas as análises de espectrometria de massa. O refinamento de dados dos arquivos de dados brutos foi realizado com os seguintes parâmetros: uma tolerância m/z do precursor de  $\leq 10$  ppm e estado de carga do precursor com um mínimo de 2 e máximo de 4. O sequenciamento *de novo* do espectro de entrada foi realizado usando o algoritmo Peaks com uma tolerância de erro de íon precursor de 10 ppm e tolerâncias de erro de íon de produto de 0,1Da. A enzima de digestão foi definida como tripsina, as modificações variáveis foram oxidação, fosforilação e desamidação, e a modificação fixa foi carbamidometilação de cisteína.

#### 4. Análise estrutural do capsídeo AAV

[00194] Foram obtidas as coordenadas atômicas AAV8, fatores estruturais e modelo de capsídeo associado do Banco de Dados de Proteínas RCSB (PDB ID: 3RA8). Realizou-se o refinamento da estrutura e gerou-se uma densidade de elétrons independente da sequência de aminoácidos primária de AAV8 VP3 para uso em análise estrutural tridimensional (3D) do capsídeo. Realizou-se esta análise a fim de observar a densidade de elétrons do ácido isoaspártico no capsídeo AAV8 que não foi enviesado pela sequência primária esperada de AAV8 VP3. Usando a estrutura resultante, modelamos as quatro asparaginas na sequência primária AAV8 VP3 com N+1 glicinas como ácidos isoaspárticos e, em seguida, refinou-se a estrutura do capsídeo AAV8 usando

software de Cristalografia e Sistema de NMR (CNS), impondo estritamente as matrizes icosaédricas não cristalográficas usando o protocolo de refinamento padrão (Brunger AT, et al. *Acta Crystallogr D Biol Crystallogr* 1998; 54(Pt 5):905-21). Foi obtido um modelo estrutural de ácido isoaspártico do banco de dados HIC-UP, seguido pela geração de um dicionário molecular em PRODRG para refinamento da estrutura (Kleywegt GJ *Acta Crystallogr D Biol Crystallogr* 2007; 63(Pt 1):94-100). Em seguida, calculou-se o mapa de densidade eletrônica média do capsídeo AAV8 (também em CNS) e foi visto usando o software COOT, seguido por pequenos ajustes do modelo resultante para ajustar os resíduos de ácido isoaspártico modelados no mapa de densidade eletrônica (Emsley P and Cowtan K *Acta Crystallogr D Biol Crystallogr* 2004; 60(Pt 12 Pt 1):2126-32). Repetimos este protocolo para modelar adicionalmente N512 na sequência primária AAV9 VP3 com N+1 glicinas (PDB ID: 3UX1). Foram geradas todas as figuras usando COOT, PyMol e UCSF Chimera (Emsley P and Cowtan K *Acta Crystallogr D Biol Crystallogr* 2004; 60(Pt 12 Pt 1):2126-32; DeLano WL *PyMOL: An Open-Source Molecular Graphics Tool* Vol. 40, 2002:82-92; Pettersen EF, et al. *J Comput Chem* 2004; 25(13):1605-12). Foi obtida uma série de estruturas de proteínas desamidadas previamente identificadas (PDB IDs: 1DY5, 4E7G, 1RTU, 1W9V, 4E7D e 1C9D) para comparação de seu mapa de densidade de elétrons para resíduos de ácido isoaspártico desamidado com nossos resíduos de ácido isoaspártico modelados de AAV8 e AAV9 (Rao FV, et al. *Chem Biol* 2005; 12(1):65-76; Noguchi S, et al. *Biochemistry* 1995; 34(47):15583-91; Esposito L, et al. *J Mol Biol* 2000; 297(3):713-32).

[00195] Determinaram-se fatores de temperatura para resíduos desamidados calculando a média dos fatores de temperatura para cada átomo de cada resíduo de asparagina relatado nas coordenadas atômicas da estrutura cristalina AAV8 ou AAV9 (PDB ID: 3RA8, 3UX1).

## 5. Estudos em animais

[00196] O Comitê Institucional de Cuidado e Uso de Animais da Universidade da Pensilvânia aprovou todos os procedimentos com animais. Para avaliar o desempenho do vetor, injetaram-se camundongos C57BL/6 de oito semanas de idade por via intravenosa por meio de injeção na veia da cauda com  $3 \times 10^{10}$  GCs de wtAAV8 ou vetor mutante do capsídeo em um volume de 100  $\mu$ L. Todos os camundongos foram sacrificados no dia 14. Para avaliação *in vivo* da expressão da luciferase, camundongos (~20g) foram anestesiados e injetados intraperitonealmente com substrato de luciferina 200  $\mu$ L ou 15mg/mL (Perkin Elmer, Waltham, MA). Os camundongos foram fotografados cinco minutos após a administração de luciferina e fotografados através de um IVIS Xenogen/*In Vivo* Imaging System. Usou-se o software Living Image 3.0 para quantificar o sinal nas regiões de interesse descritas. Foram feitas medições nos dias 7 e 14.

## 6. Avaliação do título do vetor mutante e eficiência de transdução *in vitro*

[00197] Determinaram-se os títulos do vetor por qPCR dos genomas resistentes a DNaseI. Os iniciadores qPCR recozinham com a sequência de poliadenilação do transgene empacotado. Para avaliação *in vitro* da eficiência de transdução do vetor por expressão de luciferase, semeadas  $0,9 \times 10^5$  células Huh7/poço em uma placa de 96 poços de parede preta em DMEM completo (10% de soro fetal bovino, 1% de penicilina/estreptomicina). No dia seguinte, foi removido o meio e substituído por 50  $\mu$ L de vetor bruto ou purificado diluído em meio completo. Testaram-se 4 diluições em uma série de diluições de 3 vezes para cada amostra de vetor bruto. Após 48 horas, preparou-se a luciferina (Promega, Madison, WI) em meio completo a 0,3  $\mu$ g/ $\mu$ L e adicionadas às células transduzidas em um volume de 50  $\mu$ L. Os resultados foram lidos

em um luminômetro Biotek Clarity. Descobriu-se que a atividade da luciferase/GC adicionada às células alvo é constante em um amplo intervalo de GCs, mas pode se tornar saturada em MOIs elevados. Assim, inspecionaram-se os dados da série de diluição (unidades luminescentes vs GC) para linearidade, excluiu-se o ponto mais alto se a saturação for evidente e calculou-se uma Luciferase/GC média para valores na faixa linear de cada ensaio para cada variante. Isso resulta em um valor de eficiência de transdução. Os dados são normalizados para simplificar a comparação, definindo o controle de peso para um valor de 1.

## 7. Biodistribuição

[00198] Extraíu-se DNA de amostras de fígado usando o QIAamp DNA Mini Kit (Qiagen, Hilden, Alemanha) e, em seguida, analisou-se o DNA para o vetor GC por PCR em tempo real, conforme descrito anteriormente com um conjunto de iniciador / sonda projetado contra o sinal de poliadenilação RBG do cassete de transgene (Chen SJ, et al. *Hum Gene Ther Clin Dev* 2013; 24(4):154-60).

### Sequências de iniciador para mutantes AAV8

	Sequência	Descrição
SED ID NO: 56	CGACAACCGGGCAAACcagAATAGCAACTTTGCCTGG	Iniciadores mutagênicos QC para alterar AAV8 N499 para Q
SED ID NO: 57	CCAGGCAAAGTTGCTATTCTGGTTTTGCCCGTTGTCTG	Iniciadores mutagênicos QC para alterar AAV8 N499 para Q
SED ID NO: 58	GACAACCGGGCAAACgacAATAGCAACTTTGCCTG	Iniciadores mutagênicos QC para alterar AAV8 N499 para D
SED ID NO: 59	CAGGCAAAGTTGCTATTGTCTGGTTTTGCCCGTTGTCTG	Iniciadores mutagênicos QC para alterar AAV8 N499 para D
SED ID NO: 60	GGAGGCACGGCAcagACGCAGACTCTGGG	Iniciadores mutagênicos qc para alterar AAV8 N459 para Q
SED ID NO: 61	CCCAGAGTCTGCGTCTGTGCCGTGCCTCC	Iniciadores mutagênicos qc para alterar AAV8 N459 para Q
SED ID NO: 62	CAGGAGGCACGGCAgatACGCAGACTCTGG	Iniciadores mutagênicos qc para alterar AAV8 N459 para D
SED ID NO: 63	CCAGAGTCTGCGTATCTGCCGTGCCTCCTG	Iniciadores mutagênicos qc para alterar AAV8 N459 para D
SED ID NO: 64	ctctcccgatgctgctgttgagattgc	AAV8 NA263 F
SED ID NO: 65	gcaaatctccaacgcgacatcgggaggag	AAV8 NA263 R
SED ID NO: 66	cccacggcctgactagcgttgtagtgta	AAV8 NA385 F
SED ID NO: 67	taacactcaacaacgctagtcaggccgtggg	AAV8 NA385 R
SED ID NO: 68	ggattagccaatgaattcttgcaattcagatggtattggctc	AAV8 NA514 F

SED ID NO: 69	ggaccaaataccatctgaatgcaagaaattcattggctaacc	AAV8 NA514 R
SED ID NO: 70	tttgccaaaatcaggatcgcttactgggaaaaaacg	AAV8 NA540 F
SED ID NO: 71	cgttttttcccagtaacgcgatcctgattttggcaaa	AAV8 NA540 R
SED ID NO: 72	ggacccttcaacgcactcgcacaagggg	AAV8 NA57 F
SED ID NO: 73	cccctgtcagtgctgtgaaggtcc	AAV8 NA57 R
SED ID NO: 74	tggctctcccgatgtgctgttgagattgcttg	AAV8 NS263 F
SED ID NO: 75	caagcaaatctccaacagcacatcgggaggagcca	AAV8 NS263 R
SED ID NO: 76	cccacggcctgactactgttggaggttagg	AAV8 NS385 F
SED ID NO: 77	cctaactcaacaacagtagtcagccgtggg	AAV8 NS385 R
SED ID NO: 78	ttagccaatgaattctgctattcagatggtattggcccagcag	AAV8 NS514 F
SED ID NO: 79	ctgctgggaccaaataccatctgaaatagcagaaattcattggctaa	AAV8 NS514 R
SED ID NO: 80	ttgtttgcaaaaatcaggatcgttactgggaaaaaacgctc	AAV8 NS540 F
SED ID NO: 81	gagcgtttttcccagtaacagcatcctgattttggcaaaaca	AAV8 NS540 R
SED ID NO: 82	ctccccctgtcaggtgttgaaggtccgag	AAV8 NS57 F
SED ID NO: 83	ctcggaccctcaacagcctcgacaagggggag	AAV8 NS57 R
SED ID NO: 84	cagcactcatcaacGACaactgggattccg	Iniciador QC para AAV8 N305D
SED ID NO: 85	ggaggcacggcaGATacgcagactctgg	Iniciador QC para AAV8 N459D
SED ID NO: 86	gacaaccgggcaaacGACaatagcaactttgctcg	Iniciador QC para AAV8 N499D
SED ID NO: 87	ccatctgaatggaagaGATcattgctaatcctggcatc	Iniciador QC para AAV8 N517D
SED ID NO: 88	cgaagcccaagccGACcagcaaaagcagg	Iniciador QC para AAV8 N35D
SED ID NO: 89	gtacctcggatGACcagccgagccc	Iniciador QC para AAV8 N94D
SED ID NO: 90	gatgctgagaaccggcGACaactccagttacttac	Iniciador QC para AAV8 N410D
SED ID NO: 91	cagactctgggctcagcGATggtggcctaatacaatg	Iniciador QC para AAV8 Q467D
SED ID NO: 92	ccaatcaggcaaaGACtgctcaggagc	Iniciador QC para AAV8 N479D
SED ID NO: 93	cacggacggcGACttccaccgtctc	Iniciador QC para AAV8 N630D
SED ID NO: 94	gatcctgatcaagGACacgcctgtacctgcg	Iniciador QC para AAV8 N653D
SED ID NO: 95	gtacctcggaccctcCAGggactcgacaaggg	Iniciador QC para AAV8 N57Q
SED ID NO: 96	ctacaagcaaatctccCAGgggacatcgggaggagc	Iniciador QC para AAV8 N263Q
SED ID NO: 97	gctacctaactcaacCAGggtagtcaggccgtgg	Iniciador QC para AAV8 N385Q
SED ID NO: 98	gctgggaccaaataccatctgCAGggaagaattcattggc	Iniciador QC para AAV8 N514Q
SED ID NO: 99	ggagcgtttttcccagtcagggatcctgattttggc	Iniciador QC para AAV8 N540Q
SED ID NO: 100	cggaatccccagtgtcgttgatgagtcgctg	Iniciador QC para AAV8 N305D
SED ID NO: 101	ccagagtctcgtatctgccgtgctcc	Iniciador QC para AAV8 N459D
SED ID NO: 102	caggcaaaagttgctattgctgtttgcccggtgtc	Iniciador QC para AAV8 N499D
SED ID NO: 103	gatgccaggattagccaatgaatctctccattcagatgg	Iniciador QC para AAV8 N517D
SED ID NO: 104	cctgctttgctgctggccttgggcttcg	Iniciador QC para AAV8 N35D
SED ID NO: 105	ggcgtcggcgtgtcataccgcaggtac	Iniciador QC para AAV8 N94D
SED ID NO: 106	gtaagtaactggaagttgctcgggtctcagcatc	Iniciador QC para AAV8 N410D
SED ID NO: 107	cattgtattagcccaccatcgtgaaagccagagctg	Iniciador QC para AAV8 Q467D
SED ID NO: 108	gtcctggcagccagtccttgcctgattgg	Iniciador QC para AAV8 N479D
SED ID NO: 109	gagacgggtggaagtcgcccgtccgtg	Iniciador QC para AAV8 N630D
SED ID NO: 110	cgcaggtacaggcgtgctctgatcaggatc	Iniciador QC para AAV8 N653D
SED ID NO: 111	gcagcactcatcaacGACaactgggattccggc	iniciador alternativo mais longo para fazer AAV8 N305D por mutagênese qc

SED ID NO: 112	GCCGGAATCCCCAGTTGTCGTTGATGAGTCGCTGC	iniciador alternativo mais longo para fazer AAV8 N305D por mutagênese qc
SED ID NO: 113	cagcgactcatcaacGACCaactggggattccggc	iniciador alternativo mais longo para fazer AAV8 N305D por mutagênese qc
SED ID NO: 114	GCCGGAATCCCCAGTTGTCGTTGATGAGTCGCTG	iniciador alternativo mais longo para fazer AAV8 N305D por mutagênese qc
SED ID NO: 115	gcgactcatcaacGACCaactggggattccg	iniciador alternativo mais curto para fazer AAV8 N305D por mutagênese qc
SED ID NO: 116	CGGAATCCCCAGTTGTCGTTGATGAGTCGC	iniciador alternativo mais curto para fazer AAV8 N305D por mutagênese qc
SED ID NO: 117	ctctgggcttcagcGAAggtgggcctaatac	iniciador QC mutagênico para fazer AAV8 Q467E
SED ID NO: 118	GTATTAGGCCACCTTCGCTGAAGCCCAGAG	iniciador QC mutagênico para fazer AAV8 Q467E
SED ID NO: 119	cctcggacccttcGACggactcgacaagg	Iniciador QC para AAV8 N57D
SED ID NO: 120	tacaagcaaatctccGACgggacatcggaggag	Iniciador QC para AAV8 N263D
SED ID NO: 121	ctacctaactcaacGACggtagtcaggccgtg	Iniciador QC para AAV8 N385D
SED ID NO: 122	ctgggaccaaataccatctgGATggaagaaattcattggctaatac	Iniciador QC para AAV8 N514D
SED ID NO: 123	gagcgtttttcccagtGACgggatcctgattttggc	Iniciador QC para AAV8 N540D
SED ID NO: 124	cctgtcagtcctcgaagggtccgagg	Iniciador QC para AAV8 N57D
SED ID NO: 125	ctcctcccgatgtcccgcggagattgcttga	Iniciador QC para AAV8 N263D
SED ID NO: 126	cacggcctgactaccgtcgttgagtgttaggtag	Iniciador QC para AAV8 N385D
SED ID NO: 127	gattagccaatgaatttctccatccagatggtattgggtccag	Iniciador QC para AAV8 N514D
SED ID NO: 128	gccaaaaatcaggatcccgtcactgggaaaaaacgctc	Iniciador QC para AAV8 N540D

## B. Resultados

[00199] AAV8 mostra heterogeneidade de carga substancial em suas proteínas do capsídeo

[00200] Para avaliar qualitativamente a presença de modificações pós-tradução no capsídeo do vetor AAV8 que poderiam afetar o desempenho do vetor, analisamos a proteína do capsídeo total AAV8 purificada por gradiente de iodixanol por eletroforese em gel 1D e 2D. Em um gel de SDS redutor de dodecil sulfato de sódio 1D, VP1, VP2 e VP3 resolvidos como bandas únicas nos pesos moleculares apropriados (FIG. 1B) (Rose JA, et al. *J Virol* 1971; 8(5):766-70). Quando avaliada adicionalmente por eletroforese em gel 2D, que separa as proteínas com base na carga (FIG. 1C), cada uma das proteínas do capsídeo adicionalmente resolvida como uma série de pontos distintos com diferentes pontos isoelétricos (pIs) variando de pH 6,3 a > 7,0 dependendo da isoforma de VP (FIG. 1D). Pontos individuais em cada VP foram sepa-

rados por intervalos distintos de unidades de 0,1 pI, medidos como migração em relação aos padrões de ponto isoelétrico interno da isoforma de anidrase carbônica, sugerindo uma única mudança de carga de resíduo. A presença dessas isoformas sugere que cada VP tem o potencial de sofrer muitas modificações, fazendo com que migrem de forma diferente sob focalização isoelétrica.

[00201] A desamidação, na qual uma fração de (tipicamente asparagina) grupos amida da cadeia lateral são convertidos em ácido carboxílico (FIG. 1A), é uma fonte comum de heterogeneidade de carga em preparações de proteína. Para determinar se a desamidação poderia ser responsável pela população distinta de isoformas de carga de VP, mutamos dois resíduos de asparagina AAV8 individualmente para aspartato. Essas mutações do capsídeo devem mudar a carga em uma quantidade equivalente à desamidação completa de um único resíduo de asparagina adicional. A análise de gel 2D dos mutantes indica que os pontos principais para VP1, VP2 e VP3 mudaram um local de ponto mais ácido (0,1 unidades de pH) do que os pontos equivalentes no tipo selvagem (wt) AAV8 (FIG. 1E - FIG. 1G). A magnitude dessa mudança é equivalente ao espaçamento observado entre as isoformas de carga VP em peso. Assim, a padronização de gel 2D das proteínas do capsídeo de AAV é consistente com a desamidação de múltiplos sítios.

[00202] A desamidação espontânea ocorre no capsídeo do vetor AAV8

[00203] Para identificar as modificações responsáveis pelo padrão de coloração distinto para cada proteína do capsídeo, analisou-se um painel de vetores AAV8 por espectrometria de massa. A cobertura da proteína do capsídeo AAV8 foi em média > 95% da sequência VP1 total (dados não mostrados). Detectou-se extensa desamidação de um subconjunto de resíduos de asparagina e glutamina por espectroscopia de massa, que mostrou um aumento de ~ 1 Da na massa observada dos



peptídeos individuais em comparação com os valores preditos com base na sequência codificada pelo DNA; observou-se este padrão de desamidação em todas as preparações de vetores AAV8 (FIG. 2A - FIG. 2D).

[00204] Para avaliar a heterogeneidade global de desamidação entre os métodos de purificação comumente usados e para examinar a desamidação nas regiões únicas VP1 e VP2, selecionaram-se nove lotes de AAV8 produzidos por transfecção tripla em células 293 e purificou-se por gradiente de cloreto de cézio, gradiente de iodixanol ou cromatografia de afinidade. Os vetores também variaram em relação a promotores e cassetes de transgene. Para determinar se a presença do genoma do vetor teve um impacto na desamidação, também avaliamos uma preparação AAV8 produzida por transfecção tripla em células 293 na ausência de plasmídeo cis (produzindo apenas capsídeos vazios) e purificado por gradiente de iodixanol.

[00205] Um amplo intervalo de desamidação estava presente em resíduos de asparagina e glutamina do capsídeo de AAV8, variando de indetectável a mais de 99% dos aminoácidos individuais sendo desamidados (FIG. 2E). Os níveis mais altos de desamidação (> 75%) ocorreram em resíduos de asparagina onde o resíduo N+1 era glicina (isto é, pares NG) (Tabela 1). Detectaram-se níveis mais baixos de desamidação (isto é, até 17%) em resíduos de asparagina adicionais onde o N+1 não era glicina. A desamidação média para asparaginas foi amplamente consistente entre as preparações. Também detectou-se desamidação nos resíduos de glutamina, mas em uma frequência mais baixa do que nas asparaginas; a maior porcentagem que observamos foi <2% em Q467 (FIG. 7). Esta observação foi inconsistente entre as preparações (dados não mostrados). Observaram-se as maiores diferenças de preparação para preparação no resíduo N499 (o resíduo N+1 é asparagina), com valores variando de <1% a mais de 50% de desamidação.

Independentemente disso, as variações que observamos na desamidação entre as preparações do vetor não parecem estar relacionadas ao método de purificação, identidade do transgene ou a presença do genoma do vetor, sugerindo que esses fatores não afetam as taxas de desamidação.

Tabela 1: Características dos resíduos desamidados AAV8 de interesse. Asteriscos representam resíduos selecionados para análise posterior.

	resíduo N+1	Topologia estrutural	Motivo estrutural	% Desamidação média	Fator de temperatura (Å <sup>2</sup> )
N35	Q	N/A	N/A	1	N/A
N57	G	N/A	N/A	80	N/A
N94	H	N/A	N/A	7	N/A
N254*	N	Superfície exposta	Não atribuído	9	35
N255*	H	Superfície exposta	Não atribuído	N/A	42
N263	G	Superfície exposta	HVR I	99	51
N305	N	Oculto	Hélice alfa	8	33
N385	G	Superfície exposta	HVR III	88	41
N410	N	Oculto	Não atribuído	3	33
N459	T	Superfície exposta	HVR IV	7	65
N499	N	Superfície exposta	HVR V	17	45
N514*	G	Superfície exposta	HVR V	84	36
N517*	S	Superfície exposta	HVR V	4	40
N540*	G	Oculto	HVR VII	79	40
N630*	F	Oculto	Não atribuído	1	32
N653	T	Superfície exposta	laço HI	1	35

[00206] Em seguida, executou-se uma série de experimentos para determinar se o manuseio da amostra contribuiu para os níveis observados de desamidação em AAV8. Temperatura extrema (70°C por 7 dias) ou pH (pH 2 ou pH 10 por 7 dias) não induziu significativamente desamidação adicional no capsídeo AAV8 (FIG. 4A e FIG. 4B). Dada essa resistência, raciocinou-se que era improvável que a desamidação observada ocorresse apenas na fase de purificação, que foi mais curta e relativamente leve em comparação. Tentou-se realizar uma análise de espectrometria de massa em vetor não purificado para determinar a extensão da desamidação antes e depois da purificação, mas não tivemos

sucesso. Da mesma forma, os controles de água pesada indicam que o processamento específico para o nosso fluxo de trabalho de espectrometria de massa não contribui com eventos de desamidação adicionais (FIG. 4C).

[00207] Para validar nosso fluxo de trabalho de espectrometria de massa, examinaram-se duas proteínas recombinantes que foram avaliadas anteriormente para desamidação; as descobertas (FIG. 5A e FIG. 5B) concordam com os resultados publicados [Henderson, LE, Henriks-son, D e Nyman, PO (1976). Primary structure of human carbonic anhy-drase C. *The Journal of biological chemistry* 251: 5457-5463 and Car-valho, RN, Solstad, T, Bjorgo, E, Barroso, JF, and Flatmark, T (2003). Deamidations in recombinant human phenylalanine hydroxylase. Identifi-cation of labile asparagine residues and functional characterization of Asn --> Asp mutant forms. *The Journal of biological chemistry* 278: 15142-1515]. Além disso, contratou-se uma instituição secundária para avaliar nossos dados brutos do AAV8. Esta análise independente iden-tificou os mesmos sítios como desamidados, com variação mínima na extensão da modificação em cada sítio atribuível a variações de sof-tware para software na detecção de pico e cálculo de área (FIG. 6).

[00208] A topologia estrutural, o fator de temperatura e a identidade do aminoácido N+1 contribuem para a frequência de desamidação

[00209] Como a estrutura do AAV8 foi resolvida e publicada (identifi-cador PDB: 2QA0) (Nam HJ, et al. *J Virol* 2011; 85(22):11791-99), exa-minou-se a seguir a estrutura do capsídeo do AAV8 para evidências de condições favoráveis para desamidação não enzimática e correlacionar a porcentagem de desamidação com características estruturais estabe-lecidas (Nam HJ, et al. *J Virol* 2007; 81(22):12260-71). Foram concen-trados resíduos de asparagina exclusivamente, pois os fatores que in-fluenciam a desamidação da asparagina são mais bem caracterizados na literatura e os eventos de desamidação da asparagina são muito

mais comuns do que os eventos de desamidação da glutamina (Robinson, NE e Robinson, AB (2001). Molecular clocks. *Proc Natl Acad Sci USA* 98: 944-949). Também determinou-se o fator de temperatura (ou B) para cada um desses resíduos da estrutura de cristal de AAV8; o fator de temperatura é uma medida do deslocamento de um átomo de sua posição média, com valores mais altos indicando um deslocamento maior, vibração térmica mais alta e, portanto, flexibilidade aumentada (Parthasarathy S and Murphy MR. *Protein Science: A Publication of the Protein Society* 1997; 6:2561-7). A maioria das asparaginas de interesse estava localizada em ou perto das HVRs expostas à superfície (Tabela 1), que são estruturalmente favoráveis para desamidação e fornecem um ambiente exposto a solvente (Govindasamy L, et al. *J Virol* 2013; 87(20):11187-99). Descobriu-se que os resíduos localizados nessas regiões de alça flexível foram, em média, mais frequentemente desamidados do que os resíduos em regiões menos flexíveis, como fitas beta e hélices alfa. Por exemplo, o resíduo de NG na posição N263 é parte de HVR I, tem um fator de alta temperatura e foi > 98% desamidado em média (FIG. 7A e FIG. 6, Tabela 1). N514, que foi desamidado ~85% do tempo (FIG. 3 e FIG. 6, Tabela 1), também está em uma HVR (HVR V) com uma glicina N+1; no entanto, o fator de temperatura local é relativamente baixo em comparação com o de N263 devido à sua interação com resíduos em outros monômeros VP no eixo triplo. Resíduos +1 menos favoráveis e fatores de temperatura local mais baixos correlacionados com desamidação mais baixa, mesmo para resíduos HVR. Por exemplo, N517 estava em média apenas 4% desamidado (Tabela 1); este resíduo tem um fator de temperatura equivalente ao N514 altamente desamidado, mas seu resíduo N+1 é uma serina, diminuindo a probabilidade de eventos de desamidação devido ao impedimento estérico. Isso demonstra que uma série de fatores determinam cumulativamente a extensão da desamidação em uma determinada posição do

capsídeo, embora a identidade do resíduo +1 seja aparentemente o fator mais influente.

[00210] Para testar o papel do resíduo +1 na desamidação da asparagina, geramos vetores mutantes nos quais os sítios AAV8 NG foram individualmente mutados na posição +1 para alanina ou serina. Os estudos de peptídeos modelo indicam que os peptídeos NG desamidam com meia-vida tão curta quanto 1 dia, enquanto os peptídeos NA ou NS normalmente desamidam 25 ou 16 vezes mais lentamente, respectivamente (Robinson NE e Robinson AB. *Proc Natl Acad Sci USA*. 2001; 98(8):4367-72). A análise de espectrometria de massa dos mutantes do vetor confirmou o papel central do sítio +1 na determinação da extensão da desamidação do vetor. Os sítios NG neste conjunto (> 80% de desamidação em peso) mostraram estabilização seletiva da asparagina adjacente quando o sítio +1 foi alterado para alanina (<5% de desamidação) ou serina (14% de desamidação) (Tabela 2).

Tabela 2: Extensão da desamidação (%) em cinco sítios AAV8 NG em wt e seis mutantes de sítio +1

posição/variante	WT (média)	G58S	G58A	G264A	G386S	G386A	G515A
N57	81,8	8,4	1,9	89,7	89,7	91,6	93,6
N263	99,3	98,2	98,9	4,8	100,0	94,5	97,2
N385	89,1	96,3	94,8	97,1	13,5	2,5	97,0
N514	85,2	100,0	98,0	98,8	100,0	100,0	2,2
N540	84,5	95,0	92,6	97,9	96,9	86,1	89,5

[00211] Resíduos que foram pelo menos parcialmente ocultos e não prontamente expostos ao solvente e/ou foram localizados em regiões de baixa flexibilidade local no capsídeo AAV8 intacta e totalmente montada teve uma frequência mais baixa de desamidação em comparação com aqueles localizados em um ambiente mais favorável (Tabela 1). Apesar disso, alguns dos resíduos em condições desfavoráveis foram desamidados. Por exemplo, o N630 está pelo menos parcialmente oculto, mas ainda tinha um grau detectável de desamidação. Para este resíduo, a presença de fenilalanina como o resíduo N+1 sugere que esta

região poderia ser um novo sítio de clivagem autoproteolítica não enzimática dentro da proteína AAV8 VP3.

[00212] A modelagem estrutural de AAV8 VP3 confirma eventos de desamidação

[00213] Para fornecer evidências diretas de desamidação no contexto de um capsídeo montado, avaliamos a estrutura cristalina de AAV8 (Nam HJ, et al. *J Virol* 2011; 85(22):11791-9). A resolução da estrutura cristalina disponível (isto é, 2,7Å) deste sorotipo não é alta o suficiente para identificar os átomos terminais nos grupos R e, portanto, é insuficiente para distinguir diretamente entre asparagina, resíduos de ácido aspártico e isoaspártico. Outros aspectos da estrutura do isômero do ácido aspártico que se forma sob essas condições nos deram a oportunidade de determinar a desamidação da estrutura de 2,7Å. Esta análise foi baseada em duas suposições: 1) O produto predominante da desamidação espontânea de uma asparagina é isoaspártico em vez de ácido aspártico, que é gerado em uma razão de 3:1 (Geiger T e Clarke S. *J Biol Chem* 1987; 262(2):785-94) e 2) uma asparagina ou ácido aspártico pode ser diferenciado de um ácido isoaspártico porque o mapa de densidade eletrônica correspondente ao grupo R do ácido isoaspártico é menor em comprimento. Este grupo R mais curto é criado quando o carbono beta do grupo R do ácido isoaspártico é perdido quando incorporado na cadeia principal da estrutura da proteína do capsídeo AAV8 VP3 após a resolução do intermediário succinimidil durante a reação de desamidação.

[00214] Primeiro refinou-se a própria estrutura de AAV8, gerando uma densidade de elétrons de capsídeo de AAV8 que não foi influenciada pela sequência conhecida de AAV8 VP3. Em seguida, examinou-se a estrutura de cristal AAV8 refinada para evidências de desamidação com base na presença de um grupo R mais curto associado ao ácido

isoaspártico (FIG. 3A - FIG. 3E). O mapa de densidade de elétrons confirmou um grupo R mais curto para os resíduos de glicina N+1 altamente desamidados nas posições 263 (FIG. 3C), 385 (não mostrado), 514 (FIG. 3D) e 540 (FIG. 3E) quando comparado com a asparagina em 410 que não teve desamidação detectada por espectrometria de massa (FIG. 3B). A desamidação indicada pelo mapa de densidade de elétrons é, portanto, consistente com os dados gerados por espectrometria de massa nesses sítios com > 75% de desamidação. Os modelos de ácido isoaspártico resultantes foram comparáveis aos resíduos de ácido isoaspártico observados nas estruturas cristalinas de outras proteínas desamidadas conhecidas, apoiando a validade de nossa análise de AAV8 (Rao FV, et al. *Chem Biol.* 2005; 12(1):65-76; Noguchi S, et al. *Biochemistry* 1995; 34(47):15583-91; Esposito L, et al. *J Mol Biol* 2000; 297(3):713-32). Esta análise estrutural serve como uma confirmação independente dos fenômenos de desamidação observados ao analisar o capsídeo AAV8 via espectrometria de massa.

[00215] A desamidação do capsídeo de AAV não é específica do sorotipo

[00216] Foram investigados os sorotipos além do AAV8 para evidências de desamidação do capsídeo. Examinaram-se as preparações do vetor AAV9 usando eletroforese em gel 2D (FIG. 11A) e espectrometria de massa (FIG. 11B), incluindo controles para potenciais efeitos de processamento do vetor (FIG. 11D - FIG. 11F). O padrão e a extensão da desamidação de AAV9 foram semelhantes aos de AAV8. Todos os quatro sítios AAV9 NG foram > 85% desamidados; 13 sítios não NG foram desamidados em menor grau, com alguns sítios mostrando alta variabilidade lote a lote em % de desamidação. Em seguida, aplicou-se o fluxo de trabalho de análise estrutural e reajustaram-se os dados cristalográficos AAV9 existentes (FIG. 11C, Tabela 3). Tal como acontece com AAV8, o ácido isoaspártico se encaixa melhor na densidade de elétrons

de vários sítios NG na estrutura cristalina AAV9. Estenderam-se a análise de gel 2D (dados não mostrados) e espectrometria de massa (resumida na Tabela 4) para cinco sorotipos evolutivamente diversos adicionais (rh32.33, AAV7, AAV5, AAV4, AAV3B e AAV1). Todos os capsídeos examinados contêm um padrão e extensão semelhantes de desamidação, indicando que esta modificação é difundida em vetores AAV clinicamente relevantes e é determinada por sequência primária subjacente e fatores estruturais semelhantes.

Tabela 3. Características dos resíduos desamidados de AAV9 de interesse.

[00217] Os resíduos de asparagina conservados com resíduos N+1 homólogos (em comparação com AAV8) são indicados em itálico (determinado pelo alinhamento das sequências de aminoácidos de comprimento total de AAV8 e AAV9 VP1).

	resíduo N+1	Topologia estrutural	Motivo estrutural	% Desamidação média	Fator de temperatura (Å <sup>2</sup> )
<i>N57</i>	G	N/A	N/A	97	N/A
<i>N94</i>	H	N/A	N/A	5	N/A
<i>N253</i>	N	Superfície exposta	Não atribuído	9	41
<i>N254</i>	H	Superfície exposta	Não atribuído	2	50
<i>N270</i>	D	Superfície exposta	HVR I	11	65
<i>N304</i>	N	Oculto	Hélice alfa	23	35
<i>N329</i>	G	Superfície exposta	HVR II	94	89
<i>N409</i>	N	Oculto	Não atribuído	9	36
<i>N452</i>	G	Superfície exposta	HVR IV	98	64
<i>N477</i>	Y	Oculto	Não atribuído	2	33
<i>N512</i>	G	Superfície exposta	HVR V	89	48
<i>N515</i>	S	Superfície exposta	HVR V	3	47
<i>N651</i>	T	Oculto	laço HI	1	38
<i>N663</i>	K	Superfície exposta	laço HI	4	49
<i>N668</i>	S	Superfície exposta	laço HI	13	52
<i>N704</i>	Y	Superfície exposta	HVR IX	5	68
<i>N709</i>	N	Superfície exposta	HVR IX	5	55

Tabela 4. Extensão da desamidação observada para diversos sorotipos

sorotipo	preparações de vetores analisadas	% média de cobertura de sequência por MS	# de NGs	% de desamidação média de NG	# de sítios não NG observados	% de desamidação não NG média
AAV1	3	91,4	4	95,6	19	12,9



AAV3B	1	89,8	4	97,0	9	9,4
AAV4	3	84,7	4	96,2	15	15,3
AAV5	1	88,7	3	88,7	11	15,3
AAV7	1	90,9	4	92,1	9	13
AAV8	21	93,4	5	90,5	37	7,4
AAV9	7	90,2	4	95,5	26	5,3
rh32.33	1	100	3	97,4	14	16,2

[00218] Os eventos de desamidação podem afetar a montagem do capsídeo e a eficiência da transdução

[00219] Uma abordagem para testar o impacto funcional da desamidação é substituir asparagina por aspartato por mutação genética. Gerou-se um vetor mutante de aspartato que codifica um repórter de luciferase para cada asparagina AAV8 desamidada por transfecção tripla em pequena escala de células 293 e titularam-se os vetores por qPCR de cópias de genoma resistentes a DNaseI (FIG. 8A). As mutações raramente afetaram a montagem do capsídeo em relação ao wtAAV8, e os efeitos foram limitados a sítios não NG, principalmente ocultos com baixa desamidação geral no vetor wt. Em seguida, avaliou-se o painel de mutação para a eficiência de transdução *in vitro* de células Huh7 derivadas de fígado humano (FIG. 8B). Vários mutantes mostraram eficiência de transdução prejudicada, com as posições N57, N94, N263, N305, Q467, N479 e N653 exibindo perda de transdução > 10 vezes. Observou-se um número semelhante de sítios sensíveis para AAV9 (FIG. 11G e FIG. 11H). Como tipicamente apenas uma fração de resíduos em uma determinada posição é desamidada endogenamente, esta abordagem tem o potencial de superestimar a perda funcional para proteínas, como capsídeos, onde a unidade funcional é um conjunto homomérico; a modificação endógena em um sítio do capsídeo pode ser compensada por uma subunidade vizinha com um resíduo intacto. No entanto, concluímos que o método poderia ajudar a priorizar resíduos desamidados para monitoramento futuro durante a fabricação ou estabilização mutacional. Dados funcionais de populações de vetores de desa-

midação endógena serão necessários para colocar esses dados de mutagênese de perda de função no contexto adequado.

[00220] A perda de atividade vetorial ao longo do tempo está correlacionada com a desamidação progressiva

[00221] Dada a meia-vida aparentemente curta da desamidação NG, concluiu-se que as amostras de vetores com idade diferente de 1 dia poderiam mostrar perfis de desamidação distintos, proporcionando assim uma oportunidade de correlacionar a desamidação endógena ao funcionamento. O protocolo de preparação de vetor em grande escala exige a transfecção tripla de células 293, seguida por 5 dias de incubação para produção de vetor e 1-2 dias para purificação de vetor. Para aproximar este processo, preparamos transfecções triplas em escala média (placas de cultura de células de 10 x 15 cm cada) de células 293 com AAV8 wt. Coletou-se o vetor (placas de cultura de células de 2 x 15 cm/dia) em intervalos de 1 dia por 5 dias, preservando os pontos de tempo até o final do período de 5 dias por vetor de congelamento a -80C. Em seguida, avaliou-se o título do vetor bruto e a eficiência de transdução *in vitro*, conforme descrito acima. Como esperado, o número de cópias do genoma resistentes a DNaseI montadas aumentou ao longo do tempo (FIG. 9A). Em seguida, processou-se rapidamente o vetor bruto para pontos temporais iniciais (dia 1 e 2) e tardios (dia 5) por purificação de afinidade e mediu-se a eficiência de transdução *in vitro* de células huh7. A eficiência de transdução relativa do vetor caiu progressivamente ao longo do tempo (FIG. 9B). Em termos de expressão do transgene por GC adicionado às células-alvo, o vetor do dia 5 foi apenas 40% tão eficiente quanto o material do dia 1. Esta queda de atividade também foi observada para o material bruto, indicando uma mudança na composição molecular antes da purificação (FIG.). Observou-se uma tendência semelhante na perda de atividade para AAV9 ao longo de 5 dias, com redução de aproximadamente 40% na potência do

vetor (FIG. 11I - FIG. 11K).

[00222] Em seguida, mediu-se a desamidação das amostras do curso do tempo por espectrometria de massa. A desamidação do sítio NG progrediu substancialmente ao longo de cada intervalo, com uma média de 25% de desamidação no dia 1 e > 60% dos sítios convertidos no dia 5 (FIG. 9C). A desamidação do sítio não NG geralmente progrediu ao longo de 5 dias, embora em níveis muito mais baixos e com menos consistência entre os dias 2 e 5 (FIG. 9D). Os dados correlacionam a desamidação endógena do vetor a um decaimento de ponto de tempo precoce em atividade específica e destacam uma oportunidade potencial para capturar o vetor mais ativo encurtando o ciclo de produção ou encontrando mutações do capsídeo que estabilizam as asparaginas.

[00223] Notou-se que o material usado para análise de espectrometria de massa na FIG. 2A - FIG. 2E foi de pelo menos 7 dias após a transfecção, devido a 2 dias adicionais para purificação. A maior desamidação do sítio NG nessas amostras (> 80%) indica que a desamidação provavelmente continua após o período de expressão e durante os processos de recuperação e purificação aproximadamente nas mesmas taxas até que os sítios NG sejam completamente desamidados ou a amostra do vetor seja congelada. Assim, a desamidação é amplamente determinada pela idade do vetor e não é um processo exclusivo ou causado pelo processo de recuperação e purificação. Os valores de desamidação muito mais baixos no material do dia 1 vs o material do dia 5 (ambos purificados por afinidade) sublinham este ponto.

[00224] Estabilizar asparaginas NG pode melhorar o desempenho do vetor

[00225] Dada a correlação entre a desamidação NG do vetor e a perda de eficiência de transdução, concluímos que a estabilização das amidas NG por mutagênese no sítio +1 pode melhorar a função do vetor. Produzimos vetor em pequena escala para mutantes do sítio AAV8 NG

em que cada resíduo +1 foi convertido individualmente em alanina ou serina. Mutantes +1 únicos foram bem tolerados em termos de montagem de vetor (FIG. 10A) e eficiência de transdução (FIG. 10B). Substituições de G386, localizadas perto de uma "zona morta" previamente definida na superfície do capsídeo (Aydemir F, et al. *J Virol* July 2016; 90(16):7196-204), eram defeituosas para transdução *in vitro*. A perda de função para mutantes G386 pode indicar uma preferência por uma asparagina desamidada em N385. Alternativamente, o volume adicional da cadeia lateral na posição +1 pode ter um impacto negativo na função que é independente da estabilização do grupo amida. Nenhum mutante de sítio único melhorou significativamente a transdução *in vitro*, apesar da dramática estabilização de suas asparaginas vizinhas (Tabela 2). Como as atividades de transdução *in vitro* e *in vivo* podem ser discordantes, testamos um subconjunto dos mutantes de sítio único +1 para a transdução hepática em camundongos C57BL/6. Realizamos injeção intravenosa na veia da cauda (n = 3 a 5) e examinamos a expressão da luciferase por imageamento semanal durante 2 semanas (FIG. 10C). Os dados de transdução *in vivo* e *in vitro* estavam de acordo com os erros associados de cada ensaio (ou seja, dentro do intervalo de erro). As substituições de G386 foram defeituosas para a transdução, enquanto as mutações do sítio +1 em outras posições foram amplamente toleradas, transduzindo o fígado em níveis equivalentes, mas não excedendo wtAAV8.

[00226] Como a estabilização da amida em qualquer sítio NG pode ser necessária, mas não suficiente para a restauração funcional, avaliamos a seguir as variantes do vetor com combinações de substituições de alanina no sítio +1. Nós recombinamos todos os 3 sítios AAV8 NG para os quais a alanina +1 era altamente funcional (N263, N514 e N540). Algumas combinações, incluindo o mutante triplo

G264A/G515A/G541A, montaram mal e eram disfuncionais para transdução. No entanto, ambas as combinações de pares envolvendo N263 (G246A/G515A e G264A/G541A) melhoraram a eficiência de transdução *in vitro* 2,0- e 2,6- vezes sobre wtAAV8, respectivamente) sem perda de título (FIG. 10D). Como essas mutações introduzem pelo menos duas alterações (estabilização de N-amida e substituição de cadeia lateral de resíduo +1), esses dados não ligam de forma conclusiva a desamidação de NG à perda funcional. No entanto, os dados são consistentes com o modelo estabelecido no estudo de curso de tempo, no qual a desamidação do sítio NG pode impactar a eficiência de transdução *in vitro*.

[00227] Substituições funcionais de asparagina melhoram a reprodutibilidade lote a lote na fabricação de vetores

[00228] Outro aspecto potencialmente problemático dos perfis de desamidação de vetor que relatamos é a alta variabilidade de lote a lote na desamidação em algumas posições. Para wtAAV8, esta variabilidade é mais pronunciada para N459 (desamidação observada variando de 0% a 31%) e N499 (desamidação observada variando de 0% a 53%). A variabilidade nas modificações pós-traducionais é normalmente evitada *de facto* durante o desenvolvimento de produtos biológicos, seja evitando clones que exibem essa variabilidade, monitorando e controlando cuidadosamente as cepas e condições de produção, ou por engenharia de proteína do candidato afetado.

[00229] Como não fomos capazes de determinar os fatores de produção ou processamento que contribuem para a variabilidade de desamidação de N459 e N499 (FIG. 2E), buscamos substituições de aminoácidos funcionais nessas posições. Avaliamos primeiro as preparações de vetor em pequena escala para substituições conservativas de glutamina em cada posição individualmente. Ambos N459Q e N499Q foram montados de forma eficiente no vetor e eram equivalentes à referência

wtAAV8 para eficiência de transdução *in vitro* (FIG. 7A). Em seguida, produziram-se os mutantes em grande escala e realizou-se espectrometria de massa. Consistente com nossas observações de desamidação de glutamina extremamente rara, observou-se estabilização seletiva e completa das amidas de glutamina nas posições 459 ou 499 nesses mutantes (dados não mostrados). Avaliaram-se esses lotes mutantes *in vivo* como acima para a transdução do fígado após injeção na veia da cauda em camundongos C57BL/6 (FIG. 7B e FIG. 7C). O lote do vetor wtAAV8 usado como um controle nesta experiência foi desamidado 16,8% em N499, mas nenhuma desamidação foi detectada em N459 (dados não mostrados). A transdução do fígado no dia 14 para ambos os mutantes foi equivalente a wtAAV8. Estes dados demonstram o potencial para uma abordagem de engenharia de proteínas para abordar a variabilidade molecular associada à desamidação em vetores AAV fabricados.

### C. Discussão

[00230] Foi identificada e avaliada a desamidação não enzimática de resíduos de asparagina e glutamina no capsídeo AAV8 independentemente por eletroforese em gel 2D, espectrometria de massa, modelagem de proteína *de novo* e estudos funcionais *in vitro* e *in vivo*. Foi demonstrado que a desamidação ocorre em uma ampla variedade de proteínas e tem impacto significativo na atividade de produtos biológicos, incluindo terapêuticas baseadas em anticorpos (Nebija D et al. *Int J Mol Sci* 2014; 15(4):6399-411) e vacinas baseadas em peptídeos (Verma A et al. *Clin Vaccine Immunol.* 2016; 23(5):396-402). Outras proteínas virais, como a proteína VP6 do rotavírus, foram mostradas por espectrometria de massa como sofrendo eventos de desamidação (Emslie KR et al. *Funct Integr Genomics* 2000; 1(1):12-24).

[00231] O contexto em que essas desamidações ocorreram em

AAV8 sugere que elas são o resultado de eventos não enzimáticos espontâneos. Os resíduos de asparagina são conhecidos por serem mais extensivamente desamidados do que os resíduos de glutamina; o aminoácido a jusante da asparagina influencia substancialmente a taxa de desamidação com um N+1 de glicina (isto é, NG) sendo o mais eficientemente desamidado. Observou-se a confirmação notável do papel do aminoácido N+1 na desamidação de capsídeos de AAV em que cada NG presente em VP1 foi desamidado em níveis > 75%, enquanto a desamidação nunca foi consistentemente > 20% em qualquer uma das outras asparaginas ou glutaminas no capsídeo. Praticamente todos os motivos NG nos capsídeos AAV8 e AAV9 (ou seja, 7/9) também estavam presentes na superfície do capsídeo contido nas regiões HVR que estão associadas a altas taxas de flexibilidade conformacional e vibração térmica. Isso é consistente com relatórios anteriores de motivos NG de outras proteínas que estão localizadas em regiões onde a flexibilidade pode ser necessária para a função adequada da proteína e não em estruturas mais ordenadas, como hélices alfa ou folhas beta (Yan BX and Sun YQ *J Biol Chem* 1997; 272(6):3190-4). A preferência de motivos NG em HVRs expostos à superfície aumenta ainda mais a taxa de desamidação, proporcionando acessibilidade ao solvente e flexibilidade conformacional, facilitando assim a formação do intermediário succinimidil. Como previsto, ambientes menos favoráveis levam a taxas de desamidação muito mais baixas.

[00232] Uma questão importante em relação à biologia do AAV e seu uso como vetor são as consequências funcionais dessas desamidações. A mutagênese do DNA do capsídeo para converter uma asparagina em ácido aspártico permite uma avaliação dos capsídeos nos quais todos os aminoácidos em um determinado sítio são representados como ácidos aspárticos. No entanto, não existe uma estratégia fácil para usar a mutagênese para prevenir desamidações além da mutação potencial

do resíduo N+1, que é confundido pelas consequências diretas da mutação do segundo sítio. Estudou-se um número limitado de variantes nas quais o resíduo de asparagina foi convertido em ácido aspártico por mutagênese. A análise funcional incluiu a montagem do capsídeo e a transdução *in vitro* e *in vivo*. Os efeitos mais substanciais da mutagênese na função do vetor foram aqueles envolvendo asparaginas que foram desamidadas incompletamente no início do estudo e não foram expostas à superfície. O que foi surpreendente, entretanto, foi que a mutagênese da asparagina altamente desamidada em 514 em um ácido aspártico teve algum efeito na função. Este resultado sugere que a presença de quantidades residuais da amida correspondente pode influenciar a função. Isso pode ser devido em parte à presença de interações de ligações de hidrogênio entre N514 e D531 de outro monômero VP3 relacionado a três vezes (identificado na estrutura de cristal wtAAV8) que são perdidas após a conversão deste resíduo em ácido aspártico após desamidação.

[00233] Uma melhor compreensão dos fatores que influenciam a extensão da desamidação em vetores AAV é importante ao avaliar o impacto dessas desamidações no desenvolvimento de novas terapêuticas. A incubação de vetores sob condições extremas, conhecida por acelerar marcadamente a cinética de desamidação, teve pouco efeito. Juntamente com os estudos de incorporação de isótopos, este resultado sugere que a desamidação ocorre durante a montagem do capsídeo e não é um artefato de processamento de vetor ou análise de espectrometria de massa. É improvável que desamidações em sítios NG tenham um impacto significativo no desempenho do vetor, uma vez que a reação foi virtualmente completa em todas as amostras que foram avaliadas. No entanto, os estudos funcionais iniciais sugerem que quantidades residuais de asparaginas não desamidadas podem contribuir para a função. Preocupados com os sítios onde a desamidação foi menos



completa, o que na maioria dos casos também foi associado à variação de amostra para amostra. Um exemplo é a asparagina na posição 499 que apresentou desamidação variando de 0% a 53% com média de 17%. É possível que diferenças sutis nas condições de produção do vetor contribuam para essa heterogeneidade. A notável semelhança na desamidação em AAV8 e AAV9 sugere que esta é uma propriedade de toda esta família de vírus.

[00234] Em resumo, descobriu-se heterogeneidade substancial na estrutura de aminoácidos primária das proteínas do capsídeo AAV8 e AAV9. Esses estudos têm um impacto potencial no desenvolvimento de AAV como vetores de várias maneiras. Em primeiro lugar, as sequências de aminoácidos reais das proteínas VP não são as previstas pelas sequências de DNA correspondentes. Em segundo lugar, aspectos do método de produção podem levar a variações na desamidação e mudanças correspondentes na função do vetor. Até que tenha um controle sobre os fatores que influenciam as taxas de desamidação em sítios não NG e uma melhor compreensão de suas consequências funcionais, pode ser necessário incluir a desamidação na caracterização de vetores de AAV de grau clínico. A eletroforese em gel 2D pode fornecer uma avaliação geral da desamidação líquida, embora a espectrometria de massa seja necessária para avaliar a desamidação em resíduos específicos.

#### EXEMPLO 2: DESAMIDAÇÃO AAV5.5.9

[00235] As novas sequências de AAV5.5.9 são fornecidas em SEQ ID NO: 9 e 10, respectivamente. Os vetores AAV5.5.9 foram avaliados para desamidação como descrito no Exemplo 1 para AAV9. Resíduos altamente desamidados são vistos em N57, N319, N442, N502.

Modificação SEQ ID NO: 10	WL2019CS
Enzima	Tripsina
% de cobertura	97,4
N35+Desamidação	7,8

Modificação SEQ ID NO: 10	WL2019CS
-N57+Desamidação	99,7
N113+Desamidação	3,6
-N204+Desamidação	13,9
N217+Desamidação	2,2
-N243+Desamidação	19,0
Q249+Desamidação	11,4
N293/294+Desamidação	37,3
N304+Desamidação	6,2
N309+Desamidação	0,7
Q311+Desamidação	0,3
-N319+Desamidação	83,9
N399/400+Desamidação	30,8
-N442+Desamidação	97,7
N467+Desamidação	2,6
N502+Desamidação	100,0
N505+Desamidação	18,6
-Q589+Desamidação	21,1
N618+Desamidação	6,6
- N641+Desamidação	8,1
N653+Desamidação	8,3
-N658+Desamidação	21,7
N694+Desamidação	0,6
-N699+Desamidação	8,6

### EXEMPLO 3: DESAMIDAÇÃO AAVrh79 (Clade E)

[00236] O AAVrh79 foi isolado do DNA extraído do tecido do intestino delgado de macaco rhesus. Foi caracterizado filogeneticamente como estando dentro do Clado E (FIG. 14A-14D). Suas sequências são fornecidas aqui, com as sequências de nucleotídeos estando na SEQ ID NO: 1 e a sequência de aminoácidos na SEQ ID NO: 2. Um alinhamento das sequências de aminoácidos de AAVrh79, AAVrh.10 e AAVhu.37 é fornecido na FIG. 14A. Um alinhamento das sequências de ácido nucleico de AAVrh79, AAVrh.10 e AAVhu.37 é fornecido na FIG. 14B-14D.

[00237] AAVrh79 tem três diferenças de aminoácidos em sua sequência primária. Enquanto AAVhu37 tem um Ala localizado na posição 67 e um Lys na posição 169 de sua sequência VP1 primária, AAVrh79 tem um ácido glutâmico (E) na posição 67 e um Arg na posição 169.

Diferenças nas sequências de DNA de VP1 entre rh.79, hu.37 e hu.40 são mostrados na FIG. 11B. Os vetores que expressam eGFP com base nas diversas variantes do clado E foram preparados e avaliados quanto à sua infecciosidade relativa de células Huh7 (FIG.11C). Camundongos C57BL/6 foram injetados com dois níveis de dosagem ( $3 \times 10^{10}$  e  $3 \times 10^{11}$  GC/camundongo) de AAV8 ou AAVrh expressando eGFP.79 vetores e a infectividade foram avaliados por microscopia de fluorescência (dados não mostrados).

[00238] Os vetores baseados em AAVrh79 foram preparados usando técnicas de produção conhecidas usando a sequência de nucleotídeos AAVrh79 para a produção do cap, tal como descrito anteriormente para vetores AAV8. Os resultados das avaliações de rendimento de produção e pureza de produção são fornecidos nas FIGSs15A-15B e FIG 16, respectivamente.

[00239] Para avaliar os níveis de expressão usando AAVrh79 contendo um gene marcador (luciferase de pirilampo), camundongos RAG KO machos com 6–8 semanas de idade foram injetados por via intramuscular com  $3 \times 10^{11}$  GC/camundongo do vetor realizado usando uma seringa Hamilton. A expressão de ffLuc foi visualizada por imageamento de bioluminescência de corpo inteiro, conforme descrito anteriormente [Greig JA, Peng H, Ohlstein J, Medina-Jaszek CA, Ahonkhai O, Mentzinger A, et al. (2014) Intramuscular Injection of AAV8 in Mice and Macaques Is Associated with Substantial Hepatic Targeting and Transgene Expression. PLoS ONE 9(11): e112268. doi.org/-10.1371/journal.pone.-0112268]. Os resultados são fornecidos nas FIGS 17A-17D.

[00240] A expressão dos vetores AVV8triple, AAVhu68, AAV9, AAV8 e AAVrh79 foi comparada após a administração intramuscular de  $10^{13}$  GC/kgAAVrh79 em macacos cynomolgus machos e fêmeas (FIG. 17E).

[00241] Os vetores que expressam um transgene segregado (201Ig IA) foram administrados intramuscularmente no músculo gastrocnêmio

de camundongos RAG KO machos (n = 5 / grupo) ( $3 \times 10^{10}$  ou  $3 \times 10^{11}$  GC/camundongo). Os resultados indicaram que AAV8triple expressa melhor após a injeção IM e na dose mais baixa testada a diferença na expressão de AAV8triple foi substancial. Quanto maior a dose mais elevada, AAVrh79 expresso em níveis comparáveis aos outros vetores testados (FIG. 17F).

[00242] Um macaco cynomolgus fêmea (RA2362) foi pré-rastreado para NAb (FIG. 18A) e injetado com AAVG2.TBG.eGFP.WPRE.bGH ( $1 \times 10^{13}$  ddGC/kg, por via intravenosa). O animal foi sacrificado 7 dias após o tratamento, e uma necropsia foi realizada para isolar o fígado e outros tecidos para análise. A expressão de GFP no fígado e baço foi avaliada no dia 7 (FIG. 18B e 18C). Os níveis de vetor detectados em vários tecidos são mostrados na FIG. 18D. Os níveis de expressão de GFP foram avaliados nos fígados de animais que receberam os vetores AAVG2 (RA2362) ou AAV8 e AAVG3 (FIG. 18E - FIG. 1J). FIG. 18K mostra os níveis de vetor detectados em vários tecidos desses animais.

[00243] Vários vetores AAV8 e AAVrh79 foram gerados e, em alguns casos, múltiplos lotes foram produzidos. Os rendimentos destes lotes de vetores AAV9 e AAVrh79 foram comparados (FIG. 19).

[00244] Os vetores AAVrh79 foram avaliados quanto à desamidação conforme descrito no Exemplo 1 para AAV8 e AAV9. Os resultados mostram que os vetores contêm quatro aminoácidos altamente desamidados (N57, N263, N385, N514), que correspondem às asparaginas nos pares asparagina-glicina, com base na numeração de AAVrh79 (SEQ ID NO: 1). As percentagens de desamidação mais baixas são observadas de forma consistente nos resíduos N94, N254, N410.

Modificação AAVrh79 SEQ ID NO: 2	WL1781 S	WL1784S	WL1785S	WL1781S	WL1784S	WL1785S
Enzima	Tripsina	Tripsina	Tripsina	Quimotripsina	Quimotripsina	Quimotripsina
% de cobertura	89,6	93,9	92,4	91,3	88,7	89,9
N57+Desamidação	99,3	80,9	82,9	99,6	80,1	86,4
N94+Desamidação	10,4	9,6	9,9	10,5	9,4	10,0

Modificação	WL1781 S	WL1784S	WL1785S	WL1781S	WL1784S	WL1785S
Enzima	Tripsina	Tripsina	Tripsina	Quimotripsina	Quimotripsina	Quimotripsina
-N254+Desamidação	16,0	15,8	16,4	15,3	16,3	16,7
-N263+Desamidação	84,3	93,5	95,3	82,9	89,5	90,6
-N305+Desamidação	3,2	2,5	2,4	3,2	2,6	2,3
-N385+Desamidação	79,1	100,0	100,0	76,9	96,6	92,9
-N410+Desamidação	2,0	17,8	23,9	2,0	17,6	23,0
N479+Desamidação	2,0	2,0	1,9	2,0	2,0	2,0
-N514+Desamidação	100,0	97,2	97,0	97,4	94,6	98,1
-Q601+Desamidação	0,1					
N653+Desamidação	1,3	1,1	1,4	1,3	1,1	1,5
- R487+Metilação	0,1	0,2	0,1	0,1	0,2	0,1
D97+Isomerização	1,3			1,2		
S149+Fosforilação	51,9	49,0	53,2	49,6	46,6	55,8
-S153+Fosforilação	59,7	54,3	51,0	59,8	51,7	48,5
-S474+Fosforilação		7,3	4,5		7,0	4,3
-T570+Fosforilação	46,3	36,4	21,6	45,9	35,9	21,4
-S665+Fosforilação	0,5	0,3	0,4	0,4	0,3	0,4
W248+Oxidação			0,9			
W307+Oxidação	1,7	1,2	1,5	1,8	1,3	1,5
W307+Oxidação para quinure- nina	0,3		0,4			
M405+Oxidação	5,8			6,0		
M437+Oxidação	15,0	5,2	95,4	15,1	5,3	12,0
M473+Oxidação	16,7	7,4	7,9	16,0	7,6	7,6
W480+Oxidação	4,6	0,4	4,6	4,7	0,4	4,6
W480+Oxidação para quinure- nina			0,1			
W505+Oxidação	2,4	2,2	1,3	2,5	2,1	1,3
M526+Oxidação	19,0			18,6		
M544+Oxidação	30,9	20,7		31,4	20,3	
M561+Oxidação	15,6	7,7		16,0	7,5	
W621+Oxidação para quinure- nina			0,0			
M637+Oxidação	6,9	12,5	4,5	6,9	12,6	4,7
W697+Oxidação		0,5	0,6		0,5	0,6

#### EXEMPLO 4: Preparação de AAV8.2.08

[00245] Conforme discutido em WO 2017/180854 (aqui incorporado por referência), vários mutantes AAV8 foram gerados c41, c42, c46, g110, g113, g115 e g117 com mutações na região HVR.VIII. Conforme

discutido em Gurda et al, o epítipo ADK8 principal encontra-se na região HVR.VIII (aminoácidos 586 a 591 usando numeração AAV8 vp1). Esses mutantes foram testados *in vitro* para resistência a ADK8 e alguns deles foram testados *in vivo* para resistência a ADK8. Ver, por exemplo, Lochrie 2006 citado acima.

[00246] AR2.1-9 foram escolhidos aleatoriamente. AR2.25-61 foram selecionados com base na frequência. As variantes escolhidas aleatoriamente, bem como aquelas de maior frequência, estão vivas, em termos de rendimento de placa de 6 poços e transdução Huh7. FIG. 24 - mostra expressão em vários tecidos de AAV8.AR2.08 (conjunto de barras mais à esquerda).

#### EXEMPLO 5: Desamidação de AAV8.AR2.08

[00247] As novas sequências de AAV8.AR2.08 são fornecidas na SEQ ID NO: 17 e 18, respectivamente, que foram concebidas como no Exemplo 4.

A. Modificações Os vetores AAV8.AR2.08 foram produzidos e avaliados quanto às modificações, conforme descrito no Exemplo 1 para AAV8. Os resultados mostram que os vetores contêm cinco aminoácidos altamente desamidados (N57, N263, N385, N514 e N540), que correspondem às asparaginas nos pares asparagina-glicina, com base na numeração de AAV8.AR2.08 (SEQ ID NO: 18). As percentagens de desamidação mais baixas são observadas de forma consistente nos resíduos N94, N254, N410. Em contraste com AAV8, a desamidação não é observada na posição N459 (média 7% em AAV8) ou N499 (média 17% em AAV8).

Modificação de AAV8.AR2.08 SEQ ID NO: 18	WL1846CS	WL1846CS
Enzima	Tripsina	Quimotripsina
% de cobertura	97,4	92,3
N57+Desamidação	90,7	89,5
N94+Desamidação	9,0	9,3
-N254+Desamidação	11,8	11,7
-N263+Desamidação	88,6	86,3

Modificação de AAV8.AR2.08 SEQ ID NO: 18	WL1846CS	WL1846CS
Enzima	Tripsina	Quimotripsina
% de cobertura	97,4	92,3
-N305+Desamidação	5,8	5,5
-N385+Desamidação	86,1	83,3
-N514+Desamidação	100,0	99,6
-N521+Desamidação	2,0	2,1
- N540+Desamidação	78,6	80,5
N590+Desamidação	0,4	0,4
Q601+Desamidação	0,5	0,6
N653+Desamidação	0,8	0,8
N665+Desamidação	1,2	1,2
D442+Isomerização	12,0	
D584+Isomerização	1,0	
-S149+Fosforilação	95,9	15,4
-T417+Fosforilação	0,0	
-T454+Fosforilação	0,1	
-T493+Fosforilação	0,1	
S600+Fosforilação	1,1	
-T663+Fosforilação	0,0	
-W22+Oxidação	1,0	1,0
-M204+Oxidação	0,1	0,1
-M212+Oxidação	2,9	2,9
W248+Oxidação	0,7	0,7
W307+Oxidação	0,6	0,6
M405+Oxidação	0,3	0,3
M437+Oxidação	70,3	21,0
M473+Oxidação	1,7	1,7
W480+Oxidação	0,3	0,3
W505+Oxidação	0,6	0,6
M526+Oxidação	1,0	1,0
M561+Oxidação	1,0	1,0
M607+Oxidação	2,4	2,5
-W609+Oxidação	0,1	0,1
W621+Oxidação	0,8	0,8
M637+Oxidação	2,9	3,0
W697+Oxidação	0,2	0,2

B. O RNA-seq de célula única revela a localização de tecido e assinaturas transcricionais de hepatócitos transduzidos isolados de primatas não humanos após tratamento com AAV8

[00248] O sequenciamento de RNA de célula única provou ser uma técnica poderosa para caracterizar o transcriptoma celular com resolução de célula única sem precedentes. Em nosso trabalho atual, utilizamos RNA-seq de célula única para estudar a paisagem transcricional de hepatócitos primários isolados de macacos rhesus após o tratamento com um vetor AAV8 expressando GFP. A análise do transcriptoma de células GFP+ e GFP- classificadas por FACS revela a localização do tecido de células transduzidas no lóbulo hepático, bem como genes e vias regulatórias envolvidas na transdução de hepatócitos e na regulação da expressão do transgene.

[00249] Para o projeto do nosso estudo, macacos rhesus foram tratados com  $1 \times 10^{13}$  ddGC/kg AAV8.TBG.EGFP.WPRE (n=1) ou  $1 \times 10^{13}$  ddGC/kg AAV8.2.08.TBG.EGFP.WPRE (variante AAV8, n = 1). Os animais foram sacrificados 7 dias após o tratamento, e as necropsias foram realizadas para isolar o fígado de ambos os animais. Após o tratamento com colagenase e centrifugação em gradiente, os hepatócitos isolados foram classificados por FACS por expressão de transgene GFP em placas de 96 poços BD Precise™. 192 células individuais foram isoladas de cada animal (96 GFP+ e 96 GFP-) e foram subsequentemente usadas para preparar bibliotecas de RNA-seq de célula única seguindo o protocolo padrão BD Precise™. Os dados foram analisados usando os pacotes Seurat, Scran e Scater em R para determinar transcritos diferencialmente expressos entre células classificadas GFP- e GFP+ e para realizar a reconstrução espacial de células isoladas dentro do lóbulo hepático usando assinaturas de expressão transcricional estabelecidas.

[00250] Descobriu-se que o AAV8.AR2.08 tem um tropismo hepático aumentado e exibiu um aumento de 1,5 vezes na eficiência de transdução em comparação com AAV8. A análise do transcriptoma de uma única célula de hepatócitos classificados revela que as células que ex-



pressam o transgene estão uniformemente distribuídas pelo lóbulo hepático, mostrando uma ligeira preferência pela região periportal, o que também foi observado pela histopatologia. Curiosamente, uma subpopulação de células GFP- classificadas expressa o transcrito do transgene em níveis comparáveis às células GFP+ classificadas, sugerindo que essas células são de fato transduzidas e expressam mRNA do transgene, apesar da ausência dos níveis detectáveis de proteína traduzida. A comparação dos perfis de transcrição de células GFP- e GFP+ revela transcritos diferencialmente expressos envolvidos na tradução de mRNA viral, elucidando possíveis vias envolvidas na regulação da expressão de proteínas transgênicas em células transduzidas.

C. Isolamento de uma variante do vírus adenoassociado 8 com melhor transdução do fígado e maior especificidade do fígado em primatas não humanos com evolução direcionada através de um modelo de xenoenxerto de fígado humano

[00251] Para derivar variantes de AAV com melhor transdução e maior especificidade, realizou-se a evolução dirigida por AAV, usando mutagênese de saturação direcionando sítios expostos de superfície no capsídeo de AAV8, a referência para terapia genética hepática, seguida por duas rodadas de enriquecimento in vivo em modelo de camundongo de xenoenxerto de fígado humano e isolou-se uma variante AAV8 chamada AAV8.2.08. Após a injeção intravenosa em primatas não humanos em uma dose de  $1 \times 10^{13}$  cópias do genoma (GC)/kg de peso corporal, as cópias do genoma do vetor são distribuídas por AAV8.AR2.08 em vários órgãos (incluindo pulmão, coração, estômago, pâncreas, rim e nódulos linfáticos mesentéricos) diminuiu, enquanto a distribuição aumentou no fígado, em comparação com AAV8, implicando em melhor transdução do fígado e maior especificidade do tecido. O sequenciamento de próxima geração indicou enriquecimento significativo de AAV8.AR2.08 durante as seleções in vivo, demonstrando o potencial da abordagem para

isolar capsídeos com tropismos novos e melhorados. Outras comparações são mostradas nas FIGs.21 a 27). A FIG 28 mostra a biodistribuição de AAV8.AR2.08 e AAV8.

#### D. Código de barras

[00252] Camundongos pretos 6 foram injetados I.V. com  $2 \times 10^{12}$  GC/camundongo de uma mistura de rAAVG3 derivada de 12 preparações (FIG. 18B). Cada preparação contém um código de barras separado dentro do genoma do vetor, permitindo a identificação da preparação específica (FIG. 18A). Após duas semanas, os animais foram sacrificados e os tecidos coletados. Conforme predito, a expressão de rAAVG3 foi maior no fígado do que no coração ou músculo. FIG. 18C. Os experimentos de distribuição de tecido mostram que as frequências reais correspondem às frequências teóricas de códigos de barras na mistura de vetores injetados (FIG. 18D, total; FIG. 19A, 19B, músculo; FIG. 19C, 19D, coração; e FIG. 19E, 19F, fígado), com leves anomalias em BC02 e BC06 (FIGs.20A-20C).

[00253] Todos os documentos citados neste relatório descritivo são incorporados aqui por referência. Os Pedidos de Patente Provisório U.S. 62/722.388 e 62/722.382, ambos depositados em 24 de agosto de 2018, os Pedidos de Patente Provisório U.S. 62/703.670 e 62/703.673, ambos depositados em 26 de julho de 2018, os Pedidos de Patente Provisório U.S. 62/677.471 e 62/677.474, ambos depositados em 29 de maio de 2018, o Pedido de Patente Provisório U.S. 62/667.585, depositado em 29 de maio de 2018, e o Pedido de Patente Provisório U.S. 62/635.964, depositado em 27 de fevereiro de 2018 são aqui incorporados por referência. O Pedido de Patente Provisório U.S. 62/667.881, depositado em 7 de maio de 2018, o Pedido de Patente Provisório U.S. 62/667.888, depositado em 7 de maio de 2018, o Pedido de Patente Provisório U.S. 62/667.587, depositado em 6 de maio de 2018, o Pedido de Patente

Provisório U.S. 62/663.797, depositado em 27 de abril de 2018, o Pedido de Patente Provisório U.S. 62/663.788, depositado em 27 de abril de 2018, o Pedido de Patente Provisório U.S. 62/635.968, depositado em 27 de fevereiro de 2018 são incorporados por referência. As SEQ ID NO que são aqui referidas e que aparecem na listagem de sequências anexada são incorporadas por referência. Embora a invenção tenha sido descrita com referência a modalidades particulares, será apreciado que modificações podem ser feitas sem se afastar do espírito da invenção. Tais modificações pretendem ser abrangidas pelo escopo das reivindicações anexas.

## REIVINDICAÇÕES

1. Vírus adenoassociado recombinante (rAAV), caracterizado pelo fato de que compreende:

(A) um capsídeo AAVrh79 compreendendo um ou mais dos seguintes:

(1) Proteínas do capsídeo AAVrh79 compreendendo:

uma população heterogênea de proteínas AAVrh79 vp1 selecionadas de: proteínas vp1 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 2, proteínas vp1 produzidas a partir da SEQ ID NO: 1, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 2,

uma população heterogênea de proteínas AAVrh79 vp2 selecionadas de: proteínas vp2 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 2, proteínas vp2 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 412 a 2214 da SEQ ID NO: 1, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 412 a 2214 da SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 2,

uma população heterogênea de proteínas AAVrh79 vp3 selecionadas de: proteínas vp3 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 2, proteínas vp3 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 607 a 2214 da SEQ ID NO:

1, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 607 a 2214 da SEQ ID NO: 1 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 2; e/ou

(2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos da SEQ ID NO: 2, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 2, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 da SEQ ID NO: 2, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 2 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos; e

(B) um genoma de vetor no capsídeo de AAVrh79, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

2. rAAV, de acordo com a reivindicação 1, caracterizado pelo fato de que as asparaginas desamidadas são desamidadas em ácido aspártico, ácido isoaspártico, um par interconvertido de ácido aspártico/ácido isoaspártico ou combinações dos mesmos.

3. rAAV, de acordo com a reivindicação 1 ou 2, caracterizado

pelo fato de que o capsídeo compreende ainda glutamina(s) desamidada(s) que são desamidadas em ácido ( $\alpha$ )-glutâmico, ácido  $\gamma$ -glutâmico, um par de interconversão ácido ( $\alpha$ )-glutâmico/ácido  $\gamma$ -glutâmico ou combinações dos mesmos.

4. rAAV, de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que o capsídeo AAVrh79 compreende subpopulações com um ou mais dos seguintes:

(a) pelo menos 75% das asparaginas (N) em pares asparagina - glicina localizados na posição 57 das proteínas vp1 são desamidados, com base na numeração de SEQ ID NO: 2;

(b) pelo menos 75% de N em pares de asparagina - glicina na posição 263 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 2,

(c) pelo menos 70% de N nos pares asparagina-glicina na posição 385 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 2; e/ou

(d) pelo menos 85% de N em pares de asparagina - glicina na posição 514 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 2.

5. rAAV, de acordo com qualquer uma das reivindicações 1 a 4, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende uma subpopulação de vp1 na qual 75% a 100% do N na posição 57 das proteínas vp1 são desamidados, conforme determinado usando espectrometria de massa.

6. rAAV, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas

quais 75% a 100% do N na posição 263, com base na numeração de SEQ ID NO: 2, são desamidados conforme determinado usando espectrometria de massa.

7. rAAV, de acordo com qualquer uma das reivindicações 1 a 6, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 385, com base na numeração de SEQ ID NO: 2, são desamidados conforme determinado usando espectrometria de massa.

8. rAAV, de acordo com qualquer uma das reivindicações 1 a 7, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 514, com base na numeração de SEQ ID NO: 2, são desamidados.

9. rAAV, de acordo com qualquer uma das reivindicações 1 a 8, caracterizado pelo fato de que a sequência de ácido nucleico que codifica as proteínas é a SEQ ID NO: 1 ou uma sequência pelo menos 80% a pelo menos 99% idêntica à SEQ ID NO: 1 que codifica a sequência de aminoácidos da SEQ ID NO: 2.

10. O rAAV, de acordo com a reivindicação 9, caracterizado pelo fato de que a sequência de ácido nucleico é pelo menos 80% a 97% idêntica à SEQ ID NO: 1.

11. rAAV, de acordo com qualquer uma das reivindicações 1 a 10, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende ainda pelo menos subpopulação de proteínas vp1, vp2 e/ou vp3 tendo modificações de aminoácidos de SEQ ID NO: 2 compreendendo pelo menos cerca de 50 a 100% de desamidação em pelo menos quatro posições selecionadas de um ou mais de N57, 263, 385, 514 ou combinações dos mesmos.

12. rAAV, de acordo com qualquer uma das reivindicações 1

a 11, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda 1% a cerca de 40% de desamidação em pelo menos uma ou mais das posições N94, N254, N305, N410, N479, N653 ou combinações dos mesmos.

13. rAAV, de acordo com qualquer uma das reivindicações 1 a 11, caracterizado pelo fato de que o capsídeo rAAVrh79 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda uma ou mais modificações selecionadas de uma ou mais modificações em um ou mais dos seguintes: lisina acetilada, serina fosforilada e/ou treonina, ácido aspártico isomerizado, triptofano oxidado e/ou metionina, ou um aminoácido amidado.

14. Vírus adenoassociado recombinante (rAAV), caracterizado pelo fato de que compreende:

(A) um AAV8.capsídeo AR2.08 compreendendo um ou mais dos seguintes:

(1) Proteínas do capsídeo AAV G3 compreendendo:

uma população heterogênea de AAV8.proteínas AR2.08vp1 selecionadas de: proteínas vp1 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18, proteínas vp1 produzidas a partir da SEQ ID NO: 17, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de 1 a 738 da SEQ ID NO: 18,

uma população heterogênea de AAV8.Proteínas AR2.08vp2 selecionadas de: AAV8.proteínas AR2.08vp2 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, proteínas vp2 produzidas a partir



de uma sequência que compreende pelo menos nucleotídeos 411 a 2214 da SEQ ID NO: 17, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 412 a 2214 da SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18,

uma população heterogênea de AAV8.proteínas AR2.08vp3 selecionadas de: proteínas vp3 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 18, proteínas vp3 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 607 a 2214 da SEQ ID NO: 17, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 607 a 2214 da SEQ ID NO: 17 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 204 a 738 da SEQ ID NO: 18; e/ou

(2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos da SEQ ID NO: 18, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca dos aminoácidos 138 a 738 da SEQ ID NO: 18, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 204 a 738 da SEQ ID NO: 18, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 18 e opcionalmente compreendendo ainda subpopula-

ções compreendendo outros aminoácidos desamidados, em que a desamidação resulta em uma mudança de aminoácidos; e

(B) um genoma de vetor no capsídeo AAV8.AR2.08, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

15. rAAV, de acordo com a reivindicação 14, caracterizado pelo fato de que as asparaginas desamidadas são desamidadas em ácido aspártico, ácido isoaspártico, um par interconvertido de ácido aspártico/ácido isoaspártico ou combinações dos mesmos.

16. rAAV, de acordo com a reivindicação 14 ou 15, caracterizado pelo fato de que o capsídeo compreende ainda glutamina(s) desamidada(s) que são desamidadas em ácido ( $\alpha$ )-glutâmico, ácido  $\gamma$ -glutâmico, um par de interconversão ácido ( $\alpha$ )-glutâmico/ácido  $\gamma$ -glutâmico ou combinações dos mesmos.

17. rAAV, de acordo com qualquer uma das reivindicações 14 a 16, caracterizado pelo fato de que o capsídeo AAV8.AR2.08 compreende subpopulações com um ou mais de:

(a) pelo menos 75% das asparaginas (N) em pares asparagina - glicina localizados nas posições 57 das proteínas vp1 são desamidados, com base na numeração de SEQ ID NO: 18;

(b) pelo menos 75% de N em pares de asparagina - glicina na posição 263 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 18,

(c) pelo menos 70% de N nos pares asparagina-glicina na posição 385 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID

NO: 18; e/ou

(d) pelo menos 85% de N nos pares asparagina-glicina na posição 514 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 18; e/ou

(e) pelo menos 85% de N em pares de asparagina - glicina na posição 540 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 18.

18. rAAV, de acordo com qualquer uma das reivindicações 14 a 17, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende uma subpopulação de vp1 na qual 75% a 100% do N na posição 57 das proteínas vp1 são desamidados, conforme determinado usando espectrometria de massa.

19. rAAV, de acordo com qualquer uma das reivindicações 14 a 18, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 263, com base na numeração de SEQ ID NO: 2, são desamidados conforme determinado usando espectrometria de massa.

20. rAAV, de acordo com qualquer uma das reivindicações 14 a 19, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 385, com base na numeração de SEQ ID NO: 2, são desamidados conforme determinado usando espectrometria de massa.

21. rAAV, de acordo com qualquer uma das reivindicações 14 a 20, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas

vp3 nas quais 75% a 100% do N na posição 514, com base na numeração de SEQ ID NO: 2, são desamidados.

22. rAAV, de acordo com qualquer uma das reivindicações 14 a 20, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 540, com base na numeração de SEQ ID NO: 2, são desamidados.

23. rAAV, de acordo com qualquer uma das reivindicações 14 a 22, caracterizado pelo fato de que a sequência de ácido nucleico que codifica as proteínas é a SEQ ID NO: 17 ou uma sequência pelo menos 80% a pelo menos 99% idêntica à SEQ ID NO: 17 que codifica a sequência de aminoácidos da SEQ ID NO: 18.

24. rAAV, de acordo com a reivindicação 23, caracterizado pelo fato de que a sequência de ácido nucleico é pelo menos 80% a 97% idêntica à SEQ ID NO: 17.

25. rAAV, de acordo com qualquer uma das reivindicações 1 a 10, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende ainda pelo menos subpopulação de proteínas vp1, vp2 e/ou vp3 tendo modificações de aminoácidos de SEQ ID NO: 18 compreendendo pelo menos cerca de 50 a 100% de desamidação em pelo menos quatro posições selecionadas de um ou mais de N57, 263, 385, 514, 540 ou combinações dos mesmos.

26. rAAV, de acordo com qualquer uma das reivindicações 14 a 25, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda 1% a cerca de 40% de desamidação em pelo menos uma ou mais das posições N94, N254, N305, N521, ou combinações dos mesmos.

27. rAAV, de acordo com qualquer uma das reivindicações

14 a 25, caracterizado pelo fato de que o capsídeo rAAV8.AR2.08 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda uma ou mais modificações selecionadas de uma ou mais modificações em um ou mais dos seguintes: lisina acetilada, serina fosforilada e/ou treonina, ácido aspártico isomerizado, triptofano oxidado e/ou metionina, ou um aminoácido amidado.

28. Vírus adenoassociado recombinante (rAAV), caracterizado pelo fato de que compreende:

(A) um capsídeo AAV5.5.9 compreendendo um ou mais dos seguintes:

(1) Proteínas de capsídeo AAV5.5.9 compreendendo:

uma população heterogênea de proteínas AAV5.5.9 vp1 selecionadas de: proteínas vp1 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de 1 a 736 da SEQ ID NO: 10, proteínas vp1 produzidas a partir da SEQ ID NO: 9, ou proteínas vp1 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica à SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de 1 a 726 da SEQ ID NO: 1,

uma população heterogênea de proteínas AAV5.5.9 vp2 selecionadas de: proteínas AAV5.5.9 vp2 produzidas por expressão de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 137 a 726 da SEQ ID NO: 10, proteínas vp2 produzida a partir de uma sequência compreendendo pelo menos os nucleotídeos 409 a 2178 da SEQ ID NO: 9, ou proteínas vp2 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 409 a 2178 da SEQ ID NO: 9 que codificam a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 137 a 726 da SEQ ID NO:

10,

uma população heterogênea de proteínas AAV5.5.9 vp3 selecionadas de: proteínas AAV5.5.9 vp3 produzidas por expressão a partir de uma sequência de ácido nucleico que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 193 a 726 da SEQ ID NO: 10, proteínas vp3 produzidas a partir de uma sequência que compreende pelo menos nucleotídeos 577 a 2178 da SEQ ID NO: 9, ou proteínas vp3 produzidas a partir de uma sequência de ácido nucleico pelo menos 70% idêntica a pelo menos os nucleotídeos 577 a 2178 da SEQ ID NO: 9 que codifica a sequência de aminoácidos predita de pelo menos cerca dos aminoácidos 193 a 726 da SEQ ID NO: 10; e/ou

(2) uma população heterogênea de proteínas vp1 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos da SEQ ID NO: 10, uma população heterogênea de proteínas vp2 que são o produto de uma sequência de ácido nucleico que codifica a sequência de aminoácidos de pelo menos cerca dos aminoácidos 137 a 726 da SEQ ID NO: 10, e uma população heterogênea de proteínas vp3 que são o produto de uma sequência de ácido nucleico que codifica pelo menos os aminoácidos 193 a 726 da SEQ ID NO: 10, em que: as proteínas vp1, vp2 e vp3 contêm subpopulações com modificações de aminoácidos compreendendo pelo menos duas asparaginas altamente desamidadas (N) em pares asparagina-glicina na SEQ ID NO: 10 e opcionalmente compreendendo ainda subpopulações compreendendo outros aminoácidos desamidados, em que a desamidação resulta em um mudança de aminoácidos; e

(B) um genoma de vetor no capsídeo de AAV5.5.9, o genoma de vetor compreendendo uma molécula de ácido nucleico compreendendo sequências de repetição terminal invertida de AAV e uma

sequência de ácido nucleico não AAV que codifica um produto operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira.

29. rAAV, de acordo com a reivindicação 28, caracterizado pelo fato de que as asparaginas desamidadas são desamidadas em ácido aspártico, ácido isoaspártico, um par interconvertido de ácido aspártico/ácido isoaspártico ou combinações dos mesmos.

30. rAAV, de acordo com a reivindicação 28 ou 29, caracterizado pelo fato de que o capsídeo compreende ainda glutamina(s) desamidada(s) que são desamidadas em ácido ( $\alpha$ )-glutâmico, ácido  $\gamma$ -glutâmico, um par de interconversão ácido ( $\alpha$ )-glutâmico/ácido  $\gamma$ -glutâmico ou combinações dos mesmos.

31. rAAV, de acordo com qualquer uma das reivindicações 28 a 30, caracterizado pelo fato de que o capsídeo AAV5.5.9 compreende subpopulações com um ou mais dos seguintes:

(a) pelo menos 75% das asparaginas (N) em pares asparagina - glicina localizados na posição 57 das proteínas vp1 são desamidados, com base na numeração de SEQ ID NO: 10;

(b) pelo menos 70% de N nos pares asparagina-glicina na posição 319 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 10; e/ou

(c) pelo menos 85% de N nos pares asparagina-glicina na posição 442 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 10; e/ou

(d) pelo menos 85% de N em pares de asparagina - glicina na posição 502 das proteínas vp1, v2 e vp3 são desamidados, com base na numeração de resíduos da sequência de aminoácidos de SEQ ID NO: 10.

32. rAAV, de acordo com qualquer uma das reivindicações 28 a 31, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende uma subpopulação de vp1 na qual 75% a 100% do N na posição 57 das proteínas vp1 são desamidados, conforme determinado usando espectrometria de massa.

33. rAAV, de acordo com qualquer uma das reivindicações 28 a 32, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 319, com base na numeração de SEQ ID NO: 10, são desamidados conforme determinado usando espectrometria de massa.

34. rAAV, de acordo com qualquer uma das reivindicações 28 a 33, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 442, com base na numeração de SEQ ID NO: 10, são desamidados conforme determinado usando espectrometria de massa.

35. rAAV, de acordo com qualquer uma das reivindicações 28 a 34, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende subpopulação de proteínas vp1, proteínas vp2 e/ou proteínas vp3 nas quais 75% a 100% do N na posição 502, com base na numeração de SEQ ID NO: 10, são desamidados.

36. rAAV, de acordo com qualquer uma das reivindicações 28 a 35, caracterizado pelo fato de que a sequência de ácido nucleico que codifica as proteínas é a SEQ ID NO: 9 ou uma sequência pelo menos 80% a pelo menos 99% idêntica à SEQ ID NO: 9 que codifica a sequência de aminoácidos da SEQ ID NO: 10.

37. rAAV, de acordo com a reivindicação 36, caracterizado pelo fato de que a sequência de ácido nucleico é pelo menos 80% a 97% idêntica à SEQ ID NO: 9.



38. rAAV, de acordo com qualquer uma das reivindicações 28 a 36, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda 1% a cerca de 40% de desamidação em pelo menos uma ou mais das posições N35, N113, N204, N217, N243, N249, N293/294, N304, N399.400, N467, N505, Q589, N618, N641, N653, N658, N699 ou combinações dos mesmos.

39. rAAV, de acordo com qualquer uma das reivindicações 28 a 38, caracterizado pelo fato de que o capsídeo rAAV5.5.9 compreende subpopulações de proteínas vp1, vp2 e/ou vp3 que compreendem ainda uma ou mais modificações selecionadas de uma ou mais modificações em um ou mais dos seguintes: lisina acetilada, serina fosforilada e/ou treonina, ácido aspártico isomerizado, triptofano oxidado e/ou metionina, ou um aminoácido amidado.

40. rAAV, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o rAAV compreende cerca de 60 proteínas de capsídeo total em uma razão de cerca de 1 vp1 a cerca de 1 a 1,5 vp2 a 3 a 10 vp3 proteínas.

41. rAAV, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que as sequências de AAV ITR são 5' ITR e 3' ITR de uma fonte de AAV diferente do AAV que fornece o capsídeo.

42. Composição, caracterizada pelo fato de que compreende um AAV como definido em qualquer uma das reivindicações precedentes e um veículo, excipiente e/ou diluente farmacêutico.

43. rAAV, de acordo com qualquer uma das reivindicações 1 a 41 ou composição de acordo com a reivindicação 42, caracterizados pelo fato de que distribuem um produto de gene desejado a um sujeito em necessidade do mesmo.

44. Uso de um AAV de acordo com qualquer uma das reivindicações 1 a 41 ou uma composição de acordo com a reivindicação 42, caracterizado pelo fato de que fornecem um produto de gene desejado a um sujeito em necessidade do mesmo.

45. Método para gerar um AAV recombinante compreendendo um capsídeo de AAV, caracterizado pelo fato de que compreende as etapas de cultura de uma célula hospedeira contendo: (a) uma sequência de ácido nucleico que codifica uma proteína do capsídeo de AAV selecionada de: proteína de capsídeo vp1, aminoácidos 1 a 738, SEQ ID NO: 2 (proteína de capsídeo AAVrh79); proteína do capsídeo vp2, aminoácidos 138 a 738, SEQ ID NO: 2 (proteína de capsídeo AAVrh79); ou proteína de capsídeo vp3, aminoácidos 204 a 738, SEQ ID NO: 2 (proteína de capsídeo AAVrh79); (b) um gene rep funcional; (c) um minigene compreendendo repetições terminais invertidas de AAV (ITRs) e um transgene; e (d) funções auxiliares suficientes para permitir o empacotamento do minigene no capsídeo de AAV.

46. Método para gerar um AAV recombinante, caracterizado pelo fato de que compreende um capsídeo de AAV compreendendo as etapas de cultura de uma célula hospedeira contendo: (a) uma sequência de ácido nucleico que codifica uma proteína de capsídeo de AAV selecionada: proteína de capsídeo vp1, aminoácidos 1 a 738, SEQ ID NO: 18 (proteína de capsídeo AAV8.AR2.08); proteína de capsídeo vp2, aminoácidos 138 a 738, SEQ ID NO: 18 (proteína de capsídeo AAV8.AR2.08); ou proteína de capsídeo vp3, aminoácidos 204 a 738, SEQ ID NO: 18 (proteína de capsídeo AAV8.AR2.08); (b) um gene rep funcional; (c) um minigene compreendendo repetições terminais invertidas de AAV (ITRs) e um transgene; e (d) funções auxiliares suficientes para permitir o empacotamento do minigene no capsídeo de AAV.

47. Método para gerar um AAV recombinante compreendendo um capsídeo de AAV, caracterizado pelo fato de que compreende

as etapas de cultura de uma célula hospedeira contendo: (a) uma sequência de ácido nucleico que codifica uma proteína do capsídeo de AAV selecionada de: proteína de capsídeo vp1, aminoácidos 1 a 736, SEQ ID NO: 10 (proteína de capsídeo AAV5.5.9); proteína do capsídeo vp2, aminoácidos 138 a 736, SEQ ID NO: 10 (proteína de capsídeo AAV5.5.9); ou proteína de capsídeo vp3, aminoácidos 203 a 736, SEQ ID NO: 10 (proteína de capsídeo AAV5.5.9); (b) um gene rep funcional; (c) um minigene compreendendo repetições terminais invertidas de AAV (ITRs) e um transgene; e (d) funções auxiliares suficientes para permitir o empacotamento do minigene no capsídeo de AAV.

48. Célula hospedeira, caracterizada pelo fato de que é transfectada in vitro com o AAV como definido em qualquer uma das reivindicações precedentes.

49. Molécula de ácido nucleico recombinante:

(a) codificando uma sequência de aminoácidos predita compreendendo os aminoácidos 1 a 738 de SEQ ID NO: 2; ou

(b) compreendendo os nucleotídeos 1 a 2214 da SEQ ID NO: 1, ou uma sequência de nucleotídeos pelo menos 99% idêntica aos nucleotídeos 1 a 2214 da SEQ ID NO: 1,

caracterizada pelo fato de que a molécula de ácido nucleico recombinante não contém uma repetição terminal invertida de AAV.

50. Molécula de ácido nucleico recombinante:

(a) codificando uma sequência de aminoácidos predita compreendendo os aminoácidos 1 a 738 de SEQ ID NO: 18; ou

(b) compreendendo os nucleotídeos 1 a 2214 da SEQ ID NO: 17, ou uma sequência de nucleotídeos pelo menos 99% idêntica aos nucleotídeos 1 a 2214 da SEQ ID NO: 17,

caracterizada pelo fato de que a molécula de ácido nucleico recombinante não contém uma repetição terminal invertida de AAV.

51. Molécula de ácido nucleico recombinante:

(a) codificando uma sequência de aminoácidos predita compreendendo os aminoácidos 1 a 736 de SEQ ID NO: 10; ou

(b) compreendendo os nucleotídeos 1 a 2214 da SEQ ID NO: 1, ou uma sequência de nucleotídeos pelo menos 99% idêntica aos nucleotídeos 1 a 2214 da SEQ ID NO: 1,

caracterizada pelo fato de que a molécula de ácido nucleico recombinante não contém uma repetição terminal invertida de AAV.

52. Molécula de ácido nucleico recombinante, de acordo com qualquer uma das reivindicações 49 a 51,

caracterizada pelo fato de que compreende ainda um gene rep funcional.

53. Molécula de ácido nucleico recombinante, de acordo com qualquer uma das reivindicações 49 a 51, caracterizada pelo fato de que a referida molécula de ácido nucleico recombinante é um plasmídeo.

54. Célula hospedeira, caracterizada pelo fato de que é transfectada com a molécula de ácido nucleico recombinante, como definida em qualquer uma das reivindicações 49 a 53.

55. Célula hospedeira, de acordo com a reivindicação 54, caracterizada pelo fato de que compreende um gene rep funcional, um minigene compreendendo repetições terminais invertidas (ITRs) de AAV e um transgene e funções auxiliares suficientes para permitir o empacotamento do minigene no capsídeo de AAV.

56. Sistema de produção de rAAV útil para produzir um AAV recombinante, de acordo com qualquer uma das reivindicações 1 a 41, caracterizado pelo fato de que o sistema de produção compreende:

(a) uma sequência de ácido nucleico do capsídeo AAV que codifica a sequência de aminoácidos de SEQ ID NO: 2, SEQ ID NO: 10 ou SEQ ID NO: 18;

(b) uma molécula de ácido nucleico adequada para empacotamento no capsídeo de AAV, a referida molécula de ácido nucleico

compreendendo pelo menos uma repetição terminal invertida de AAV (ITR) e uma sequência de ácido nucleico não AAV que codifica um produto de gene operacionalmente ligado a sequências que direcionam a expressão do produto em uma célula hospedeira; e

(c) funções rep de AAV e funções auxiliares suficientes para permitir o empacotamento da molécula de ácido nucleico no capsídeo de AAV recombinante.

57. Sistema, de acordo com a reivindicação 56, caracterizado pelo fato de que a sequência de ácido nucleico de (a) compreende pelo menos SEQ ID NO: 1, ou uma sequência pelo menos 70% a pelo menos 99% idêntica à SEQ ID NO: 1 que codifica a sequência de aminoácidos da SEQ ID NO: 2.

58. Sistema, de acordo com a reivindicação 56, caracterizado pelo fato de que a sequência de ácido nucleico de (a) compreende pelo menos SEQ ID NO: 9, ou uma sequência pelo menos 70% a pelo menos 99% idêntica à SEQ ID NO: 9 que codifica a sequência de aminoácidos da SEQ ID NO: 10.

59. Sistema, de acordo com a reivindicação 56, caracterizado pelo fato de que a sequência de ácido nucleico de (a) compreende pelo menos SEQ ID NO: 17, ou uma sequência pelo menos 70% a pelo menos 99% idêntica à SEQ ID NO: 17 que codifica a sequência de aminoácidos da SEQ ID NO: 18.

60. Sistema, de acordo com qualquer uma das reivindicações 56 a 59, caracterizado pelo fato de que a cultura de células compreende células 293 de rim embrionário humano.

61. Sistema, de acordo com qualquer uma das reivindicações 56 a 60, caracterizado pelo fato de que o rep de AAV é de um AAV diferente daquele que fornece o capsídeo.

62. Sistema, de acordo com a reivindicação 61, caracterizado pelo fato de que o rep de AAV é de AAV2.

63. Composição, caracterizada pelo fato de que compreende uma população de rAAV com título, potência ou transdução aumentados, a referida composição compreendendo rAAV com capsídeos que são modificados para ter desamidação total diminuída em comparação com um rAAV com um padrão de desamidação com um padrão de desamidação de capsídeo de acordo com qualquer um da Tabela A (AAVrh79), Tabela B (AAV8.AR2.08), ou Tabela C (AAV5.5.9), desde que o rAAV não seja AAVhu68.

64. Composição, de acordo com a reivindicação 63, caracterizada pelo fato de que os capsídeos de rAAV são modificados em um local diferente de:

(a) N57, N263, N398, e/ou N514, com base na numeração de SEQ ID NO: 2, com base na numeração da sequência de aminoácidos vp1 predita com o M inicial, para um capsídeo de AAVrh79;

(b) N57, N263, N385, N514 e/ou N540, com referência à numeração da SEQ ID NO: 18, com base na numeração da sequência de aminoácidos vp1 predita com o M inicial, para um capsídeo AAV8.AR2.08;

(c) N57, N319, N442 e/ou N502, com referência à numeração da SEQ ID NO: 10, com base na numeração da sequência de aminoácidos vp1 predita com o M inicial, para um capsídeo AAV5.5.9;

FIG. 1A

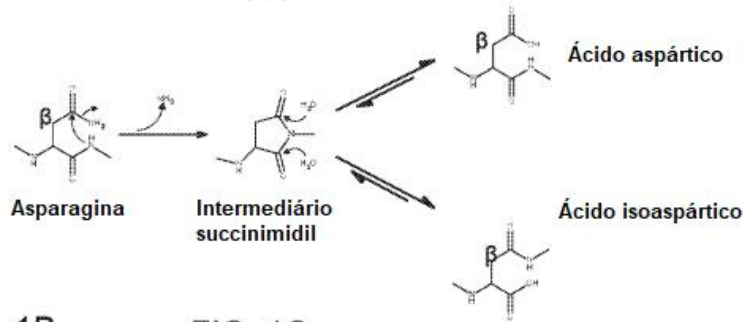


FIG. 1B

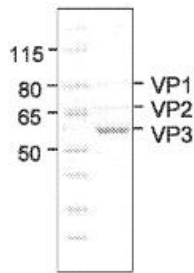


FIG. 1C

Local	pI	Local	pI
1	7.30	11	5.70
2	7.11	12	5.63
3	6.97	13	5.54
4	6.75	14	5.46
5	6.61	15	5.37
6	6.45	16	5.27
7	6.27	17	5.20
8	6.11	18	5.14
9	6.01	19	5.08
10	5.86	20	4.99

FIG. 1D

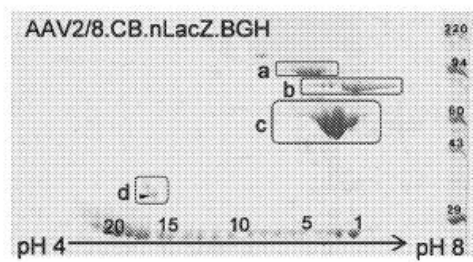
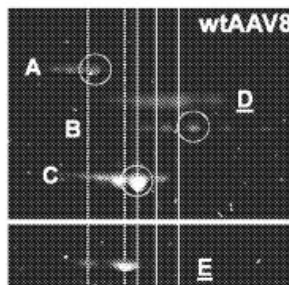
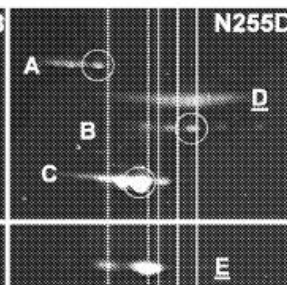


FIG. 1E



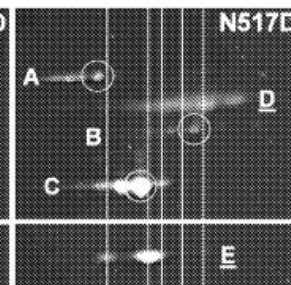
A-AAV8 VP1: 87kDa  
D-Conalbumina: 76kDa

FIG. 1F



B-AAV8 VP2: 73kDa  
E-Turbonuclease: 27kDa

FIG. 1G



C-AAV8 VP3: 61kDa

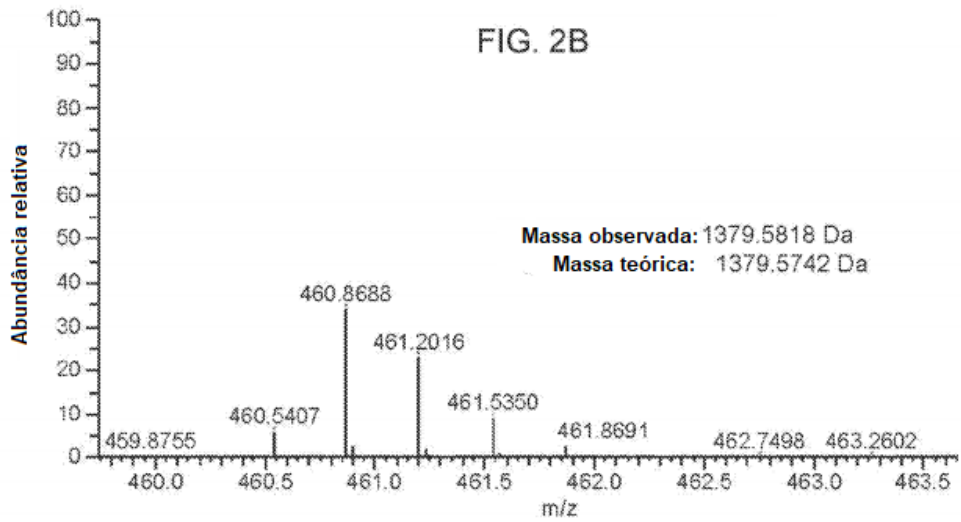
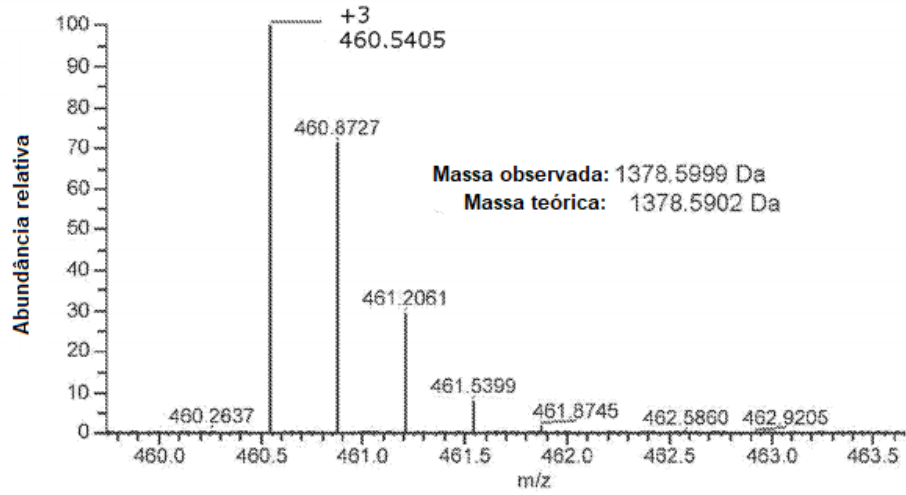
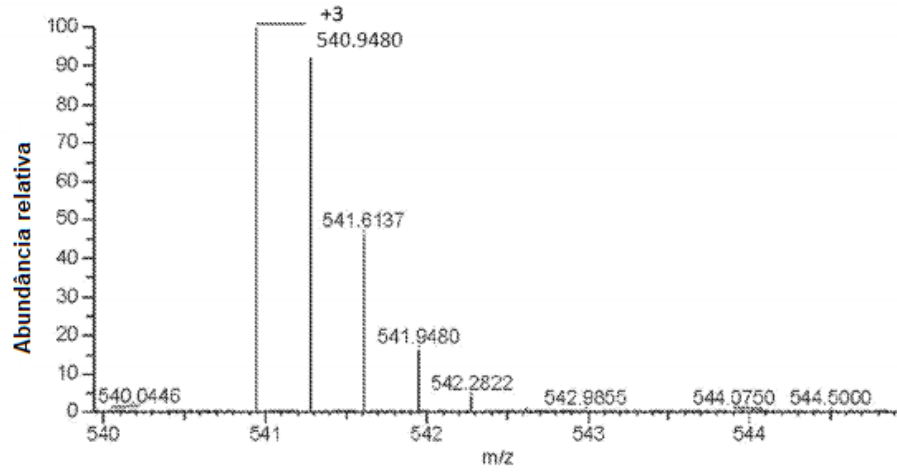


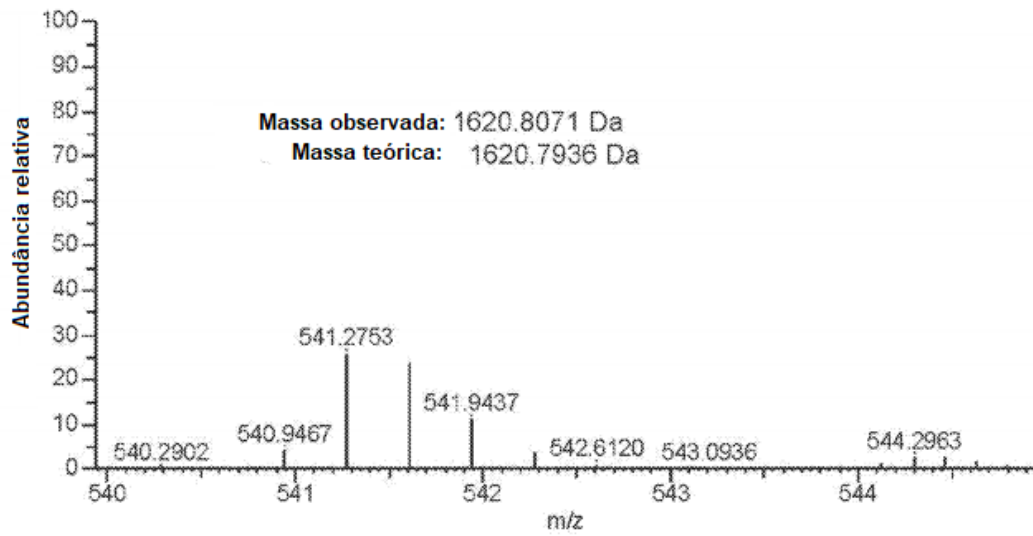


FIG. 2C

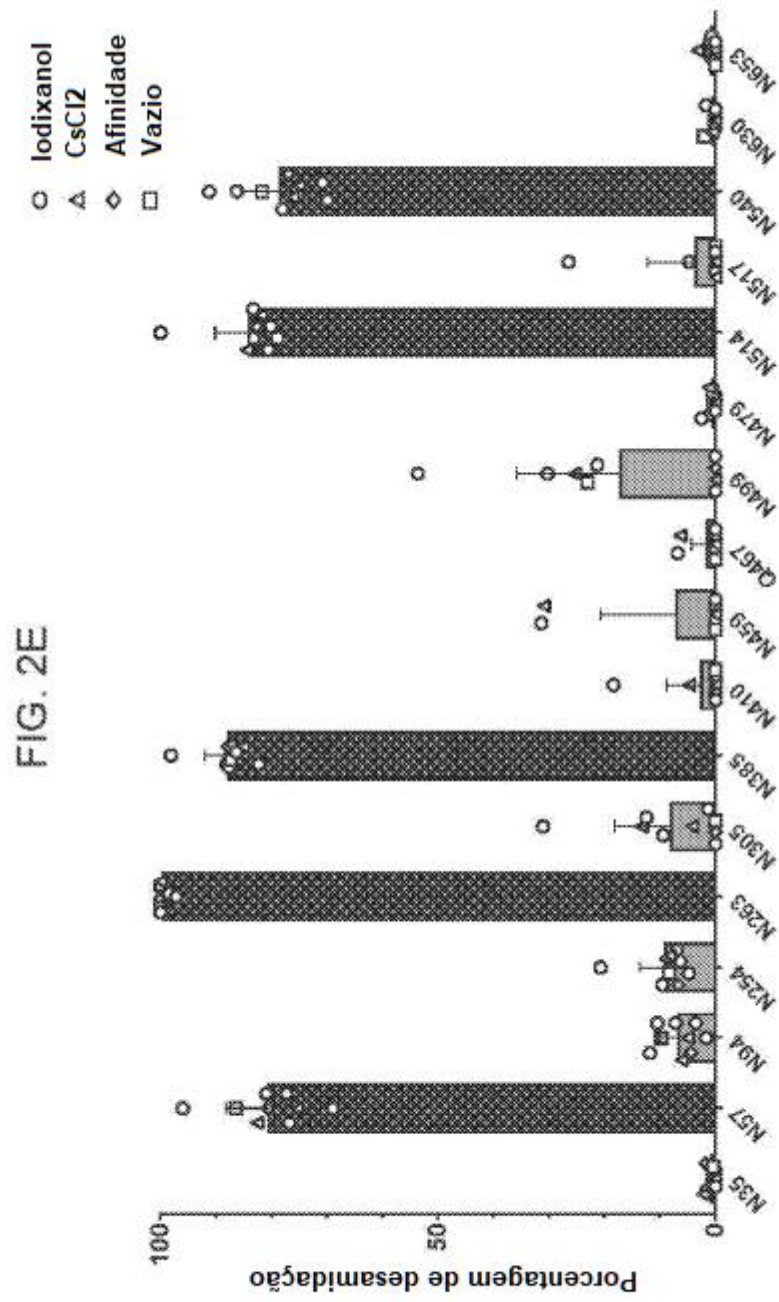


Massa observada: 1619.8213 Da  
Massa teórica: 1619.8096 Da

FIG. 2D



Massa observada: 1620.8071 Da  
Massa teórica: 1620.7936 Da



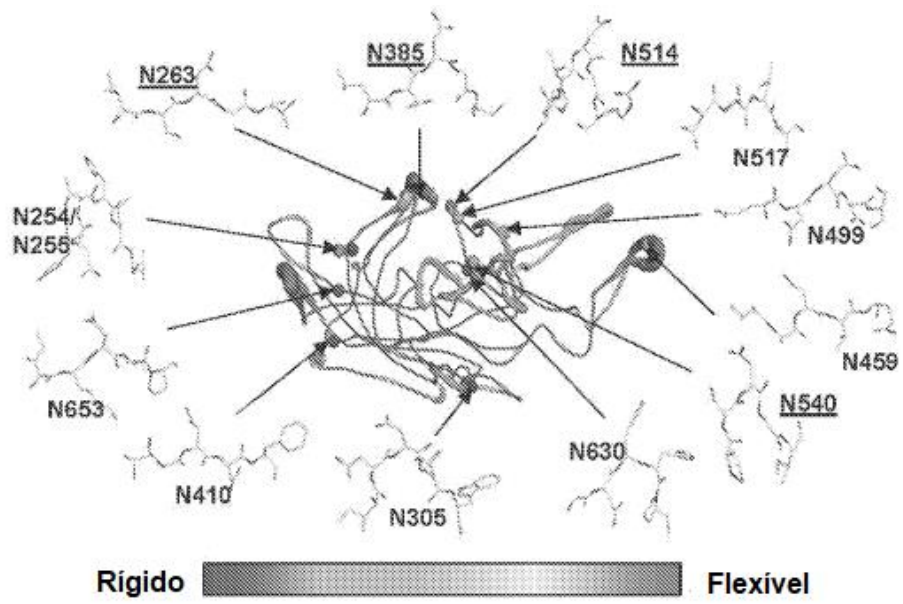


FIG. 3B

FIG. 3C

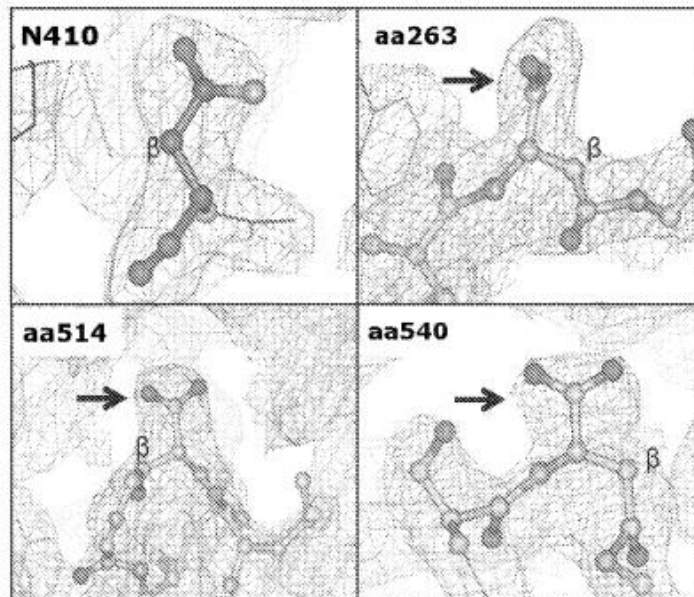


FIG. 3D

FIG. 3E

FIG. 4A

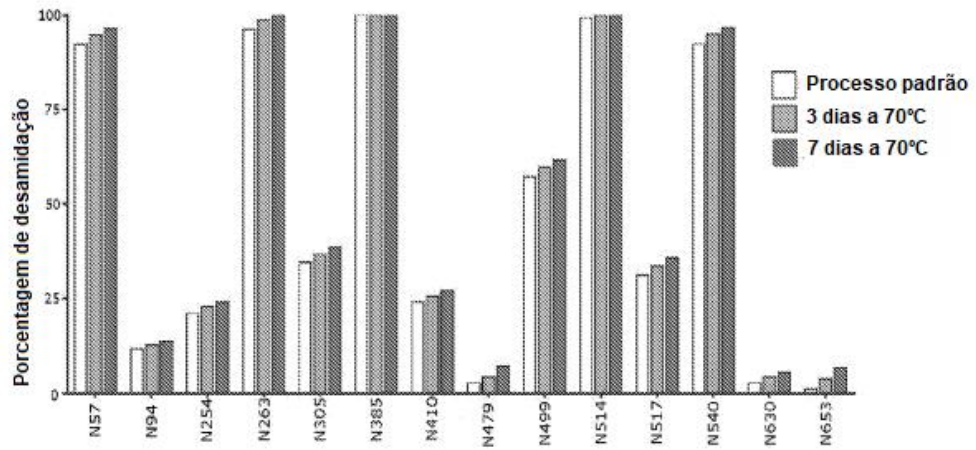


FIG. 4B

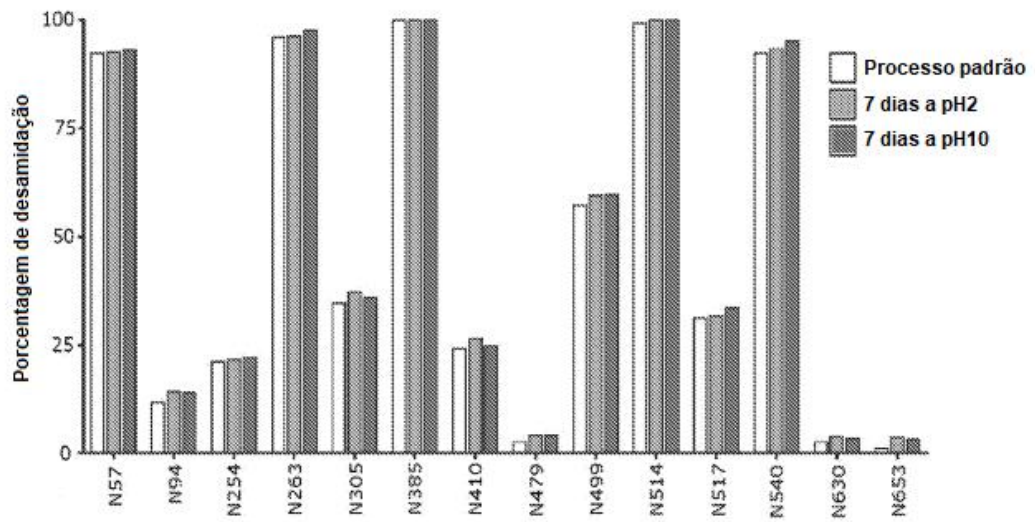


FIG. 4C

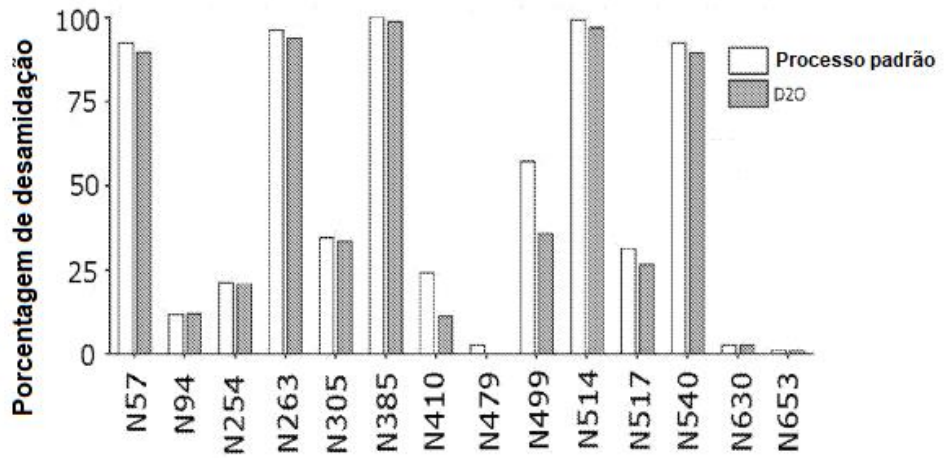


FIG. 4D

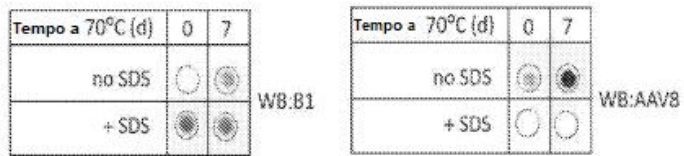


FIG. 5A

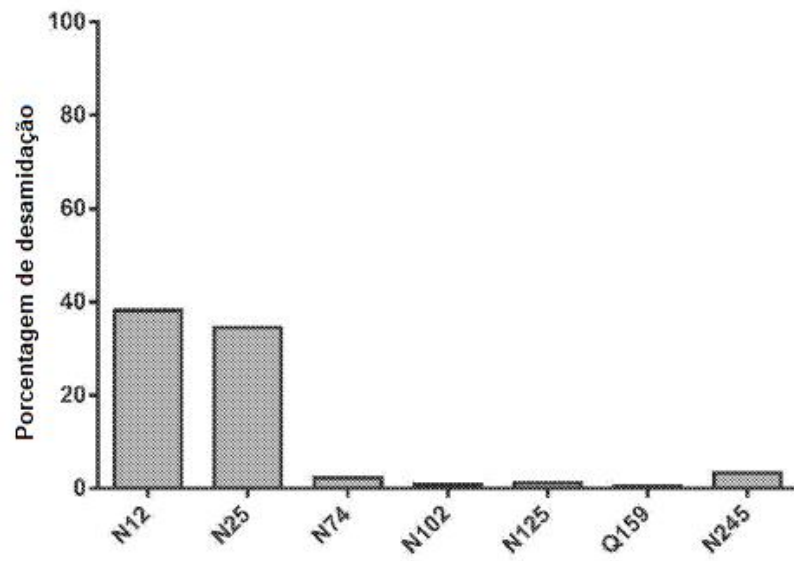
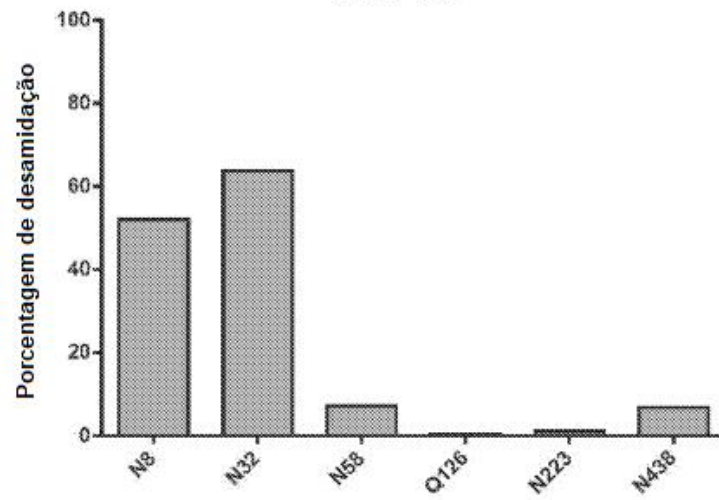


FIG. 5B



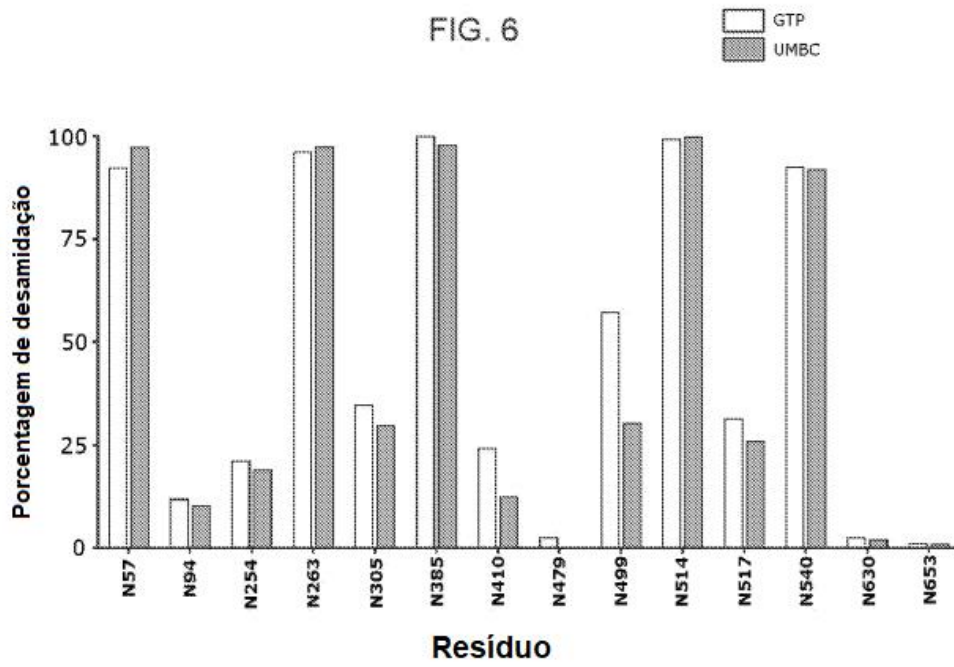


FIG. 7B

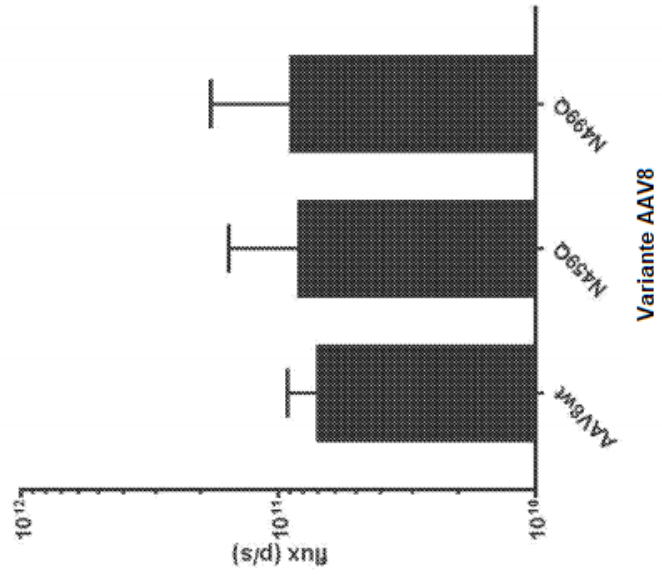
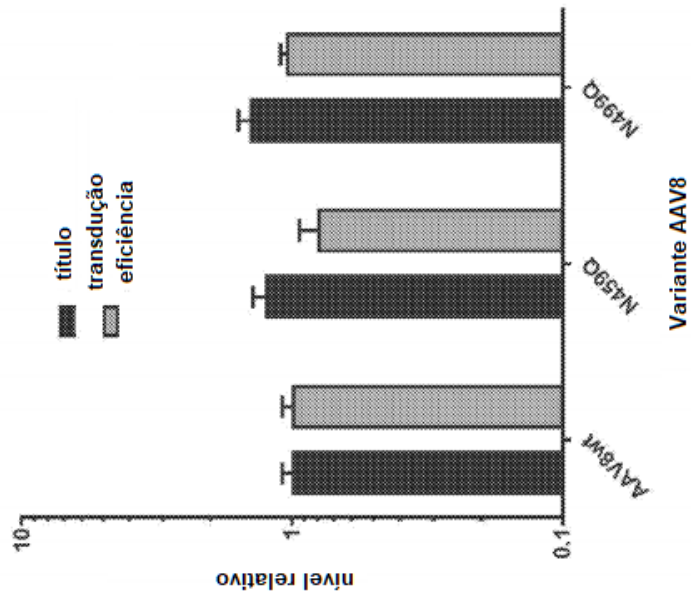


FIG. 7A





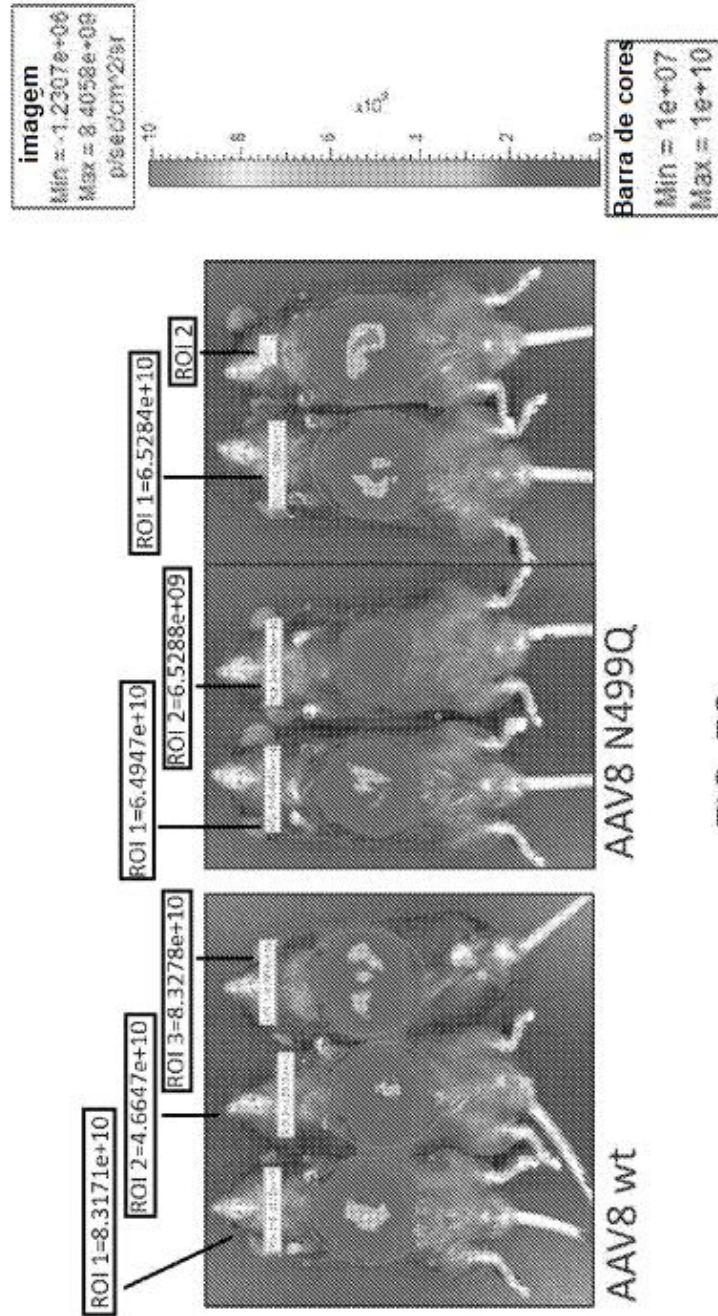


FIG. 7C

FIG. 8A

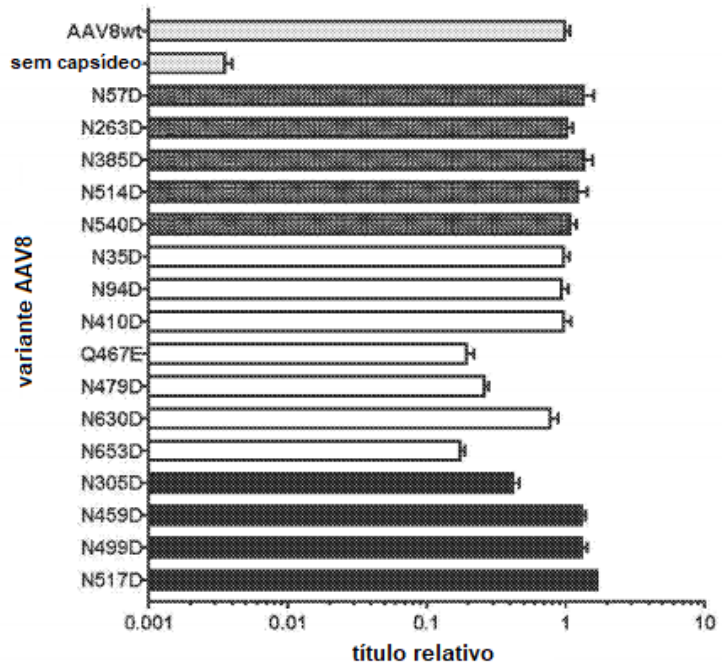


FIG. 8B

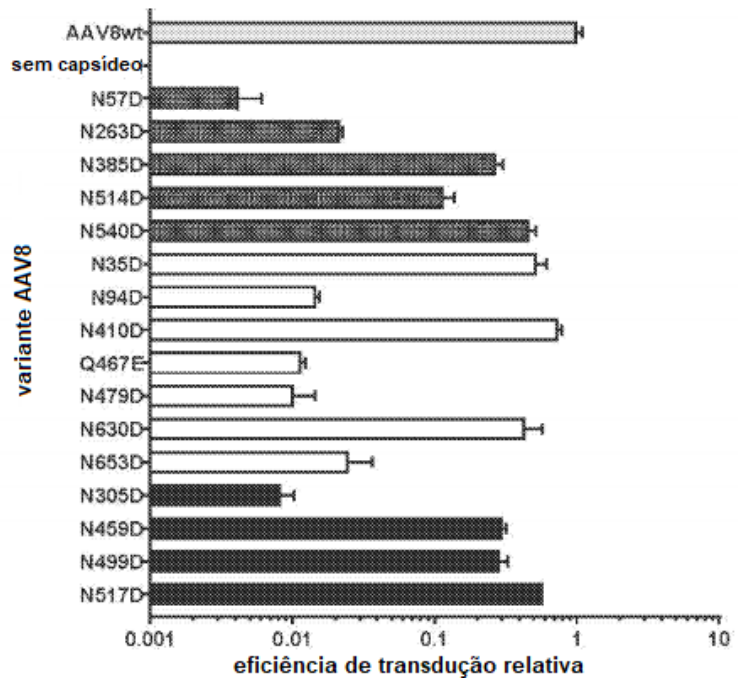


FIG. 9A

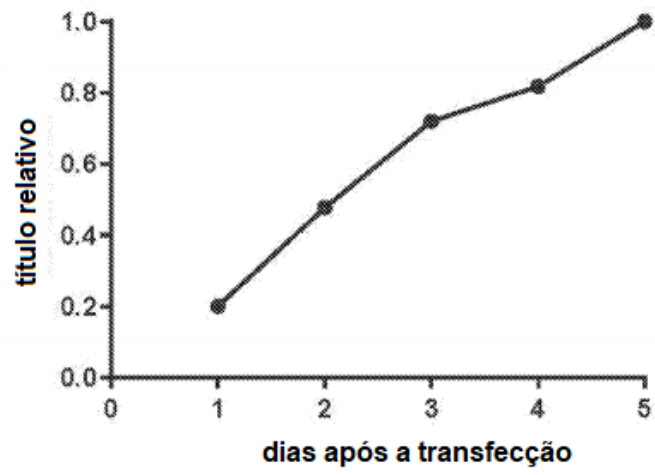


FIG. 9B

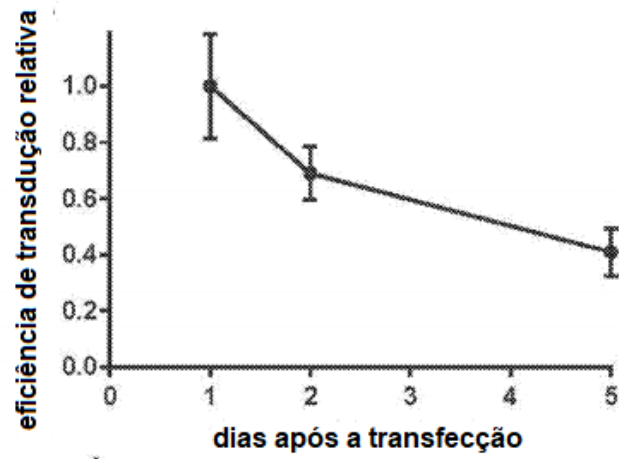


FIG. 9C

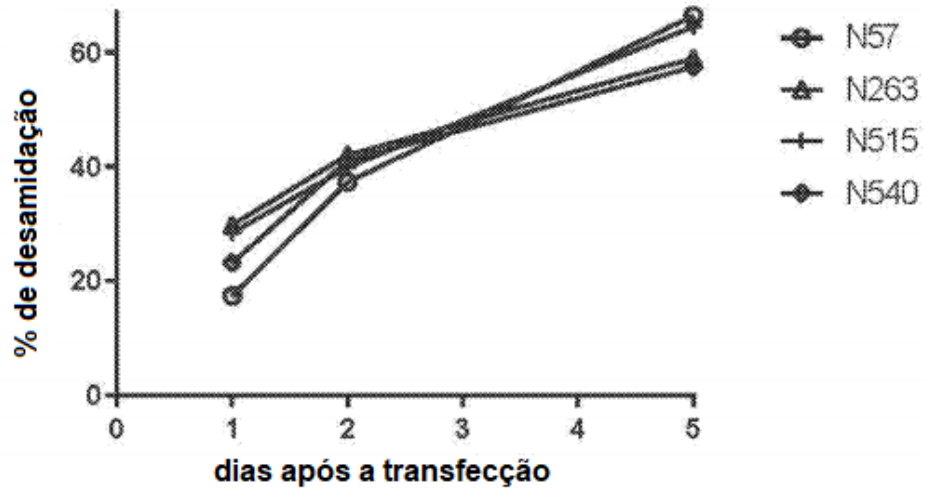
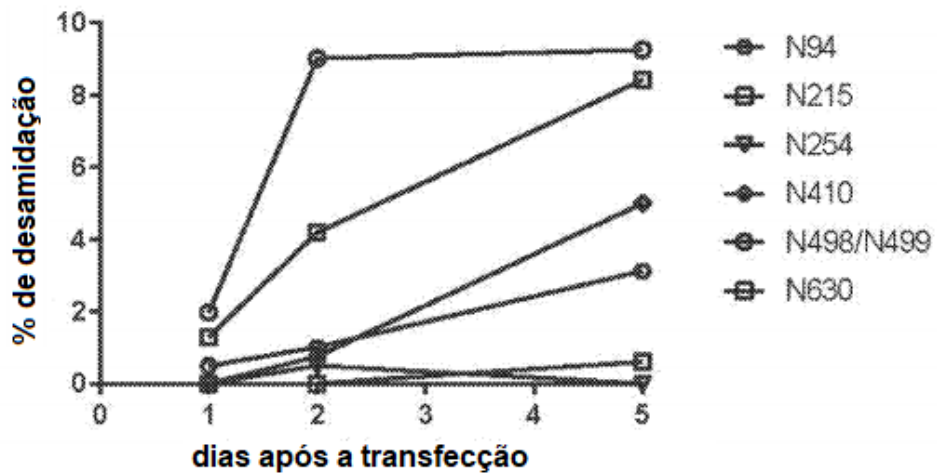
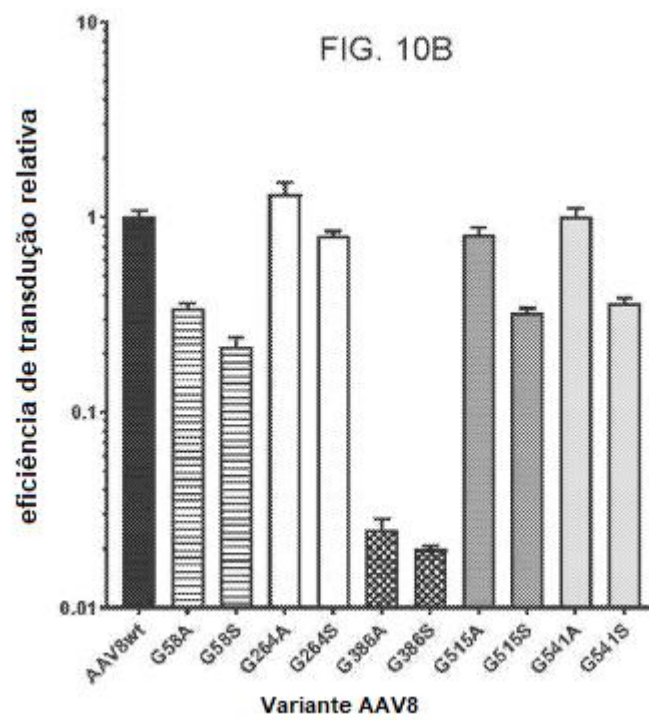
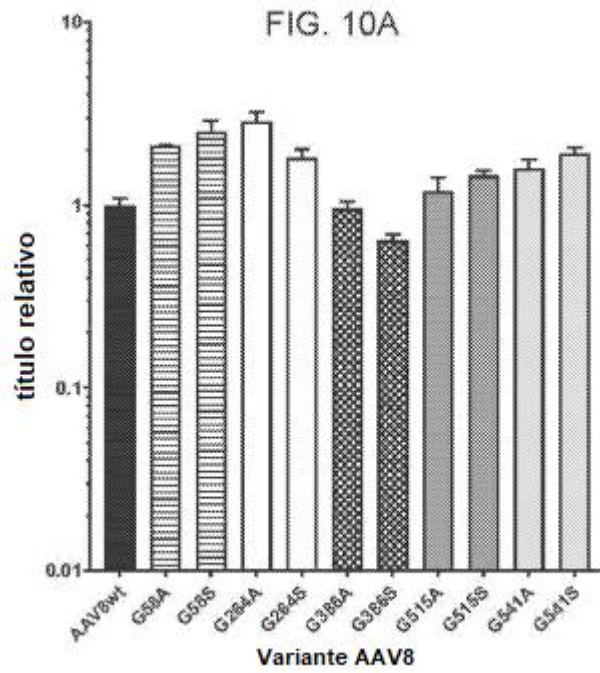


FIG. 9D





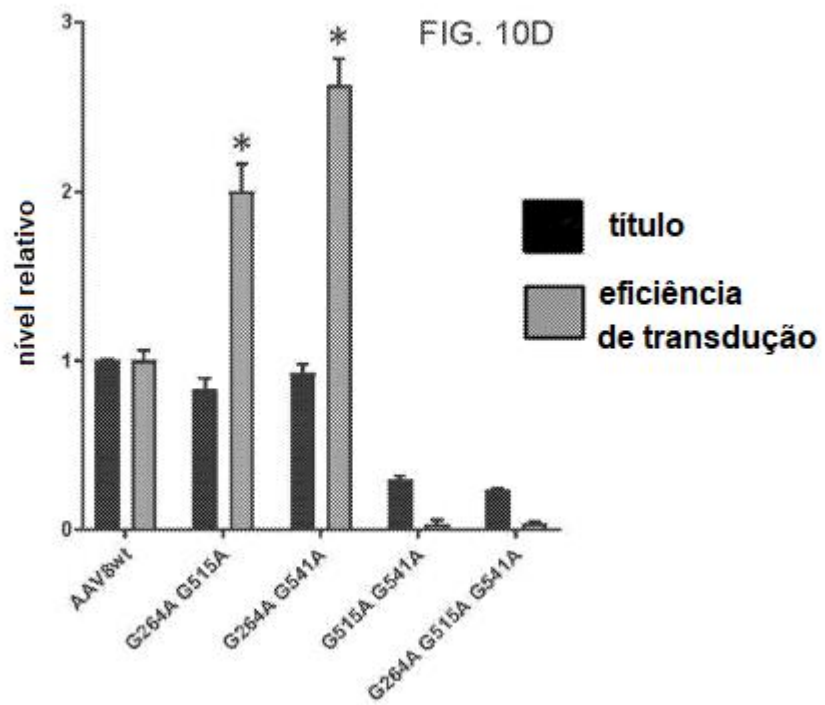
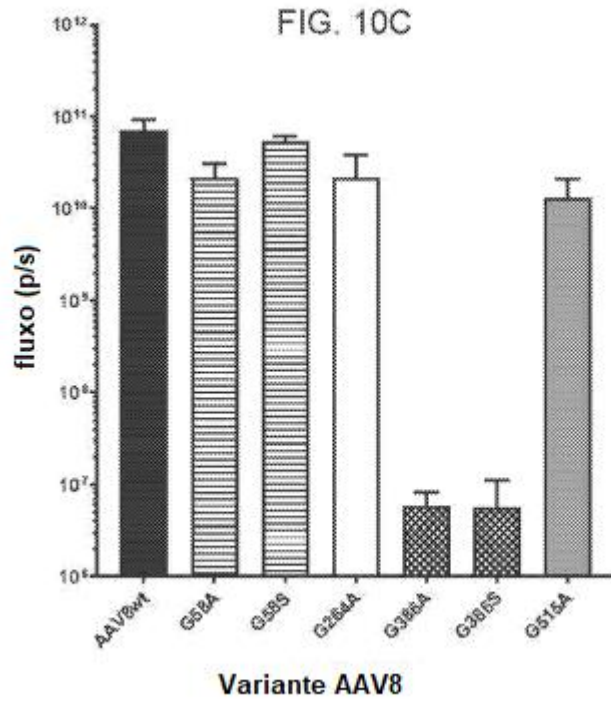


FIG. 11 A

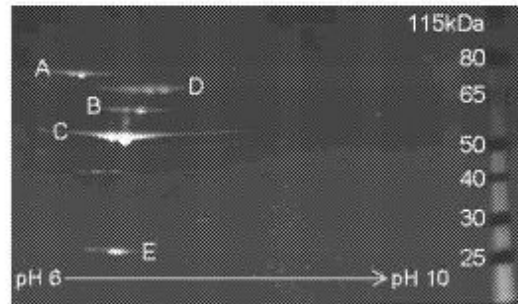


FIG. 11 C

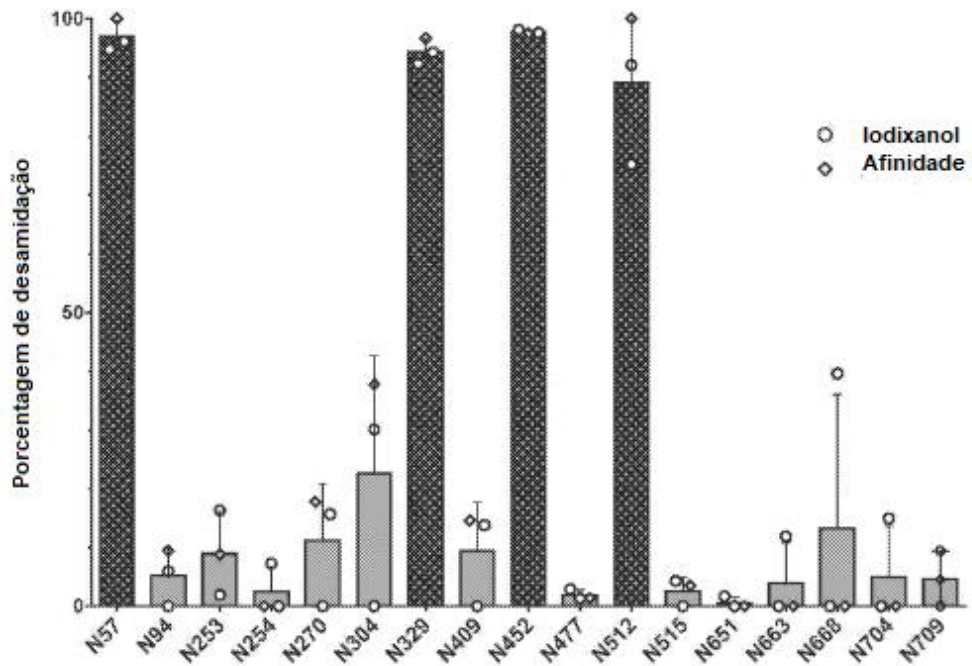
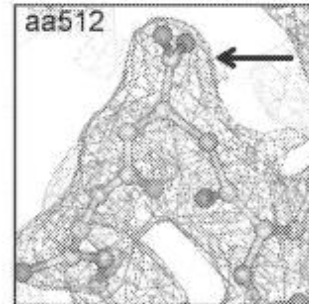


FIG. 11 B

FIG. 11D

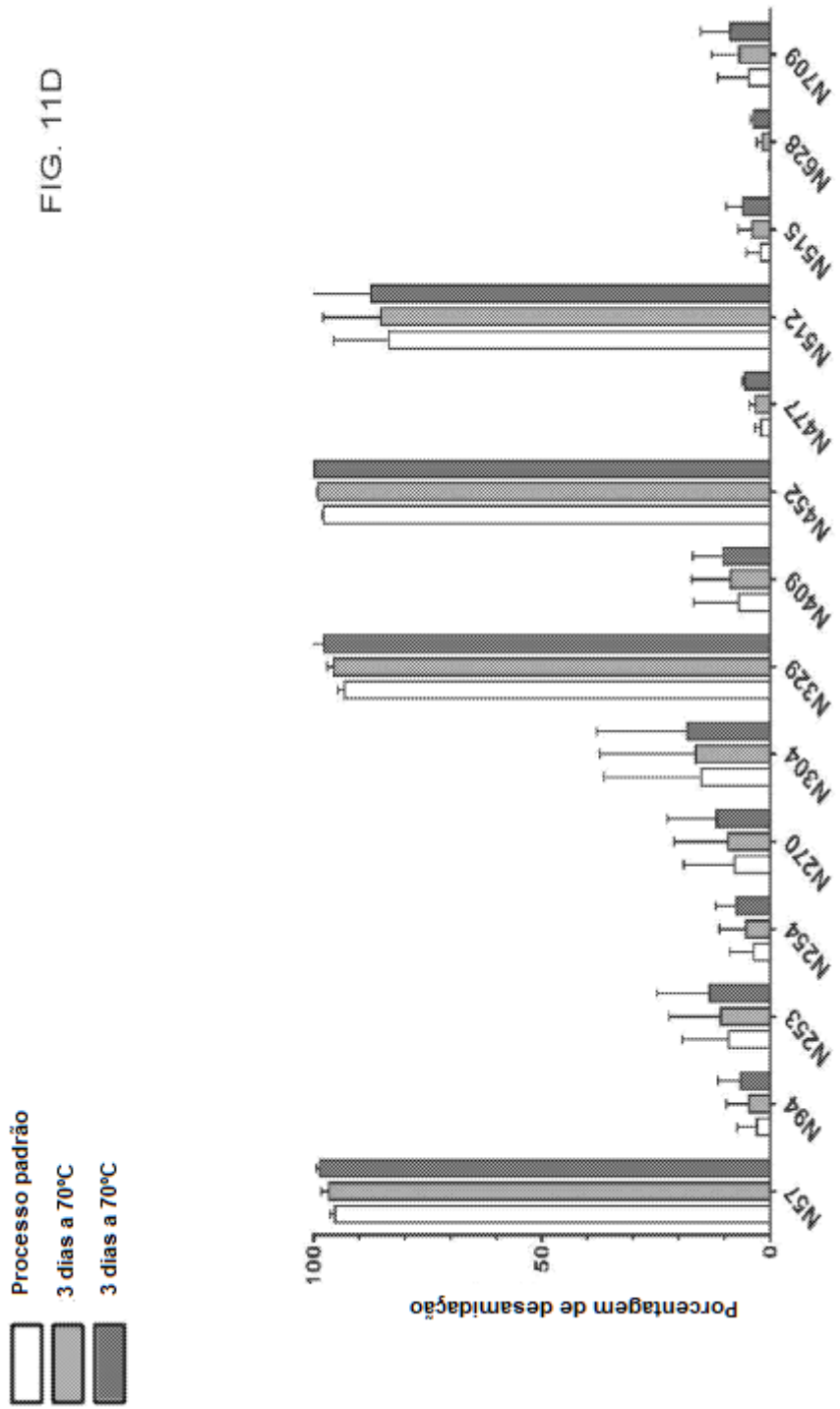




FIG. 11E

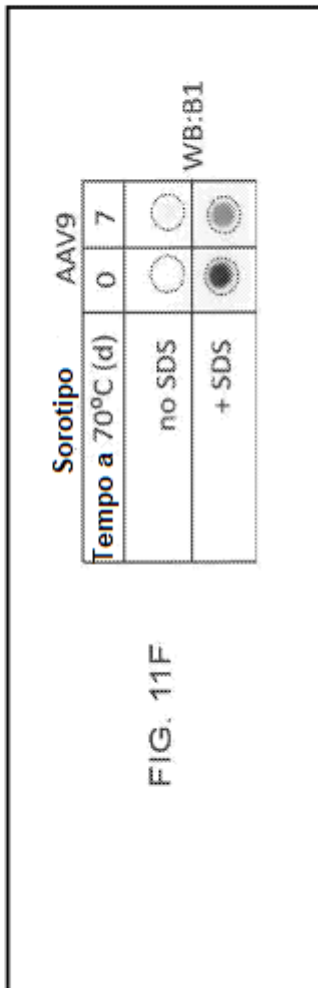
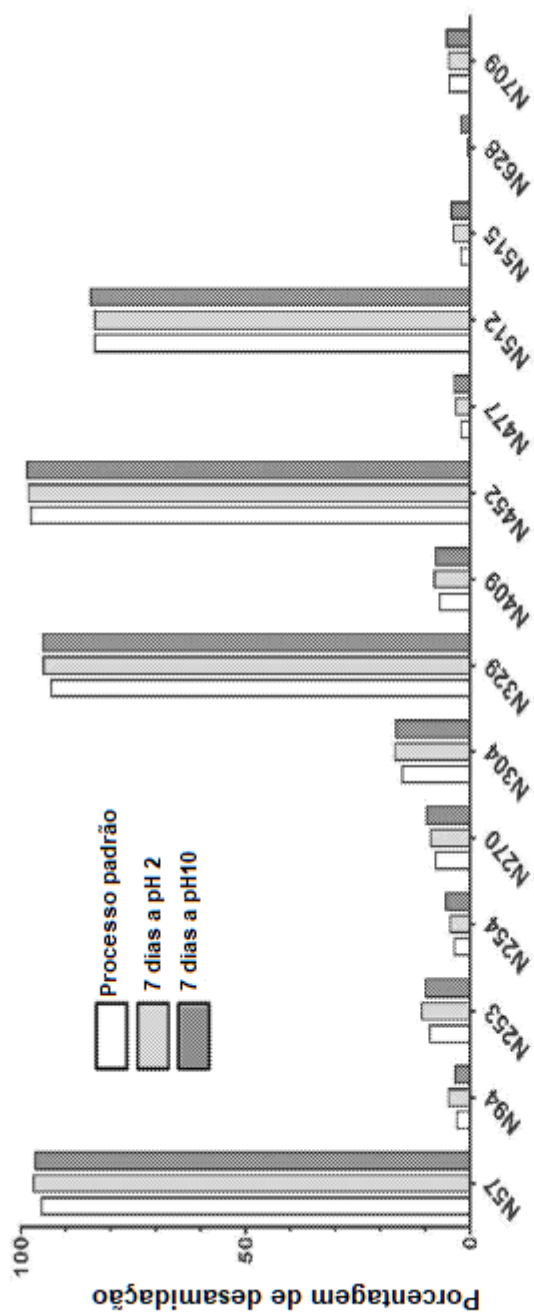


FIG. 11F

FIG. 11G

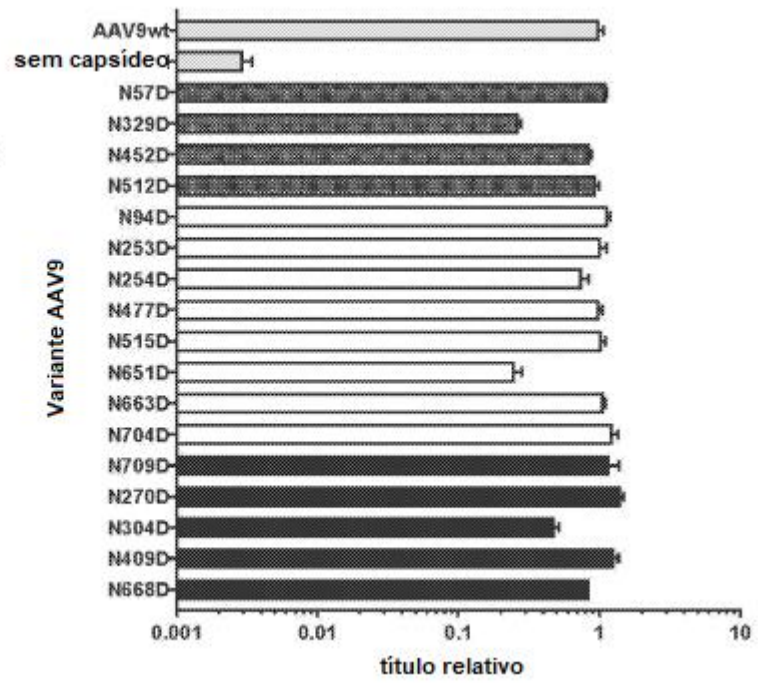


FIG. 11H

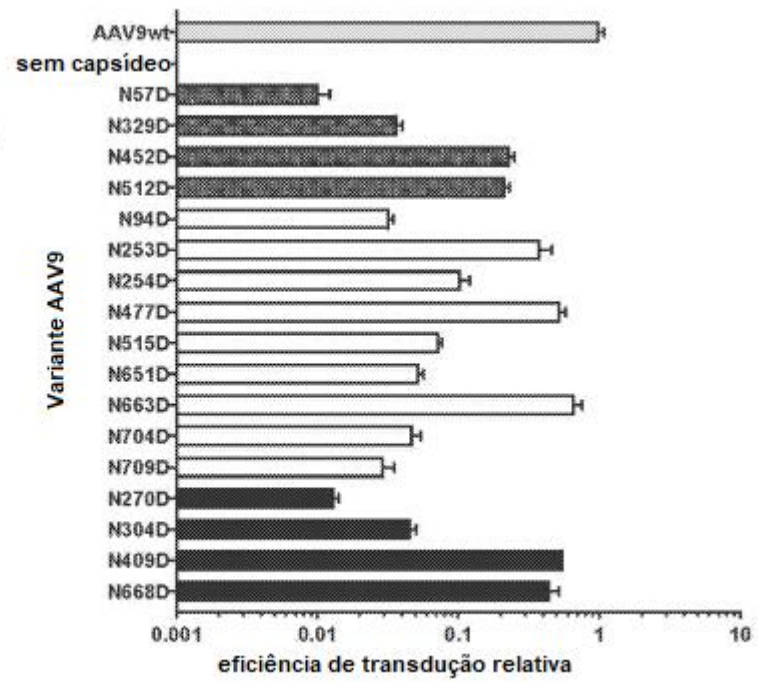


FIG. 11I

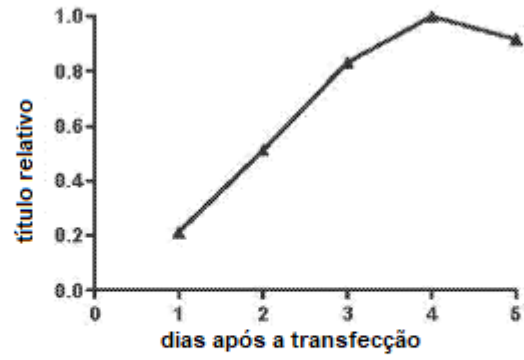


FIG. 11J

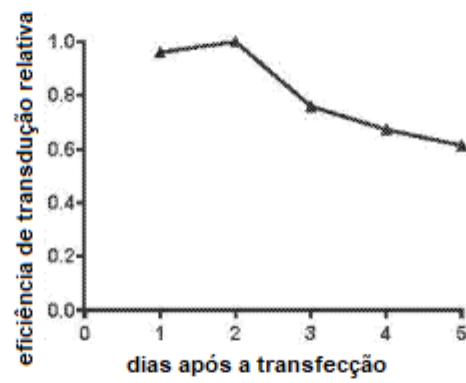


FIG. 11K

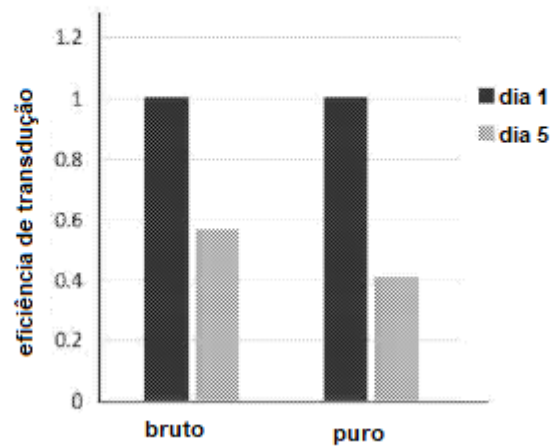


FIG. 12B ( Clado F nt)

AAVGS 1 atgtcttttggatcacctccagattgg--ttggaagaagtgtggaaggttctcgagtttttgggcttgaagggggccacgaaacccaaac  
 PHP .B 1 atggctccgatgggttatctccagattggctcgagacaaccttagtgaagaaatcgagtggtggctttgaaacctggagccctcaaccaagg  
 AAV9 1 atggctccgatgggttatctccagattggctcgagacaaccttagtgaagaaatcgagtggtggctttgaaacctggagccctcaaccaagg  
 hu68 1 atggctccgatgggttatctccagattggctcgagacaaccttagtgaagaaatcgagtggtggctttgaaacctggagccctcaaccaagg

AAVGS 101 ccaatcagcagcatcaagatcaagcccgtggtcttctgctgctggttaaaactatctcggaaccggaaacggctctgatcgagagagcctgtcaacag  
 PHP .B 101 caaatcaacaatcaagacaacgctcgaggtcttctgcttccgggttaaaactatctcggaaccggaaacggctctgatcgagagagcctgtcaacag  
 AAV9 101 caaatcaacaatcaagacaacgctcgaggtcttctgcttccgggttaaaactatctcggaaccggaaacggctctgatcgagagagcctgtcaacag  
 hu68 101 caaatcaacaatcaagacaacgctcgaggtcttctgcttccgggttaaaactatctcggaaccggaaacggctctgatcgagagagcctgtcaacag

AAVGS 201 ggcagcagaggtcgcgcgagcagcatctctcaaacagcagcagcttgagggcgggagacaacccctacctcaagtacaaccacgcggacgcagttt  
 PHP .B 201 agcagacggcggcctcgagcaacagcctacgacagcagctcaagccggagagacaacccctacctcaagtacaaccacgcggacgcagttt  
 AAV9 201 agcagacggcggcctcgagcaacagcctacgacagcagctcaagccggagagacaacccctacctcaagtacaaccacgcggacgcagttt  
 hu68 201 agcagacggcggcctcgagcaacagcctacgacagcagctcaagccggagagacaacccctacctcaagtacaaccacgcggacgcagttt

AAVGS 301 caggagaagctcgcgcgacacacatccttgggggaaacctcggaaagcagcttctcagggccaagaaagggctctcgaacctttggcctggtgag  
 PHP .B 301 cagggcggctcaagaagatacgtcttgggggcaacctcggcgagcagctctccaggccaagaaagggctcttgaacctcttggctggtgag  
 AAV9 301 caggagcggctcaagaagatacgtcttgggggcaacctcggcgagcagctctccaggccaagaaagggctcttgaacctcttggctggtgag  
 hu68 301 caggagcggctcaagaagatacgtcttgggggcaacctcggcgagcagctctccaggccaagaaagggctcttgaacctcttggctggtgag

AAVGS 401 agyytgctaagacggccctaccggaagcggatagcagacacacttccaaaagaagaagcccggaagagagactcca-----a-----  
 PHP .B 401 aagcggctaagacggctcctggaagaagagggcctctagagagctcctcaggaacggactcctccggggatttggcaaatcgggtgcaacagccgc  
 AAV9 401 aagcggctaagacggctcctggaagaagagggcctctagagagctcctcaggaacggactcctccggggatttggcaaatcgggtgcaacagccgc  
 hu68 401 aagcggctaagacggctcctggaagaagagggcctctagagagctcctcaggaacggactcctccggggatttggcaaatcgggtgcaacagccgc

AAVGS 501 -----gcctccacctctcagacccggaagctggaccagcggatccagagctgcaaatccagcccaacagcctcaagtttggggagct  
 PHP .B 501 taaaagagactcaatttgggtcagactggcgacacagagctcagttccagacctcaaccatcgggagacctcccgagccctcaggttggggatct  
 AAV9 501 taaaagagactcaatttgggtcagactggcgacacagagctcagttccagacctcaaccatcgggagacctcccgagccctcaggttggggatct  
 hu68 501 taaaagagagactcaatttgggtcagactggcgacacagagctcagttccagacctcaaccatcgggagacctcccgagccctcaggttggggatct

FIG. 12A ( Ciado F aa)

	<b>  vp1 inicio</b>	
AAV95 1	MSFVDPHPDMLLE - VEGELREFLGEAGPKPKFNQHQDQARGLVLPGNYLGFNGELDRGEPVNRADDEVAREHHDISVNEQLEAGDMFPYLKYNHADAEP	
AAV9 1	MAADGYLPDWLEEDNLSEGIREMVALKPGAPQKANKQHQHONARGVLVPGYKYLSPNGELDKGEPVNAADAALAHDKAYDQQLKAGDNPYLKYNHADAEP	
PHP. B 1	MAADGYLPDWLEEDNLSEGIREMVALKPGAPQKANKQHQHONARGVLVPGYKYLSPNGELDKGEPVNEADAALAHDKAYDQQLKAGDNPYLKYNHADAEP	
	<b>  vp2 inicio</b>	
AAV95 101	QEKLADDTSEGNLQKAVFOAKKXVLEFFGLVEEGAKTAPFGKRIDDHFPKRKARTEED - - - SKPS - - - - - TSSDAEAGPSGSSQLQIIPQAPASSLG	
AAV9 101	QERLKEDTSEGNLGRAVFOAKKXVLEFFGLVEEAAKTAFGKRFVVEQSPQEPDSSAGTIGKSGAQPAKRRKLNFGQGTDTES - VPDFPQIIGEPPAAPSPGVG	
PHP. B 101	QERLKEDTSEGNLGRAVFOAKKXVLEFFGLVEEAAKTAFGKRFVVEQSPQEPDSSAGTIGKSGAQPAKRRKLNFGQGTDTES - VPDFPQIIGEPPAAPSPGVG	
	<b>  vp3 inicio</b>	
AAV95 201	ADTMASGGGAPVAADNNEGADGVGSSSGNWHCDSQWLGDRLVITTSRTWALPTYNHLYKQISNSTSGGSSNDNAVFGYSTPWGTFDNRPHCHFSPRDQ	
AAV9 201	SLTMASGGGAPVAADNNEGADGVGSSSGNWHCDSQWLGDRLVITTSRTWALPTYNHLYKQISNSTSGGSSNDNAVFGYSTPWGTFDNRPHCHFSPRDQ	
PHP. B 201	SLTMASGGGAPVAADNNEGADGVGSSSGNWHCDSQWLGDRLVITTSRTWALPTYNHLYKQISNSTSGGSSNDNAVFGYSTPWGTFDNRPHCHFSPRDQ	
AAV95 301	RLINNNWGFPRKRLNFKLFNIQVKEVTDNNGVKTIANNLTSVQVFTDSYQLPVLGSAHESECLPPFPADVFMIPOYGYLTLNDGSAVGRSSFYCLEY	
AAV9 301	RLINNNWGFPRKRLNFKLFNIQVKEVTDNNGVKTIANNLTSVQVFTDSYQLPVLGSAHESECLPPFPADVFMIPOYGYLTLNDGSAVGRSSFYCLEY	
PHP. B 301	RLINNNWGFPRKRLNFKLFNIQVKEVTDNNGVKTIANNLTSVQVFTDSYQLPVLGSAHESECLPPFPADVFMIPOYGYLTLNDGSAVGRSSFYCLEY	
AAV95 401	FPSQMLRTGNNFQFSYEFENVPFHSSVAHSQSLDRMLNPLIDQYLYLSKTINGSGQOQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTVTTQNNNS	
AAV9 401	FPSQMLRTGNNFQFSYEFENVPFHSSVAHSQSLDRMLNPLIDQYLYLSKTINGSGQOQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTVTTQNNNS	
PHP. B 401	FPSQMLRTGNNFQFSYEFENVPFHSSVAHSQSLDRMLNPLIDQYLYLSKTINGSGQOQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTVTTQNNNS	
AAV95 501	EFAPFGASSWALANGRNSLNNPGPAMASHKEGEDRFPFLSGSLIFGKQGTGRDNDVADKVMITNEEIKTTPVATESYGOVATNHQSAQ - - - - - AQAQ	
AAV9 501	EFAPFGASSWALANGRNSLNNPGPAMASHKEGEDRFPFLSGSLIFGKQGTGRDNDVADKVMITNEEIKTTPVATESYGOVATNHQSAQ - - - - - AQAQ	
PHP. B 501	EFAPFGASSWALANGRNSLNNPGPAMASHKEGEDRFPFLSGSLIFGKQGTGRDNDVADKVMITNEEIKTTPVATESYGOVATNHQSAQ - - - - - AQAQ	
AAV95 601	TGWVQNGQILPGWVWQDRDVIYLGQPIWAKIPIHTDGNFHPSPLMGFGMKHPPFOILIKNTVPADPPTAFNKDKLMSFITQYSTGQVSVIEIWELOKENS	
AAV9 601	TGWVQNGQILPGWVWQDRDVIYLGQPIWAKIPIHTDGNFHPSPLMGFGMKHPPFOILIKNTVPADPPTAFNKDKLMSFITQYSTGQVSVIEIWELOKENS	
PHP. B 601	TGWVQNGQILPGWVWQDRDVIYLGQPIWAKIPIHTDGNFHPSPLMGFGMKHPPFOILIKNTVPADPPTAFNKDKLMSFITQYSTGQVSVIEIWELOKENS	
AAV95 701	KRWNPFIQYTSNYKSNVFEAVNTBGVYSEPRPIGTRYLTRNL	
AAV9 701	KRWNPFIQYTSNYKSNVFEAVNTBGVYSEPRPIGTRYLTRNL	
PHP. B 701	KRWNPFIQYTSNYKSNVFEAVNTBGVYSEPRPIGTRYLTRNLMAADGYLPDWLEEDNLSEGIREMVALKPGAPQKANKQHQHONARGVLVPGYKYLGP	

FIG. 12C (Clado F nt - cont)

AAVG5 601 gatacaatggcttcaggtggggcaccagtcggcagacaataaacgaaggtgccgatgggttagtctcctcgggaaattggcattcggattccccat  
 PHP. B 601 cttacaatggcttcaggtggggcaccagtcggcagacaataaacgaaggtgccgatgggttagtctcctcgggaaattggcattcggattccccat  
 AAV9 601 cttacaatggcttcaggtggggcaccagtcggcagacaataaacgaaggtgccgatgggttagtctcctcgggaaattggcattcggattccccat  
 hu68 601 cttacaatggcttcaggtggggcaccagtcggcagacaataaacgaaggtgccgatgggttagtctcctcgggaaattggcattcggattccccat  
  
 AAVG5 701 ggctgggggacagagtcatacaccacagcaccgaaacctgggcccctgccacccaataacacctacaagcaaaatctccaacagcaatctcgggagg  
 PHP. B 701 ggctgggggacagagtcatacaccacagcaccgaaacctgggcccctgccacccaataacacctacaagcaaaatctccaacagcaatctcgggagg  
 AAV9 701 ggctgggggacagagtcatacaccacagcaccgaaacctgggcccctgccacccaataacacctacaagcaaaatctccaacagcaatctcgggagg  
 hu68 701 ggctgggggacagagtcatacaccacagcaccgaaacctgggcccctgccacccaataacacctacaagcaaaatctccaacagcaatctcgggagg  
  
 AAVG5 801 atcttcaaatgacaacgcctacttgggtacagcaccacctgggggtatcttgaacttaaacagattccactgccactctcaaccctgactggcagcga  
 PHP. B 801 atcttcaaatgacaacgcctacttgggtacagcaccacctgggggtatcttgaacttaaacagattccactgccactctcaaccctgactggcagcga  
 AAV9 801 atcttcaaatgacaacgcctacttgggtacagcaccacctgggggtatcttgaacttaaacagattccactgccactctcaaccctgactggcagcga  
 hu68 801 atcttcaaatgacaacgcctacttgggtacagcaccacctgggggtatcttgaacttaaacagattccactgccactctcaaccctgactggcagcga  
  
 AAVG5 901 ctcatcaacaacaactccgggctaaaggagctcaactcaagctcttcaaacattcaggtcaaaagggttacgggacaacaataaggagtcagagacca  
 PHP. B 901 ctcatcaacaacaactccgggctaaaggagctcaactcaagctcttcaaacattcaggtcaaaagggttacgggacaacaataaggagtcagagacca  
 AAV9 901 ctcatcaacaacaactccgggctaaaggagctcaactcaagctcttcaaacattcaggtcaaaagggttacgggacaacaataaggagtcagagacca  
 hu68 901 ctcatcaacaacaactccgggctaaaggagctcaactcaagctcttcaaacattcaggtcaaaagggttacgggacaacaataaggagtcagagacca  
  
 AAVG5 1001 tcgccataaacttaccagcaggtccaggtcttcaaggactcagactatcagctcccgtacgtgctgggtcgggtcagaggggtgectcccgcggtt  
 PHP. B 1001 tcgccataaacttaccagcaggtccaggtcttcaaggactcagactatcagctcccgtacgtgctgggtcgggtcagaggggtgectcccgcggtt  
 AAV9 1001 tcgccataaacttaccagcaggtccaggtcttcaaggactcagactatcagctcccgtacgtgctgggtcgggtcagaggggtgectcccgcggtt  
 hu68 1001 tcgccataaacttaccagcaggtccaggtcttcaaggactcagactatcagctcccgtacgtgctgggtcgggtcagaggggtgectcccgcggtt  
  
 AAVG5 1101 ccagcgggacgttttccatgattcctcagtcaggggtatctgccttaataatgaaagccagcggcgtgggttcgttccttttactgctcgtggaatatttc  
 PHP. B 1101 ccagcgggacgttttccatgattcctcagtcaggggtatctgccttaataatgaaagccagcggcgtgggttcgttccttttactgctcgtggaatatttc  
 AAV9 1101 ccagcgggacgttttccatgattcctcagtcaggggtatctgccttaataatgaaagccagcggcgtgggttcgttccttttactgctcgtggaatatttc  
 hu68 1101 ccagcgggacgttttccatgattcctcagtcaggggtatctgccttaataatgaaagccagcggcgtgggttcgttccttttactgctcgtggaatatttc  
  
 AAVG5 1201 ccgtcgcaaatgctaaagaacgggtaacaactccagttcagtaacgcttggagaaagctacccttccatagcagctacgctcagcccaaaagcctggacc  
 PHP. B 1201 ccgtcgcaaatgctaaagaacgggtaacaactccagttcagtaacgcttggagaaagctacccttccatagcagctacgctcagcccaaaagcctggacc  
 AAV9 1201 ccgtcgcaaatgctaaagaacgggtaacaactccagttcagtaacgcttggagaaagctacccttccatagcagctacgctcagcccaaaagcctggacc  
 hu68 1201 ccgtcgcaaatgctaaagaacgggtaacaactccagttcagtaacgcttggagaaagctacccttccatagcagctacgctcagcccaaaagcctggacc

FIG. 12D (Clado F nt - cont)

AAV95 1301 gactaatgaatccactcactgaccataacttgaactatctcctcaagactattaacgggttcttggacagagaatcaaaaacgctaaaaatccagtgtggccgg  
 PHP.B 1301 gactaatgaatccactcactcactgaccataacttgaactatctcctcagaaactatcaaggttcttggacagagaatcaaaaacgctaaaaatccagtgtggccgg  
 AAV9 1301 gactaatgaatccactcactgaccataacttgaactatctcctcaagactattaacgggttcttggacagagaatcaaaaacgctaaaaatccagtgtggccgg  
 hu68 1301 gactcaatgaatccactcactgaccataacttgaactatctcctcaagactattaacgggttcttggacagagaatcaaaaacgctaaaaatccagtgtggccgg  
 AAV95 1401 acccagcaaacatggctgtccagggaagaactacatacactggaccctaccggaacaacactgtctcaaacactgtgactcaaaaacaacaacagcgaa  
 PHP.B 1401 acccagcaaacatggctgtccagggaagaactacatacactggaccctaccggaacaacactgtctcaaacactgtgactcaaaaacaacaacagcgaa  
 AAV9 1401 acccagcaaacatggctgtccagggaagaactacatacactggaccctaccggaacaacactgtctcaaacactgtgactcaaaaacaacaacagcgaa  
 hu68 1401 acccagcaaacatggctgtccagggaagaactacatacactggaccctaccggaacaacactgtctcaaacactgtgactcaaaaacaacaacagcgaa  
 AAV95 1501 ttctgttggcctggagcttcttcttgggctctcaatggacgttaatagcttgaatccttggaccctgctatggccaccacaagaaggagagaccggtt  
 PHP.B 1501 ttctgttggcctggagcttcttcttgggctctcaatggacgttaatagcttgaatccttggaccctgctatggccaccacaagaaggagagaccggtt  
 AAV9 1501 ttctgttggcctggagcttcttcttgggctctcaatggacgttaatagcttgaatccttggaccctgctatggccaccacaagaaggagagaccggtt  
 hu68 1501 ttctgttggcctggagcttcttcttgggctctcaatggacgttaatagcttgaatccttggaccctgctatggccaccacaagaaggagagaccggtt  
 AAV95 1601 tctttcctttgtctggatctttaaatttttggcaacaaggaaactggaagacaacactggatgggacaacaaagtcatgataaccaacaagaagaataata  
 PHP.B 1601 tctttcctttgtctggatctttaaatttttggcaacaaggaaactggaagacaacactggatgggacaacaaagtcatgataaccaacaagaagaataata  
 AAV9 1601 tctttcctttgtctggatctttaaatttttggcaacaaggaaactggaagacaacactggatgggacaacaaagtcatgataaccaacaagaagaataata  
 hu68 1601 tctttcctttgtctggatctttaaatttttggcaacaaggaaactggaagacaacactggatgggacaacaaagtcatgataaccaacaagaagaataata  
 AAV95 1701 aactactaacccggtagcaacgggagctcctatggacaagtggccacaaccaccagagtgcccaag-----caacaggccagagacc  
 PHP.B 1701 aactactaacccggtagcaacgggagctcctatggacaagtggccacaaccaccagagtgcccaactttggcgggtgctttaaaggcacaggccagagacc  
 AAV9 1701 aactactaacccggtagcaacgggagctcctatggacaagtggccacaaccaccagagtgcccaag-----caacaggccagagacc  
 hu68 1701 aactactaacccggtagcaacgggagctcctatggacaagtggccacaaccaccagagtgcccaag-----caacaggccagagacc  
 AAV95 1801 ggctgggttcaaaaacaaaggaaatacttccgggtatggtttggcaggacagagatgtgtaacctgcaagaccatttgggccaataatcctcacacgggagc  
 PHP.B 1801 ggctgggttcaaaaacaaaggaaatacttccgggtatggtttggcaggacagagatgtgtaacctgcaagaccatttgggccaataatcctcacacgggagc  
 AAV9 1801 ggctgggttcaaaaacaaaggaaatacttccgggtatggtttggcaggacagagatgtgtaacctgcaagaccatttgggccaataatcctcacacgggagc  
 hu68 1801 ggctgggttcaaaaacaaaggaaatacttccgggtatggtttggcaggacagagatgtgtaacctgcaagaccatttgggccaataatcctcacacgggagc

FIG. 12E (Clado F - cont)

AAVG5 1901 gcaactttcacccttcccgctgctgagtgagggtttgggaatgaagcaccgcccctccctcagatcctcatcaaaaaacacacctgtacctgggatcctccaac  
 PHP.B 1901 gcaactttcacccttcccgctgctgagtgagggtttgggaatgaagcaccgcccctccctcagatcctcatcaaaaaacacacctgtacctgggatcctccaac  
 AAV9 1901 gcaactttcacccttcccgctgctgagtgagggtttgggaatgaagcaccgcccctccctcagatcctcatcaaaaaacacacctgtacctgggatcctccaac  
 hu68 1901 gcaactttcacccttcccgctgctgagtgagggtttgggaatgaagcaccgcccctccctcagatcctcatcaaaaaacacacctgtacctgggatcctccaac  
  
 AAVG5 2001 ggccttcaacaaggacaagctgaactcttccatccccagtattctactggccaagtcacgctggagatcgagctgggagctgcagaaggaaaaacagcaag  
 PHP.B 2001 ggccttcaacaaggacaagctgaactcttccatccccagtattctactggccaagtcacgctggagatcgagctgggagctgcagaaggaaaaacagcaag  
 AAV9 2001 ggccttcaacaaggacaagctgaactcttccatccccagtattctactggccaagtcacgctggagatcgagctgggagctgcagaaggaaaaacagcaag  
 hu68 2001 ggccttcaacaaggacaagctgaactcttccatccccagtattctactggccaagtcacgctggagatcgagctgggagctgcagaaggaaaaacagcaag  
  
 AAVG5 2101 cgtggaaacccgggagatccagtcacacttccaactattacaagttataatgttgaatttgctgttaactgaaggtgtatatagtgaaaccccccccc  
 PHP.B 2101 cgtggaaacccgggagatccagtcacacttccaactattacaagttataatgttgaatttgctgttaactgaaggtgtatatagtgaaaccccccccc  
 AAV9 2101 cgtggaaacccgggagatccagtcacacttccaactattacaagttataatgttgaatttgctgttaactgaaggtgtatatagtgaaaccccccccc  
 hu68 2101 cgtggaaacccgggagatccagtcacacttccaactattacaagttataatgttgaatttgctgttaactgaaggtgtatatagtgaaaccccccccc  
  
 AAVG5 2201 ttggcaccagataacctgactcgttaactctg  
 PHP.B 2201 ttggcaccagataacctgactcgttaactctg  
 AAV9 2201 ttggcaccagataacctgactcgttaactctg  
 hu68 2201 ttggcaccagataacctgactcgttaactctg



FIG. 13A (AAV8)

	<b>  vp1 inicio</b>	
AAV8T 1	MAADGYL	PDWLEDNLS
AAV8T 1	MAADGYL	PDWLEDNLS
AAV8T 1	MAADGYL	PDWLEDNLS
AAV8T 101	QERLQEDTS	FGGNLGRAV
AAV8T 101	QERLQEDTS	FGGNLGRAV
AAV8T 101	QERLQEDTS	FGGNLGRAV
	<b>  vp3 inicio</b>	
AAV8T 201	PNTMAAGGG	APMADNNEGAD
AAV8T 201	PNTMAAGGG	APMADNNEGAD
AAV8T 201	PNTMAAGGG	APMADNNEGAD
AAV8T 301	RLINNNW	GFRPKRLS
AAV8T 301	RLINNNW	GFRPKRLS
AAV8T 301	RLINNNW	GFRPKRLS
AAV8T 401	FPSQMLR	TGNNFQ
AAV8T 401	FPSQMLR	TGNNFQ
AAV8T 401	FPSQMLR	TGNNFQ
AAV8T 501	SNFAWT	AGTKYHLN
AAV8T 501	SNFAWT	AGTKYHLN
AAV8T 501	SNFAWT	AGTKYHLN
AAV8T 601	QGALP	GMVWQNR
AAV8T 601	QGALP	GMVWQNR
AAV8T 601	QGALP	GMVWQNR
AAV8T 701	IQYTSN	YKYKST
AAV8T 701	IQYTSN	YKYKST
AAV8T 701	IQYTSN	YKYKST

MAADGYL PDWLEDNLS EGIREWAL KFGAPKPKANQQKDDGRGLV LFGYKYL GPFNGLDKGEPVNAADA AAL EHDKA YDQQLQAGDN PYLRYNHADA EF  
 MAADGYL PDWLEDNLS EGIREWAL KFGAPKPKANQQKDDGRGLV LFGYKYL GPFNGLDKGEPVNAADA AAL EHDKA YDQQLQAGDN PYLRYNHADA EF  
 MAADGYL PDWLEDNLS EGIREWAL KFGAPKPKANQQKDDGRGLV LFGYKYL GPFNGLDKGEPVNAADA AAL EHDKA YDQQLQAGDN PYLRYNHADA EF  
 QERLQEDTS FGGNLGRAV FQAKKRVL EPLGLVEGAKTAPGKKRPVEPSPQRS PDSSTGIGKKGQ PAKRKLNFQGTGDSSEVPDPQPLGEPAPA PSGVG  
 QERLQEDTS FGGNLGRAV FQAKKRVL EPLGLVEGAKTAPGKKRPVEPSPQRS PDSSTGIGKKGQ PAKRKLNFQGTGDSSEVPDPQPLGEPAPA PSGVG  
 QERLQEDTS FGGNLGRAV FQAKKRVL EPLGLVEGAKTAPGKKRPVEPSPQRS PDSSTGIGKKGQ PAKRKLNFQGTGDSSEVPDPQPLGEPAPA PSGVG  
 PNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSRTWALPTYNHLYKQISSGTH - GAINDN TYFGYSTPWGYDFNRFHCHFS PRDWQ  
 PNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSRTWALPTYNHLYKQISSGTH - GAINDN TYFGYSTPWGYDFNRFHCHFS PRDWQ  
 PNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSRTWALPTYNHLYKQISSGTH - GAINDN TYFGYSTPWGYDFNRFHCHFS PRDWQ  
 RLINNNW GFRPKRLS FKL FNIQVKEVTQNEGKTTIANNLTSTIQVFTDSEYQLPYVILGSAHQGCLPPFPADVFMIPQYGYLTLNNGSAVGRSSFYCLEY  
 RLINNNW GFRPKRLS FKL FNIQVKEVTQNEGKTTIANNLTSTIQVFTDSEYQLPYVILGSAHQGCLPPFPADVFMIPQYGYLTLNNGSAVGRSSFYCLEY  
 RLINNNW GFRPKRLS FKL FNIQVKEVTQNEGKTTIANNLTSTIQVFTDSEYQLPYVILGSAHQGCLPPFPADVFMIPQYGYLTLNNGSAVGRSSFYCLEY  
 FPSQMLRTGNNFQFTYTFEDVPPFHSSYAHSSQSLDRLANPLIDQYLYLSRTQTGGSRPTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNN  
 FPSQMLRTGNNFQFTYTFEDVPPFHSSYAHSSQSLDRLANPLIDQYLYLSRTQTGGSRPTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNN  
 FPSQMLRTGNNFQFTYTFEDVPPFHSSYAHSSQSLDRLANPLIDQYLYLSRTQTGGSRPTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNN  
 SNFAWTAGTKYHLNGRNSLANPGLAMATHKDDERFPFSSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGI VGDNLQLYNTAPGSVVFVNS  
 SNFAWTAGTKYHLNGRNSLANPGLAMATHKDDERFPFSSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGI VGDNLQLYNTAPGSVVFVNS  
 SNFAWTAGTKYHLNGRNSLANPGLAMATHKDDERFPFSSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGI VGDNLQLYNTAPGSVVFVNS  
 QGALP GMVWQNRDVYLQGF IWAKI PHTDGNFHPSP LMGGFGLKHPPPQ ILIKNTVPADPPTTFNQSKLNSFI TOYSTQGVSE IEWELQKENS KRWNPE  
 QGALP GMVWQNRDVYLQGF IWAKI PHTDGNFHPSP LMGGFGLKHPPPQ ILIKNTVPADPPTTFNQSKLNSFI TOYSTQGVSE IEWELQKENS KRWNPE  
 QGALP GMVWQNRDVYLQGF IWAKI PHTDGNFHPSP LMGGFGLKHPPPQ ILIKNTVPADPPTTFNQSKLNSFI TOYSTQGVSE IEWELQKENS KRWNPE  
 IQYTSN YKYKSTVDFEAVNTEGVSEPRPIGTRYLTRNL  
 IQYTSN YKYKSTVDFEAVNTEGVSEPRPIGTRYLTRNL  
 IQYTSN YKYKSTVDFEAVNTEGVSEPRPIGTRYLTRNL

FIG. 13B (AAV8)

	vp1 inicio	
AAV8T 1	atggctgccgatgggttatctccagattggctcgaggacaacctctctgagggcattcgcgagtggtgggctgaaacctggagccccgaagccccaaaag	
AAVG3 1	atggctgccgatgggttatctccagattggctcgaggacaacctctctgagggcattcgcgagtggtgggctgaaacctggagccccgaagccccaaaag	
AAV8 1	atggctgccgatgggttatctccagattggctcgaggacaacctctctgagggcattcgcgagtggtgggctgaaacctggagccccgaagccccaaaag	
AAV8T 101	ccaaccagcaaaaagcaggacgcggccggggtctcgggtgctccctgggtacaagtacctcggaccctcaacggactcgacaaggggggagccccgtcaacgc	
AAVG3 101	ccaaccagcaaaaagcaggacgcggccggggtctcgggtgctccctgggtacaagtacctcggaccctcaacggactcgacaaggggggagccccgtcaacgc	
AAV8 101	ccaaccagcaaaaagcaggacgcggccggggtctcgggtgctccctgggtacaagtacctcggaccctcaacggactcgacaaggggggagccccgtcaacgc	
AAV8T 201	ggcggacgcagcggccctcgagcgcgacaaggctcgcaccagcagctgcaggcgggtgacaatcctcctggtacatccgtaaccacgcgcgcgagcttt	
AAVG3 201	ggcggacgcagcggccctcgagcgcgacaaggctcgcaccagcagctgcaggcgggtgacaatcctcctcctggtacatccgtaaccacgcgcgcgagcttt	
AAV8 201	ggcggacgcagcggccctcgagcgcgacaaggctcgcaccagcagctgcaggcgggtgacaatcctcctggtacatccgtaaccacgcgcgcgagcttt	
AAV8T 301	caggagcgtctcgaagaagatacgtctcttgggggcaacctcggggcagcagctctccaggccaagaagcgggtctcggaaacctcggctcggttgagg	
AAVG3 301	caggagcgtctcgaagaagatacgtctcttgggggcaacctcggggcagcagctctccaggccaagaagcgggtctcggaaacctcggctcggttgagg	
AAV8 301	caggagcgtctcgaagaagatacgtctcttgggggcaacctcggggcagcagctctccaggccaagaagcgggtctcggaaacctcggctcggttgagg	
	vp2 inicio	
AAV8T 401	aaggcgtcaagcaggtcctcctggaaagaagagaccggtagagcctaccaccaggttctccagactctcaccgggcatcgggcaagaaggcccaacagcc	
AAVG3 401	aaggcgtcaagcaggtcctcctggaaagaagagaccggtagagcctaccaccaggttctccagactctcaccgggcatcgggcaagaaggcccaacagcc	
AAV8 401	aaggcgtcaagcaggtcctcctggaaagaagagaccggtagagcctaccaccaggttctccagactctcaccgggcatcgggcaagaaggcccaacagcc	
AAV8T 501	cgccagaaaaagactcaatttgggtcagactggcgactcagactcaggttcagaccctcaacctcgggagaacctccagcagcccccttgggtgaggga	
AAVG3 501	cgccagaaaaagactcaatttgggtcagactggcgactcagactcaggttcagaccctcaacctcgggagaacctccagcagcccccttgggtgaggga	
AAV8 501	cgccagaaaaagactcaatttgggtcagactggcgactcagactcaggttcagaccctcaacctcgggagaacctccagcagcccccttgggtgaggga	
	vp3 inicio	
AAV8T 601	cctaatacaattggctgcaggcgtggcgaccacatggcagacaataaacgaaggcgccgagtggttagttcctcggggaattggcattgcgattcca	
AAVG3 601	cctaatacaattggctgcaggcgtggcgaccacatggcagacaataaacgaaggcgccgagtggttagttcctcggggaattggcattgcgattcca	
AAV8 601	cctaatacaattggctgcaggcgtggcgaccacatggcagacaataaacgaaggcgccgagtggttagttcctcggggaattggcattgcgattcca	
AAV8T 701	catggctggcgcagagtgctaccaccagcaccctgggccccgccacatacaaacaccactctcaagaataatcctcctcgggtactca ---	
AAVG3 701	catggctggcgcagagtgctaccaccagcaccctgggccccgccacatacaaacaccactctcaagaataatcctcctcgggtactca ---	
AAV8 701	catggctggcgcagagtgctaccaccagcaccctgggccccgccacatacaaacaccactctcaagaataatcctcctcgggtactca ---	
AAV8T 801	tggagccaccaaacgacaacactacttgggtcagcagccccctgggggtatttggactttaacagattccacttccactttccaccagtgactggcag	
AAVG3 801	tggagccaccaaacgacaacactacttgggtcagcagccccctgggggtatttggactttaacagattccacttccactttccaccagtgactggcag	
AAV8 801	tggagccaccaaacgacaacactacttgggtcagcagccccctgggggtatttggactttaacagattccacttccactttccaccagtgactggcag	

FIG. 13C (AAV8 - cont)

AAV8T 901 cgactcaacaacaactgggggattccggcccagaagactcagctcaagctcttcaacatcccaaggtcaaggaggtcaacgcagaatgaaggcaccacaaga  
 AAVG3 901 cgactcaacaacaactgggggattccggcccagaagactcagctcaagctcttcaacatcccaaggtcaaggaggtcaacgcagaatgaaggcaccacaaga  
 AAV8 901 cgactcaacaacaactgggggattccggcccagaagactcagctcaagctcttcaacatcccaaggtcaaggaggtcaacgcagaatgaaggcaccacaaga  
  
 AAV8T 1001 ccatacgaataaactccaccagaccatccaggtgtttacggactcggagtagcagctgacctgctgctctcctggctcctggcccaacagggctgctgctcctcc  
 AAVG3 1001 ccatacgaataaactccaccagaccatccaggtgtttacggactcggagtagcagctgacctgctgctctcctggctcctggcccaacagggctgctgctcctcc  
 AAV8 1001 ccatacgaataaactccaccagaccatccaggtgtttacggactcggagtagcagctgacctgctgctctcctggctcctggcccaacagggctgctgctcctcc  
  
 AAV8T 1101 gttccggcggagcgtgttcatgattccccagtagcggtaacctaaactcaacaacggttagtcaggccgtggagcgtcctctctctactcctgggataac  
 AAVG3 1101 gttccggcggagcgtgttcatgattccccagtagcggtaacctaaactcaacaacggttagtcaggccgtggagcgtcctctctctactcctgggataac  
 AAV8 1101 gttccggcggagcgtgttcatgattccccagtagcggtaacctaaactcaacaacggttagtcaggccgtggagcgtcctctctctactcctgggataac  
  
 AAV8T 1201 tttccttcgcagatgctgagaaccggcaaacacttccagtttacttacaccttcgaggacgtgcttccacagcagctacgcccacagccagagcttgg  
 AAVG3 1201 tttccttcgcagatgctgagaaccggcaaacacttccagtttacttacaccttcgaggacgtgcttccacagcagctacgcccacagccagagcttgg  
 AAV8 1201 tttccttcgcagatgctgagaaccggcaaacacttccagtttacttacaccttcgaggacgtgcttccacagcagctacgcccacagccagagcttgg  
  
 AAV8T 1301 accggctgatgaatcctctgattgaccagtagctactactgtctcggactcaacaacaggtggagtaggctcagcagactctgggcttccagcca  
 AAVG3 1301 accggctgatgaatcctctgattgaccagtagctactactgtctcggactcaacaacaggtggagtaggctcagcagactctgggcttccagcca  
 AAV8 1301 accggctgatgaatcctctgattgaccagtagctactactgtctcggactcaacaacaggtggagtaggctcagcagactctgggcttccagcca  
  
 AAV8T 1401 aggtggcctaaatacaatggcccaatcaggcaagaactggctgccaaggacctgttaccggccaacacgctcacaacgcaacccgggcaaaaacaacaat  
 AAVG3 1401 aggtggcctaaatacaatggcccaatcaggcaagaactggctgccaaggacctgttaccggccaacacgctcacaacgcaacccgggcaaaaacaacaat  
 AAV8 1401 aggtggcctaaatacaatggcccaatcaggcaagaactggctgccaaggacctgttaccggccaacacgctcacaacgcaacccgggcaaaaacaacaat  
  
 AAV8T 1501 agcaacttgcctggactgctgggcccataaccatctgaatggaagaaattcattggctaatcctggcaatcctggcaacacacaagaacgacgaggg  
 AAVG3 1501 agcaacttgcctggactgctgggcccataaccatctgaatggaagaaattcattggctaatcctggcaatcctggcaacacacaagaacgacgaggg  
 AAV8 1501 agcaacttgcctggactgctgggcccataaccatctgaatggaagaaattcattggctaatcctggcaatcctggcaacacacaagaacgacgaggg  
  
 AAV8T 1601 agcgtttttttcccagtaacgggactcctgatttttggcaaacaaaatgctccagagacaatgctgggattacagcgtatgctcaccagcggagaaga  
 AAVG3 1601 agcgtttttttcccagtaacgggactcctgatttttggcaaacaaaatgctccagagacaatgctgggattacagcgtatgctcaccagcggagaaga  
 AAV8 1601 agcgtttttttcccagtaacgggactcctgatttttggcaaacaaaatgctccagagacaatgctgggattacagcgtatgctcaccagcggagaaga

FIG. 13D (AAV8- cont)

AAV8T 1701	aatcaaaaaaacactaacccctgtggctacagagggaatacgggtatcgtgggtgataaacttgcaattgataaacaaggctcctctgggttgggtgttctcaacagc
AAVG3 1701	aatcaaaaaaacactaacccctgtggctacagagggaatacgggtatcgtgggtgataaacttgcaattgataaacaaggctcctctgggttgggtgttctcaacagc
AAV8 1701	aatcaaaaaaacactaacccctgtggctacagagggaatacgggtatcgtgggtgataaacttgcaattgataaacaaggctcctctgggttgggtgttctcaacagc
AAV8T 1801	cagggggccttaccgggtatggctggcagaaccggggaactgtacctgagggtcccatctgggccaagatlcctcacacgggaacggcaacttccaccgt
AAVG3 1801	cagggggccttaccgggtatggctggcagaaccggggaactgtacctgagggtcccatctgggccaagatlcctcacacgggaacggcaacttccaccgt
AAV8 1801	cagggggccttaccgggtatggctggcagaaccggggaactgtacctgagggtcccatctgggccaagatlcctcacacgggaacggcaacttccaccgt
AAV8T 1901	ctccgctgatggggcgtttggcctgaaaaatcctccgctcagatcctgatcaagaacacgctgtacctgagggtacctgagggtacctccgaccctcaaccagtc
AAVG3 1901	ctccgctgatggggcgtttggcctgaaaaatcctccgctcagatcctgatcaagaacacgctgtacctgagggtacctgagggtacctccgaccctcaaccagtc
AAV8 1901	ctccgctgatggggcgtttggcctgaaaaatcctccgctcagatcctgatcaagaacacgctgtacctgagggtacctgagggtacctccgaccctcaaccagtc
AAV8T 2001	aaagctgaactcttcatcacgcaatacagcaccggacaggtcagcgtggaaattgaaatgggagctgcagaagggaacacagcaagcgcgtggaaacccccgag
AAVG3 2001	aaagctgaactcttcatcacgcaatacagcaccggacaggtcagcgtggaaattgaaatgggagctgcagaagggaacacagcaagcgcgtggaaacccccgag
AAV8 2001	aaagctgaactcttcatcacgcaatacagcaccggacaggtcagcgtggaaattgaaatgggagctgcagaagggaacacagcaagcgcgtggaaacccccgag
AAV8T 2101	atccagtaaacctccaactactacaatactacaagtgtggactttgctgttaatacagaaggcgtgtactctgaacccccccccatggcaccocggttacc
AAVG3 2101	atccagtaaacctccaactactacaatactacaagtgtggactttgctgttaatacagaaggcgtgtactctgaacccccccccatggcaccocggttacc
AAV8 2101	atccagtaaacctccaactactacaatactacaagtgtggactttgctgttaatacagaaggcgtgtactctgaacccccccccatggcaccocggttacc
AAV8T 2201	tcacccgtaaatctg
AAVG3 2201	tcacccgtaaatctg
AAV8 2201	tcacccgtaaatctg

FIG. 14A (AAVrh10)

		<b>  vp1 inicio</b>	
AAVG2 1		MAADGYLFDWLEDNLSEGIREWWDLKP	GAPKPKANQQKDDGRLVLP
rh10 1		MAADGYLFDWLEDNLSEGIREWWDLKP	GAPKPKANQQKDDGRLVLP
hu37 1		MAADGYLFDWLEDNLSEGIREWWDLKP	GAPKPKANQQKDDGRLVLP
			<b>  vp2 inicio</b>
AAVG2 101		QERLQEDTSFGGNLGRAVFOAKKRVLE	PLGLVEAAKTA
rh10 101		QERLQEDTSFGGNLGRAVFOAKKRVLE	PLGLVEEAGAKTAP
hu37 101		QERLQEDTSFGGNLGRAVFOAKKRVLE	PLGLVEEAAKTAP
			<b>  vp3 inicio</b>
AAVG2 201		SGTMAAGGAPMADNNEGADGVGSSGN	WHCDSTWLGDRVITTS
rh10 201		SGTMAAGGAPMADNNEGADGVGSSGN	WHCDSTWLGDRVITTS
hu37 201		SGTMAAGGAPMADNNEGADGVGSSGN	WHCDSTWLGDRVITTS
AAVG2 301		RLINNNWGFRPKRLSFKLFNIQVKEVT	QNEGKTIANNL
rh10 301		RLINNNWGFRPKRLNFKLFNIQVKEVT	QNEGKTIANNL
hu37 301		RLINNNWGFRPKRLSFKLFNIQVKEVT	QNEGKTIANNL
AAVG2 401		FPSQMLRTGNNFEEFSTYFEDVPPHSS	YAHSQSLDRLMNPL
rh10 401		FPSQMLRTGNNFEEFSTYFEDVPPHSS	YAHSQSLDRLMNPL
hu37 401		FPSQMLRTGNNFEEFSTYFEDVPPHSS	YAHSQSLDRLMNPL
AAVG2 501		SNFAWTGATKYHLNAGRDSLVPNGVA	MATHKDDDEERFPSS
rh10 501		SNFAWTGATKYHLNAGRDSLVPNGVA	MATHKDDDEERFPSS
hu37 501		SNFAWTGATKYHLNAGRDSLVPNGVA	MATHKDDDEERFPSS
AAVG2 601		QGALPGMWQNRDVLQGFIAWAKIPHTD	GNFHPSPLMGGFGLKHP
rh10 601		QGALPGMWQNRDVLQGFIAWAKIPHTD	GNFHPSPLMGGFGLKHP
hu37 601		QGALPGMWQNRDVLQGFIAWAKIPHTD	GNFHPSPLMGGFGLKHP
AAVG2 701		IQYTSNYYSKTNVDFAVNTEGTYSE	FRPIGTRYLTRNL
rh10 701		IQYTSNYYSKTNVDFAVNTEGTYSE	FRPIGTRYLTRNL
hu37 701		IQYTSNYYSKTNVDFAVNTEGTYSE	FRPIGTRYLTRNL

FIG. 14B (AAVrh10)

			vp1 inicio	
AAVG2	1			a t g g c t g c t g a c g g t t a t c t t c c a g a t t g g c t c g a g a c a a c c t c t c g a g g g c a t t c g c a g t g g t g g g a c c t g a a c c t g g a c c c c c c a a g c c c a a g g
AAVhu37	1			a t g g c t g c t g a c g g t t a t c t t c c a g a t t g g c t c g a g a c a a c c t c t c g a g a g t g g t g g g a c c t g a a c c t g g a c c c c c c a a g c c c c a a g g
AAVrh10	1			a t g g c t g c c g a t g g t t a t c t t c c a g a t t g g c t c g a g a c a a c c t c t c g a g g g c a t t c g c a g t g g t g g g a c t g a a c c t g g a c c c c c c a a g c c c a a a g
AAVG2	101			c c a a c c a g a a g a g c a g g a c g a c g g g c c g g g t c t g t g t c t c t g g t a c a a g t a c c t c g g a c c t t c a a c g g a c c t c a a c g g g a c t c g a c a a g g g g g a g c c c g t c a a c g a
AAVhu37	101			c c a a c c a g c a g a a g a g a c a g a c g g c c g g g t c t g t g t c t c t g g c t a c a a g t a c c t c g g a c c t t c a a c g g a c t c g a c a a g g g g g a g c c c g t c a a c g c
AAVrh10	101			c c a a c c a g c a a a a g c a g g a c g a c g g c c g g g t c t g t g t c t c t g g c t a c a a g t a c c t c g g a c c t t c a a c g g a c t c g a c a a g g g g a g c c c c g t c a a c g c
AAVG2	201			g g c g g a c g c g c g g c c t c g a g e c a c a a g g c t a c g a c a g c a g t c a a a g c g g g t g a c a a t c c g t a c c t c g g t a t a a c c a c g c c g a c g c c g a g t t t
AAVhu37	201			g g c g g a c g c g c g g c c t c g a g e c a c a a g g c t a c g a c a g c a g t c a a a g c g g g t g a c a a t c c g t a c c t c g g t a t a a c c a c g c c g a c g c c g a g t t t
AAVrh10	201			g g c g g a c g c g c g g c c c t c g a g e a c g a c a a g g c t a c g a c c a g c a g t c a a a g c g g g t g a c a a t c c g t a c c t c g g t a t a a c c a c g c c g a c g c c g a g t t t
AAVG2	301			c a g g a c g t c t g c a a g a a g a t a e g t c t t t t g g g g g e a a c c t e g g g c g a g c a g t t t c c a g g c c a a g a g c g g g t t c t c g a a c c t c t c g g t c t g g t t g a g g
AAVhu37	301			c a g g a c g t c t g c a a g a a g a t a e g t c t t t t g g g g g e a a c c t e g g g c g a g c a g t t t c c a g g c c a a g a g c g g g t t c t c g a a c c t c t c g g t c t g g t t g a g g
AAVrh10	301			c a g g a c g t c t g c a a g a a g a t a e g t c t t t t g g g g g e a a c c t e g g g c g a g c a g t t t c c a g g c c a a g a a g c g g g t t c t c g a a c c t c t c g g t c t g g t t g a g g
AAVG2	401			a a g c t g c t a a g a c g g c t c c t g g a a a g a a g a g a c c g g t a g a a c c g t c a c c t c a g c g a t c c c c g a c t c c t c a c g g g c a t c g g c a a a a a a g g c c a g c a g c c
AAVhu37	401			a a g c t g c t a a g a c g g c t c c t g g a a a g a a g a c c g t a g a a c c g t c a c c t c a g c g t t e c c c e g a c t c c t c a c g g c a t c g g c a a g a a a g g c c a g c a g c c
AAVrh10	401			a a g g c t a a g a c g g c t c c t g g a a a g a a g a g a c c g g t a g a g c c a t c a c c c a g g t t c t c c a g a c t c t e t a c g g g c a t c g g c a a g a a a g g c c a g c a g c c
AAVG2	501			c g c g a a a g a g a c t g a a c t t t g g g c a g a c t g g c g a c t c a g a g t c a g t c c c c g a c c c t c a a c c a a t c g g a g a a c c a c c a g c a g g c c c c t c t g g t c t g g g a
AAVhu37	501			c g t a a a a a g a g a c t g a a c t t t g g t c a g a c t g g c a c t c a g a g t c a g t c c c c g a c c c t c a a c c a a t c g g a a a c c a c c a g c a g c c c c t c t g g t c t g g g a
AAVrh10	501			c g c a a a a a g a g a c t c a a c t t t g g g c a g a c t g g c g a c t c a g a g t c a g t g c c c g a c c c t c a a c c a a t c g g a g a a c c c c c c g a g g c c c c t c t g g t c t g g g a
AAVG2	601			t c t g g t a c a a t g g c t g c a g g c g g t g g c g c t c c a a t g g c a g a a t a a c g a a g g c g c g a c g g a g t g g g t a g t t c c t c a g g a a a t t g g e a t t g c g a t t c c a
AAVhu37	601			t c t g g t a c a a t g g c t g c a g g c g g t g g c g c t c c a a t g g c a g a a t a a c g a g a c a a t a a c g a a g g c g c a c g g a g t g g g t a g t t c c t c a g a a a t t g g e a t t g c g a t t c c a
AAVrh10	601			t c t g g t a c a a t g g c t g c a g g c g g t g g c g c t c c a a t g g c a g a a a t a a c g a a g g c g c g a c g g a g t g g g t a g t t c c t c a g g a a a t t g g e a t t g c g a t t c c a
AAVG2	701			c a t g g c t g g c g a c a g a g t c a t c a c c a c c a g a c c c g a a c c t g g g c c t g c c c a c c t a c a a c c c a c c t c a c a a c c c c t c a a g c a a a t c t c c a a t g g g a c a t c g g g
AAVhu37	701			c a t g g c t g g c g a c a g a g t c a t c a c c a c c a g a c c c g a a c c t g g g c c c t g c c c c c t c a a c a a c c c c t a c a a g c a a a t a t c c a a t g g g a c a t c g g g
AAVrh10	701			c a t g g c t g g c g a c a g a g t c a t c a c c a c c a g a c c c g a a c c t g g g c c c t c c c c a c c c t a c a a c c a c c c t c t a c a a g c a a a t c t c c a a g c g g a c t t c g g g

| vp2 inicio

| vp3 inicio

FIG. 14C (AAVrh10)

AAVG2	801	aggaagcaaccaacgacacactactctgggtacagcaccctggggg-atttgacttcaacagattccacttcccttccaccctgactggcag
AAVrh37	801	aggaagcaaccaacgacacactactctgggtacagcaccctggggg-atttgacttcaacagattccacttcccttccaccctgactggcag
AAVrh10	801	aggaagcaaccaacgacacactactctgggtacagcaccctggggg-atttgacttcaacagattccacttcccttccaccctgactggcag
AAVG2	901	agactcatcaacaactggggaatccggcccaagagactcagctcaagctcttcaacatccaggtaaggaggtcaCGcagaatgaaggccaacaaga
AAVrh37	901	cgactcatcaacaactggggaatccggcccaagagactcagctcaagctcttcaacatccaggtaaggaggtcaCGcagaatgaaggccaacaaga
AAVrh10	901	cgactcatcaacaactggggaatccggcccaagagactcaactcaagctcttcaacatccaggtaaggaggtcaCGcagaatgaaggccaacaaga
AAVG2	1001	ccatcgccaataaccttaaccagcagattcaggattttacggactcgggaataccagctcctcctcctggctccggccaccagggctgcctgcctcc
AAVrh37	1001	ccatcgccaataaccttaaccagcagattcaggattttacggactcgggaataccagctcctcctcctggctccggccaccagggctgcctgcctcc
AAVrh10	1001	ccatcgccaataaccttaaccagcagattcaggattttacggactcgggaataccagctcctcctcctggctccggccaccagggctgcctgcctcc
AAVG2	1101	gttccggcggtgtcttcatgattccccagtaaggctacctaactgaacaacgggaagtcaagccctagggcgttccctcattctactgcctgggaatat
AAVrh37	1101	gttccggcggtgtcttcatgattccccagtaaggctacctaactgaacaacgggaagtcaagccctagggcgttccctcattctactgcctgggaatat
AAVrh10	1101	gttccggcggtgtcttcatgattccccagtaaggctacctaactgaacaacgggaagtcaagccctagggcgttccctcattctactgcctgggaatat
AAVG2	1201	tttccatctcaaatgctgggactggaacaactttgaaattagctacaccttgaggagcgtgcccttccacagcagcaccacacagcagcagcctgg
AAVrh37	1201	tttccatctcaaatgctgggactggaacaactttgaaattagctacaccttgaggagcgtgcccttccacagcagcaccacacagcagcagcctgg
AAVrh10	1201	tttccatctcaaatgctgggactggaacaactttgaaattagctacaccttgaggagcgtgcccttccacagcagcaccacacagcagcagcctgg
AAVG2	1301	accggctgatgaaccctctcatcgaccagtaactgtattactatccagaactcagctccacagggaggaactcaaggtaacacagcaaatgttattttctca
AAVrh37	1301	accggctgatgaaccctctcatcgaccagtaactgtattactatccagaactcagctccacagggaggaactcaaggtaacacagcaaatgttattttctca
AAVrh10	1301	accggctgatgaaccctctcatcgaccagtaactgtattactatccagaactcagctccacagggaggaactcaaggtaacacagcaaatgttattttctca
AAVG2	1401	agccggcctgcaaatatgtcggctcagggcccaagaactggctacctggacttgcctcagcagcagcagctctccacgacactgtcgcaaaacaacaac
AAVrh37	1401	agctggcctgcaaacatgtcggctcagggcccaagaactggctacctggacttgcctcagcagcagcagctctccacgacactgtcgcaaaacaacaac
AAVrh10	1401	gfgccggcctcaataacatgtcggctcagggcccaaaaactggctacccggccctcctacccgggagcaacgctctccacgacactgtcgcaaaataacaac
AAVG2	1501	agcaacttgcctggactggcccaaaaatcatctgaaacgggaagactcttggatgaatcccggtgtgctatggcaacgcatcaaggacgaggg
AAVrh37	1501	agcaacttgcctggactggcccaaaaatcatctgaaacgggaagactcttggatgaatcccggtgtgctatggcaacgcatcaaggacgaggg
AAVrh10	1501	agcaacttgcctggactggcccaaaaatcatctgaaacgggaagactcttggatgaatcccggtgtgctatggcaacgcatcaaggacgaggg

FIG. 14D (AAVrh10)

AAVG2	1601	aacgtttcttccatcgagcggagtcctgatgtttggaacaacagggctctggaagagacaatgtgactatagcagcgttatgctaaccagcgaggaga
AAVhu37	1601	aacgtttcttccatcgagcggagtcctgatgtttggaacaacagggctctggaagagacaatgtgactatagcagcgttatgctaaccagcgaggaga
AAVrh10	1601	agcgtatcttccgtccagcggagtcctaatgtttggaacaacagggctggaagagacaacgtggaactatagcagcgttatgctaaccagcgaggaga
AAVG2	1701	aatcaagaccactaacctgtagccactgaacaatacggcgtggctgataaacttgcagcaaaacccaatacaggacctatcgtgggaaatgtcaaccgc
AAVhu37	1701	aatcaaaaccactaacctgtagccacagaacaatacggctggctgacaacttgcagcaaaacccaatacaggacctatcgtgggaaatgtcaaccgc
AAVrh10	1701	aattcaaaaccacacccagtgccacagaaacagtcaggctggctggccgat.aacctgcaacagcaaaacgcgcctctctattgttaggggcccgtcaaccgt
AAVG2	1801	caaggagccttaacctggcctgttgccagaaacogagacgtgtacctgcaagggtcccatctgggccaagattctcacacggacggcaactttcaaccgt
AAVhu37	1801	caaggagccttaacctggcctgttgccagaaacogagacgtgtacctgcaagggtcccatctgggccaagattctcacacggacggcaactttcaaccgt
AAVrh10	1801	caaggagccttaacctggcctgttgccagaaacogggacgtgtacctgcaagggtcccatctgggccaagattctcacacggacggcaaaactttcaacct
AAVG2	1901	ctcctctgatgggcgcttggactgaaacacccgcctcctcaaaatcctgatcaagaaacactcccttctctgggatctctcaaacgacctcagccaggc
AAVhu37	1901	caccgctaabtgaggagatttggactgaagcaaccaacctctcagatcctgatcaagaaacacgcggctcctcgggatctctcaaacgacctcagccaggc
AAVrh10	1901	cgcctctgatgggagcttggactgaaacacccgcctcctcagatcctgatcaagaaacactcccttctctgggatctctcaaacgacctcagtcaggc
AAVG2	2001	gaaatbgycttccctcctcaocgcaglatagtaaccggccaggctcagcgtggagatcgagtgaggagctgcagaaggagaaacagcaagcgtggaacccagaa
AAVhu37	2001	gaaatbgycttccctcctcaatcagcagtaacagcaccggacaggtcagcgtggaaatcgagtgaggagctgcagaaggagaaacagcaaaicgctggaacccagag
AAVrh10	2001	taagctggcgtcgttcatcaccgcaatcagcaaccggacaggtcagcgtggaaatgaaatgggagctgcagaagaaacagcaaaicgctggaacccagag
AAVG2	2101	attcagtaacttccaaactactaataatctacaatgtggactttgtctgtaaatcccgaggggtacatatccagagcctgcctccatggaaactcgttacc
AAVhu37	2101	attcagtaacttccaaactactaataatctacaatgtggactttgtctgtaaatcccgaggggtacatatccagagcctgcctccatggaaactcgttacc
AAVrh10	2101	attcaatacaacttccaaactactaataatctacaatgtggactttgtctgtaaacacagatggcacttattctgagcctgcctccatggcaaccgttacc
AAVG2	2201	tcaaccgtaactctg
AAVhu37	2201	tcaaccgtaactctg
AAVrh10	2201	tcaaccgtaactctg



FIG. 15A

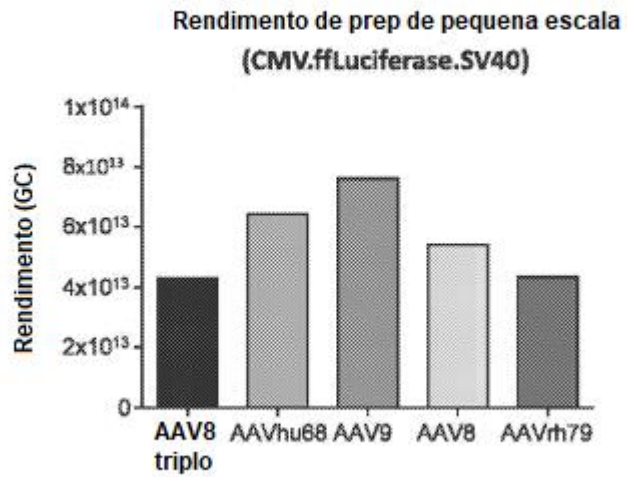


FIG. 15B

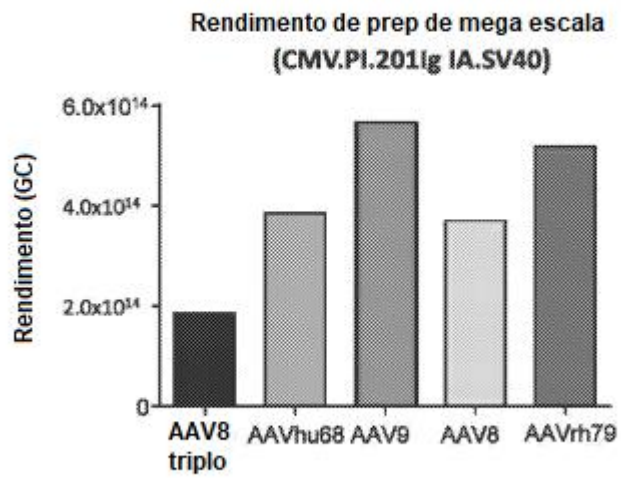


FIG. 16

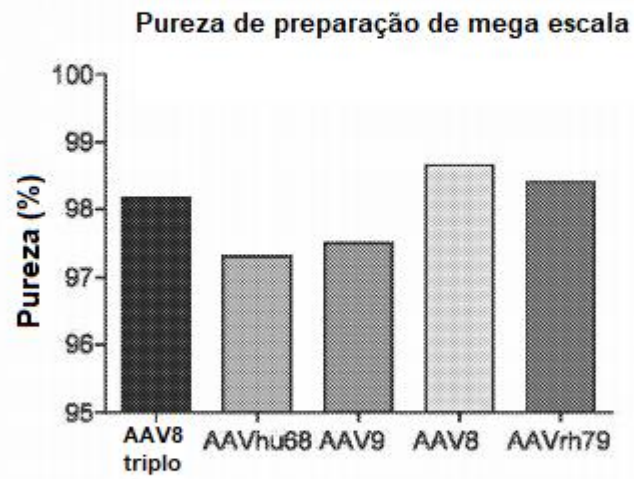


FIG. 17A

Fígado

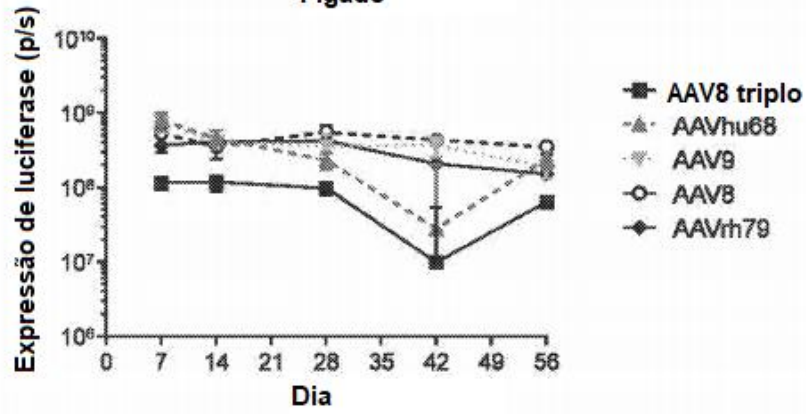


FIG. 17B

Músculo

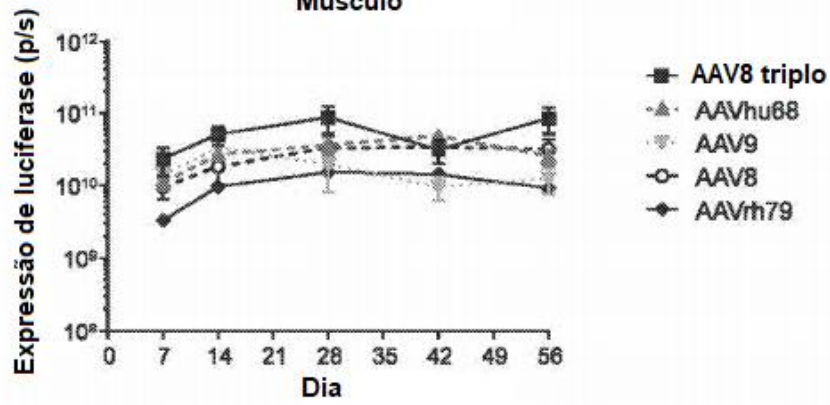


FIG. 17C

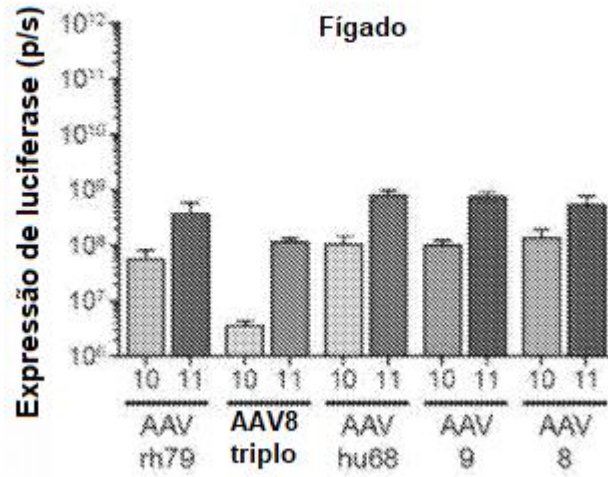


FIG. 17D

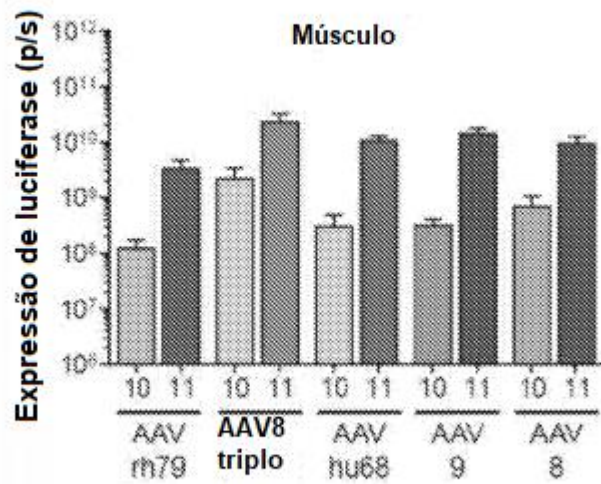


FIG. 17E

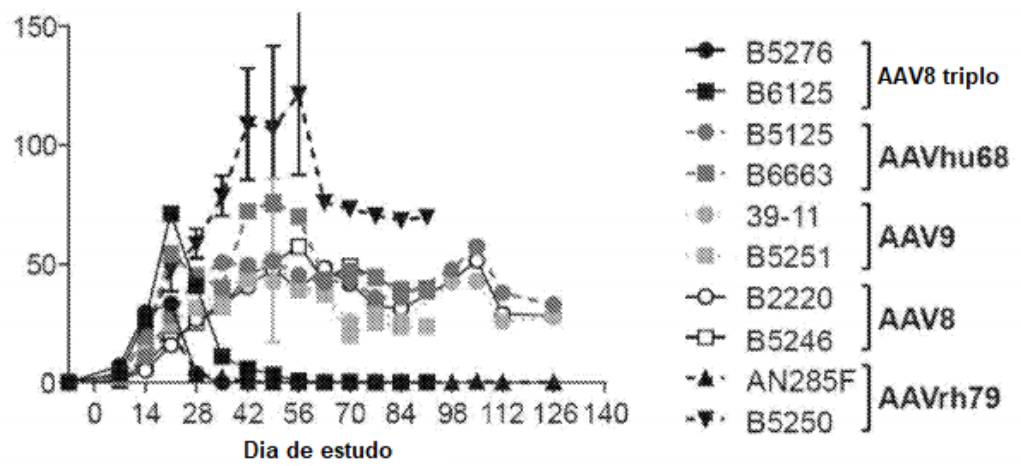
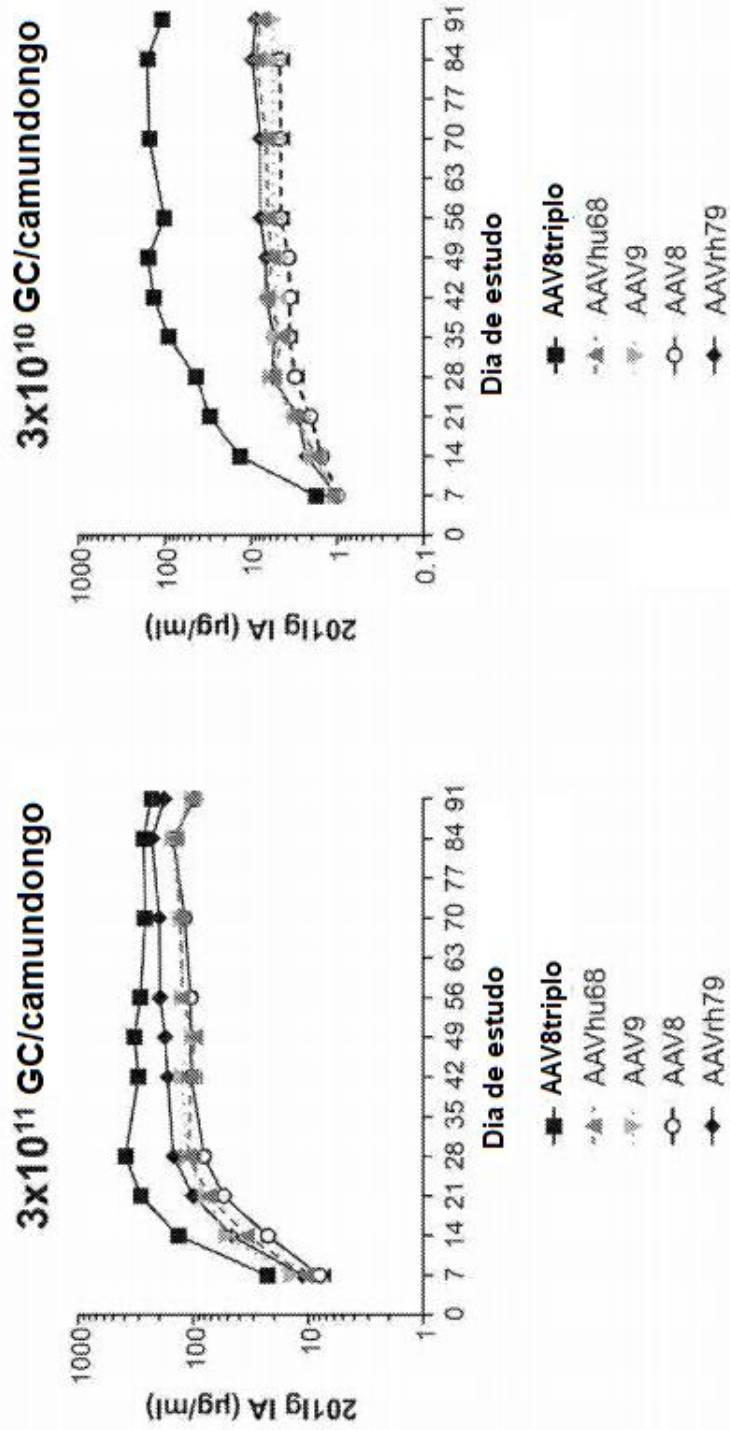
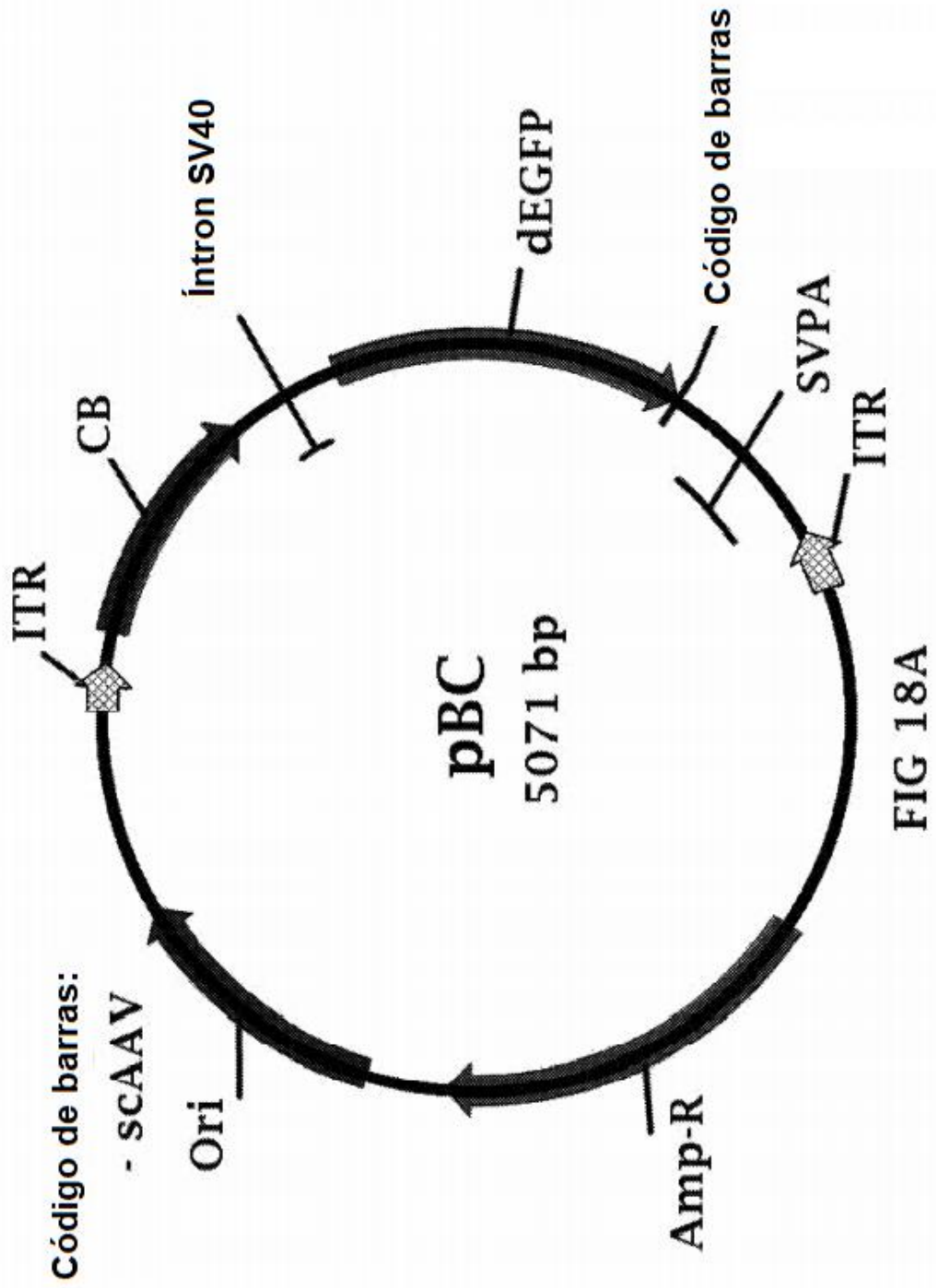


FIG. 17F





- **Teste in vivo, camundongos B6**
- **AAV8.AR2.08, 2e12 gc/camundongo, 2 semanas, i.v.**

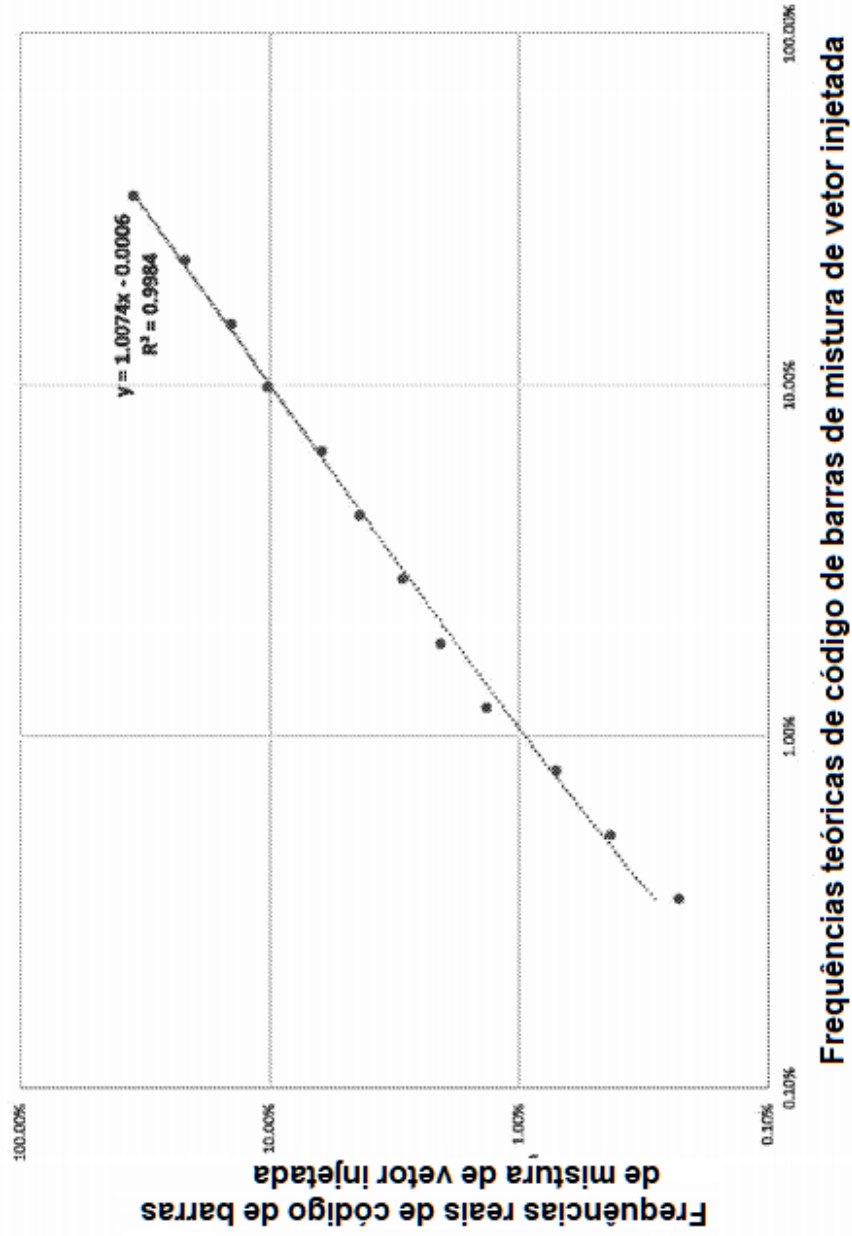
<b>Código</b>	<b>Quantidade por camundongo (GC)</b>
BC01	6.87E+09
BC02	1.04E+10
BC03	1.59E+10
BC04	2.41E+10
BC05	3.67E+10
BC06	5.57E+10
BC07	8.47E+10
BC08	1.29E+11
BC09	1.96E+11
BC10	2.97E+11
BC11	4.52E+11
BC12	6.87E+11
<b>Total de gc da mistura</b>	<b>2.00E+12</b>

FIG 18B

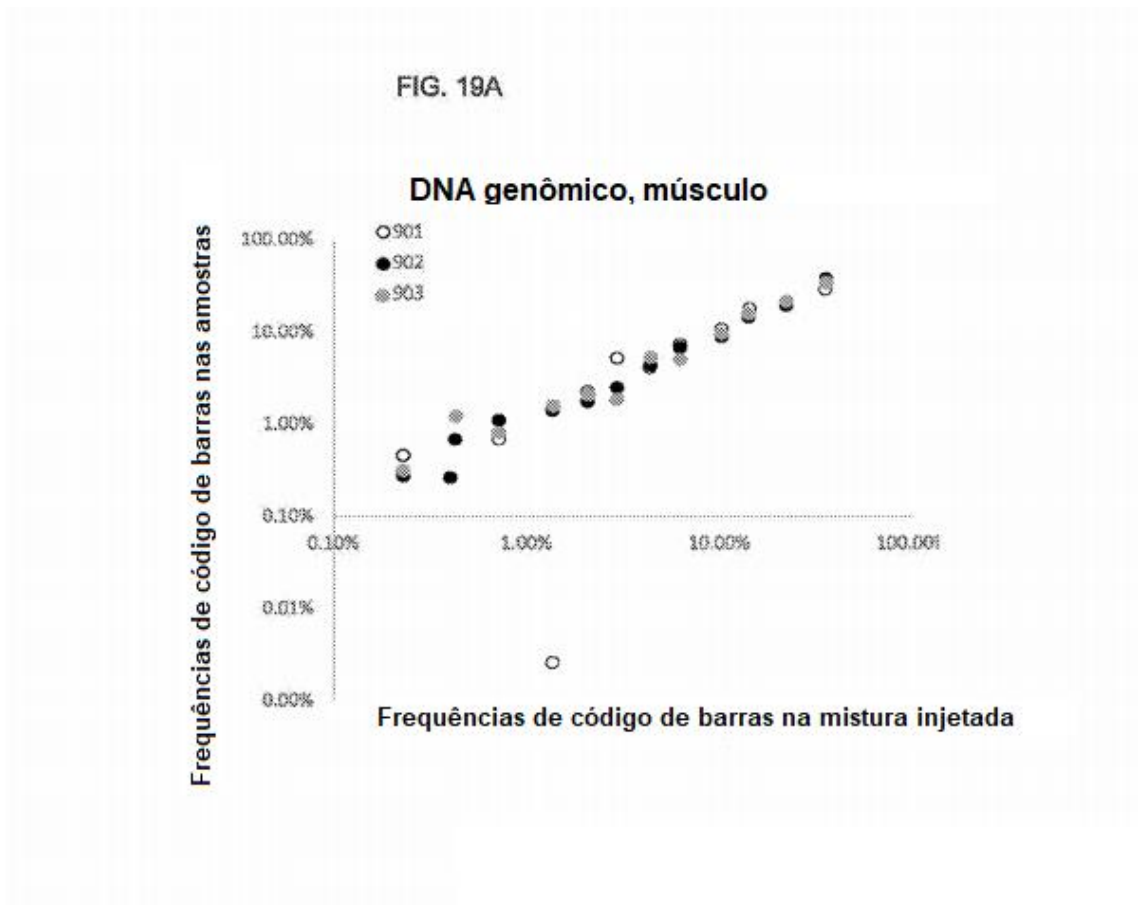


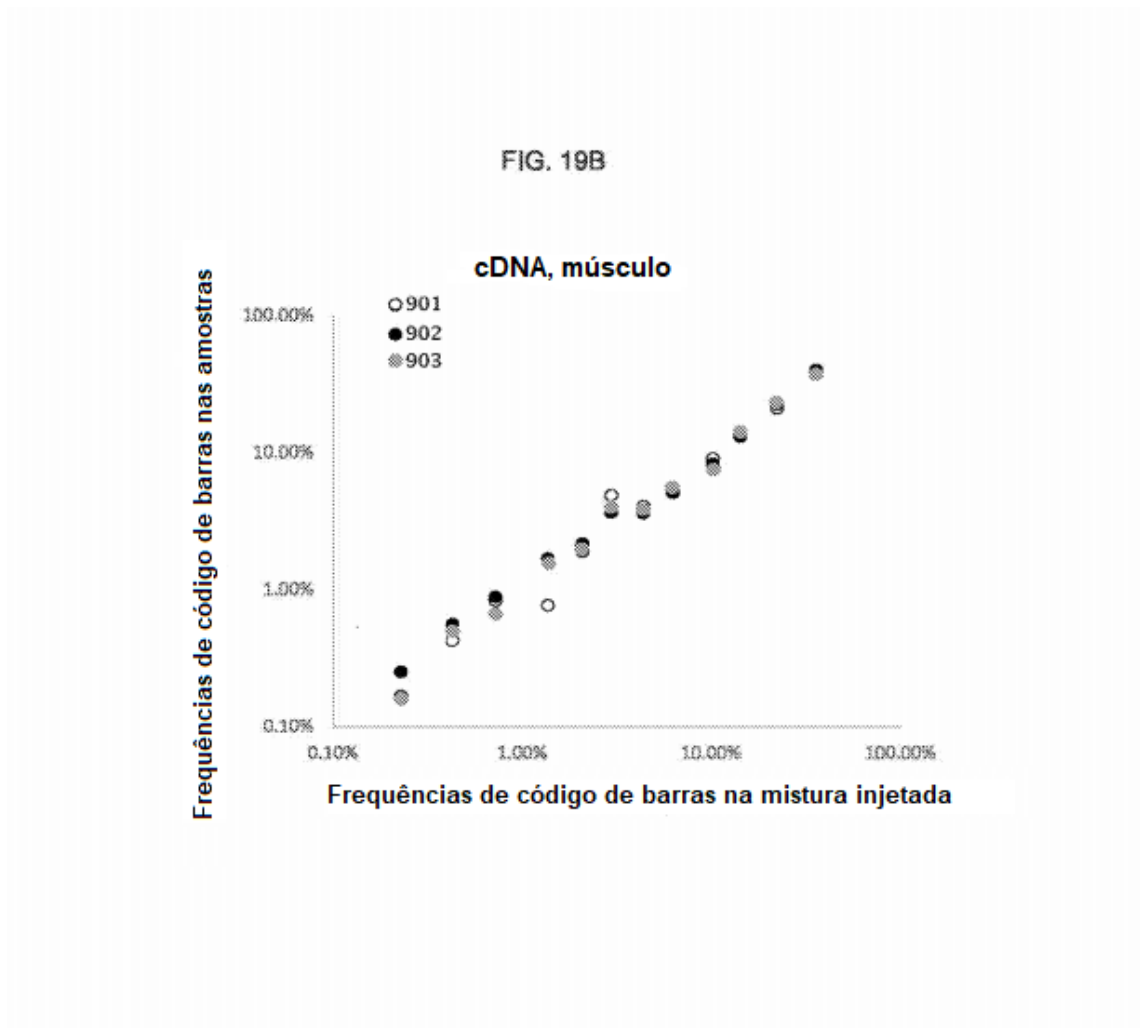


**FIG 18D**



**FIG. 18D**





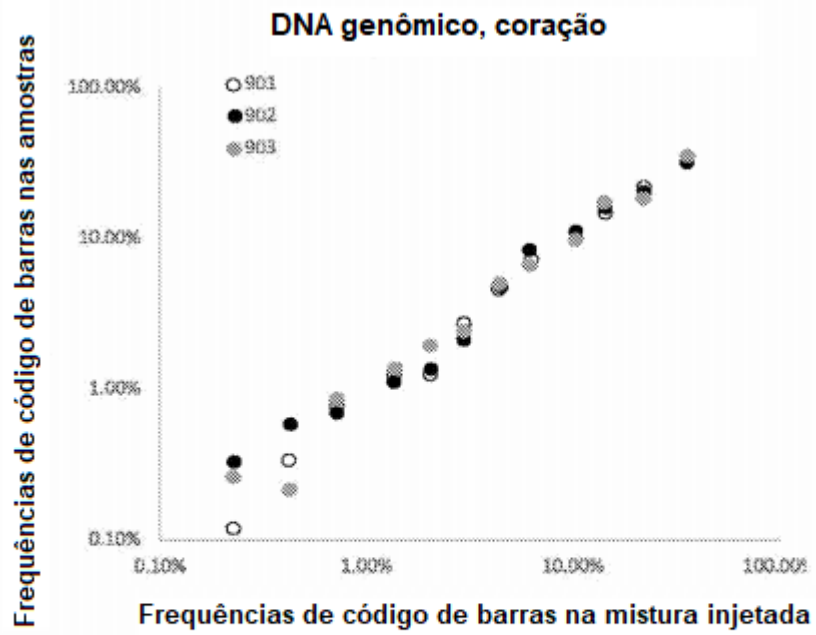


FIG. 19C

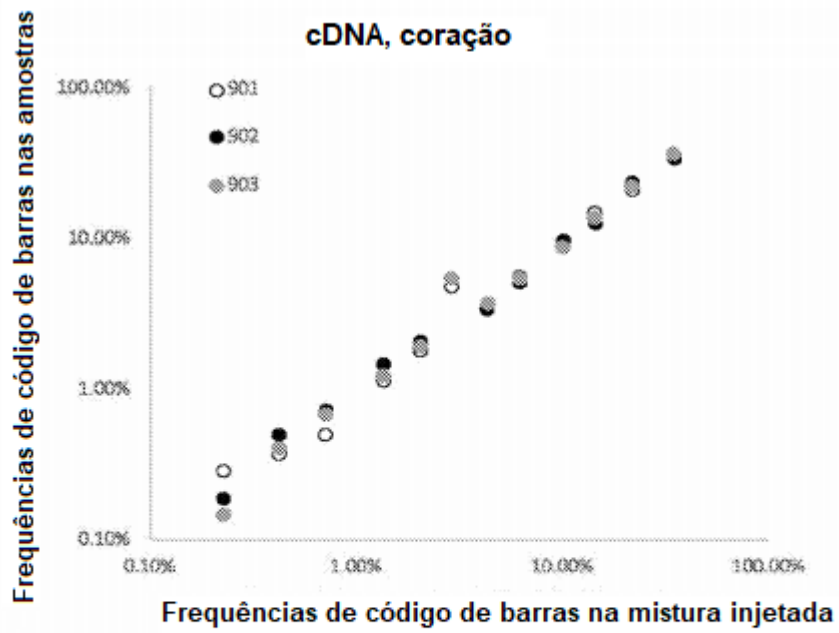
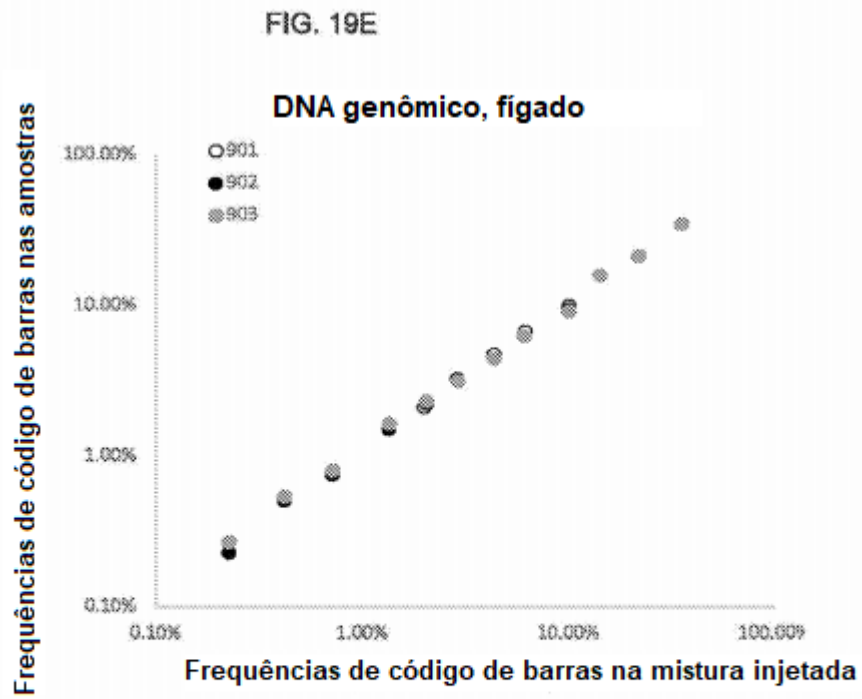
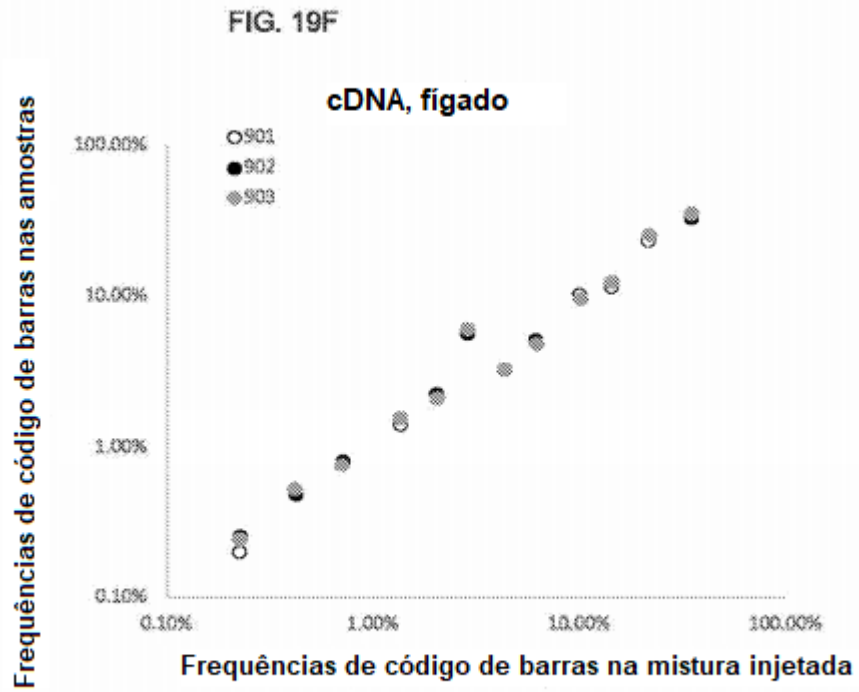


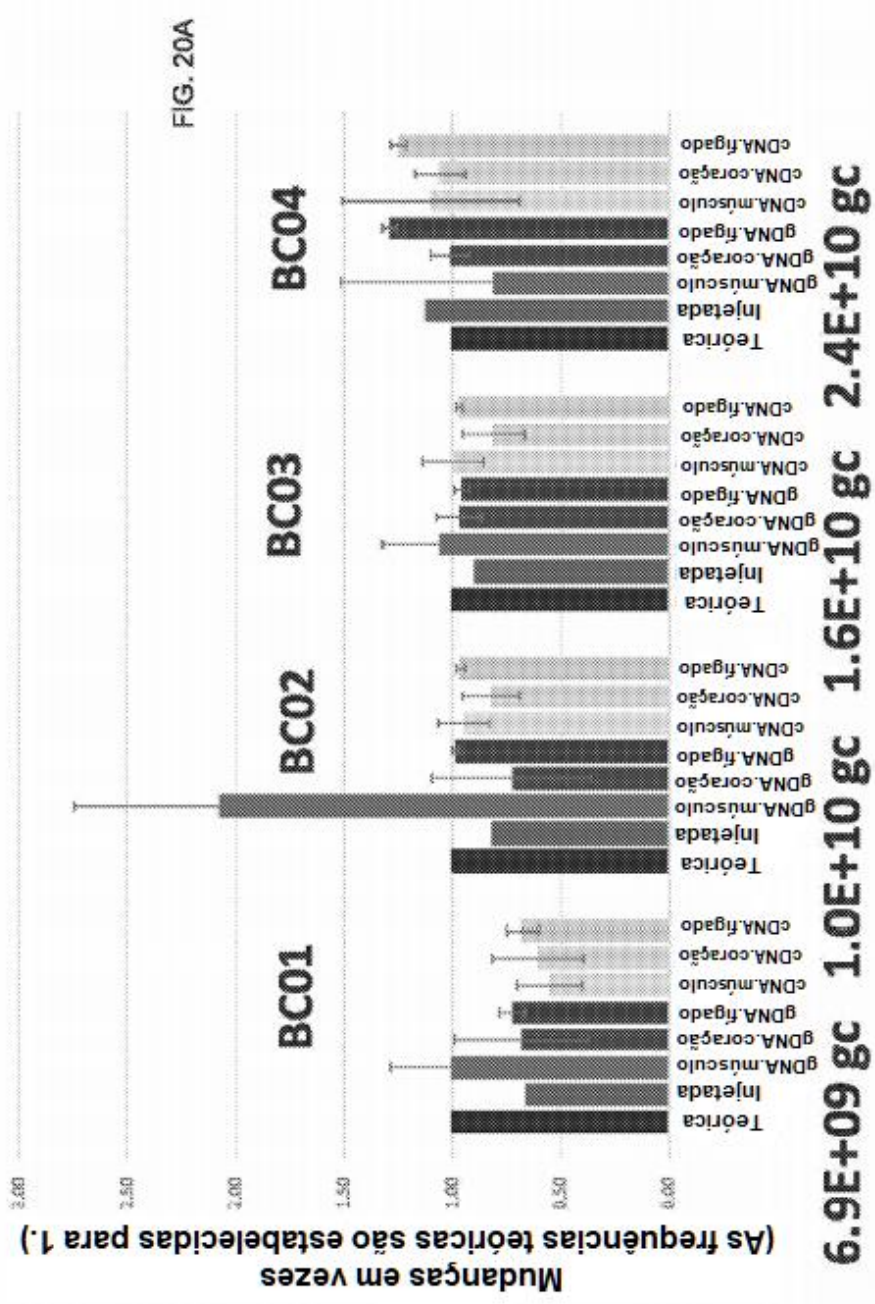
FIG. 19D





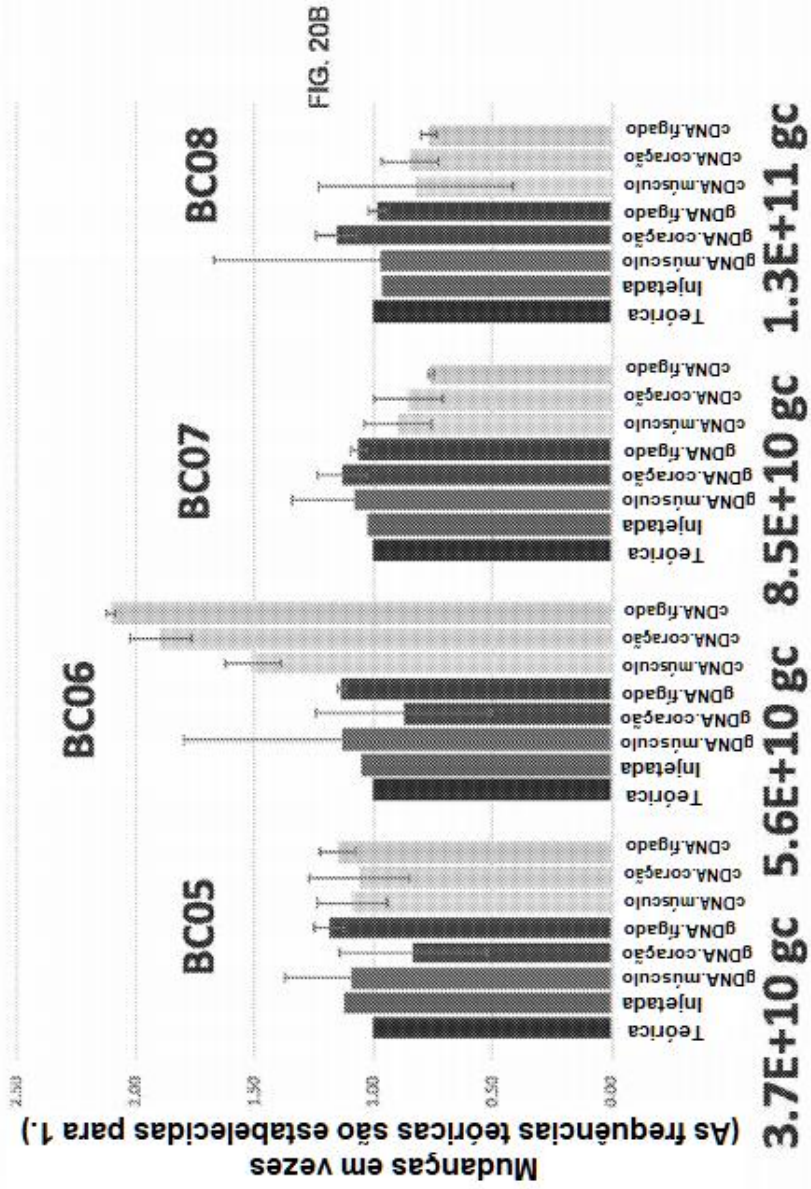


Mudança em vezes (em comparação com as frequências teóricas)



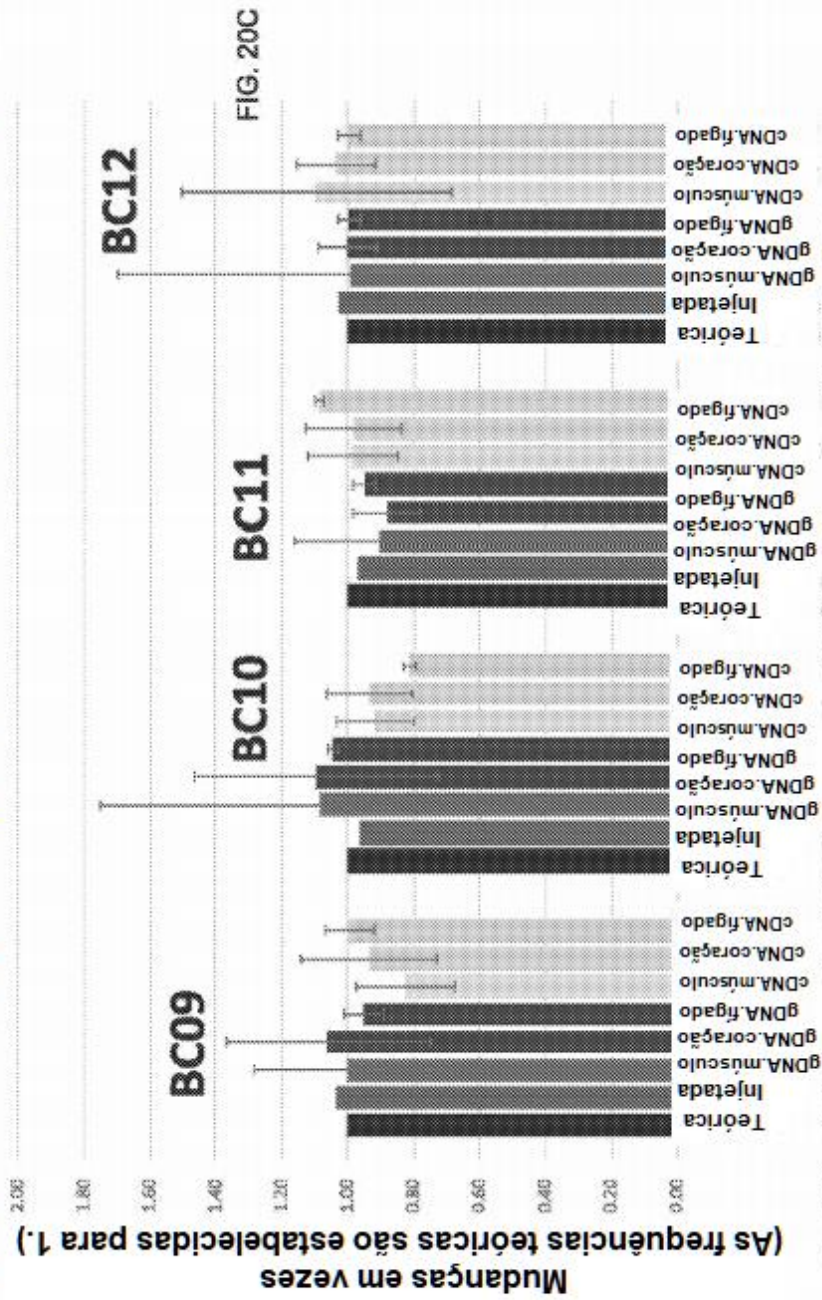
Os códigos de barra e suas injeções de frequências genômicas.

**Mudança em vezes (em comparação com as frequências teóricas)**



**Os códigos de barra e suas injeções de frequências genômicas.**

**Mudança em vezes (em comparação com as frequências teóricas)**



**2.0E+11 gc 3.0E+11 gc 4.5E+11 gc 6.9E+11 gc**

**Os códigos de barra e suas injeções de frequências genômicas.**

**Biodistribuição em camundongos**

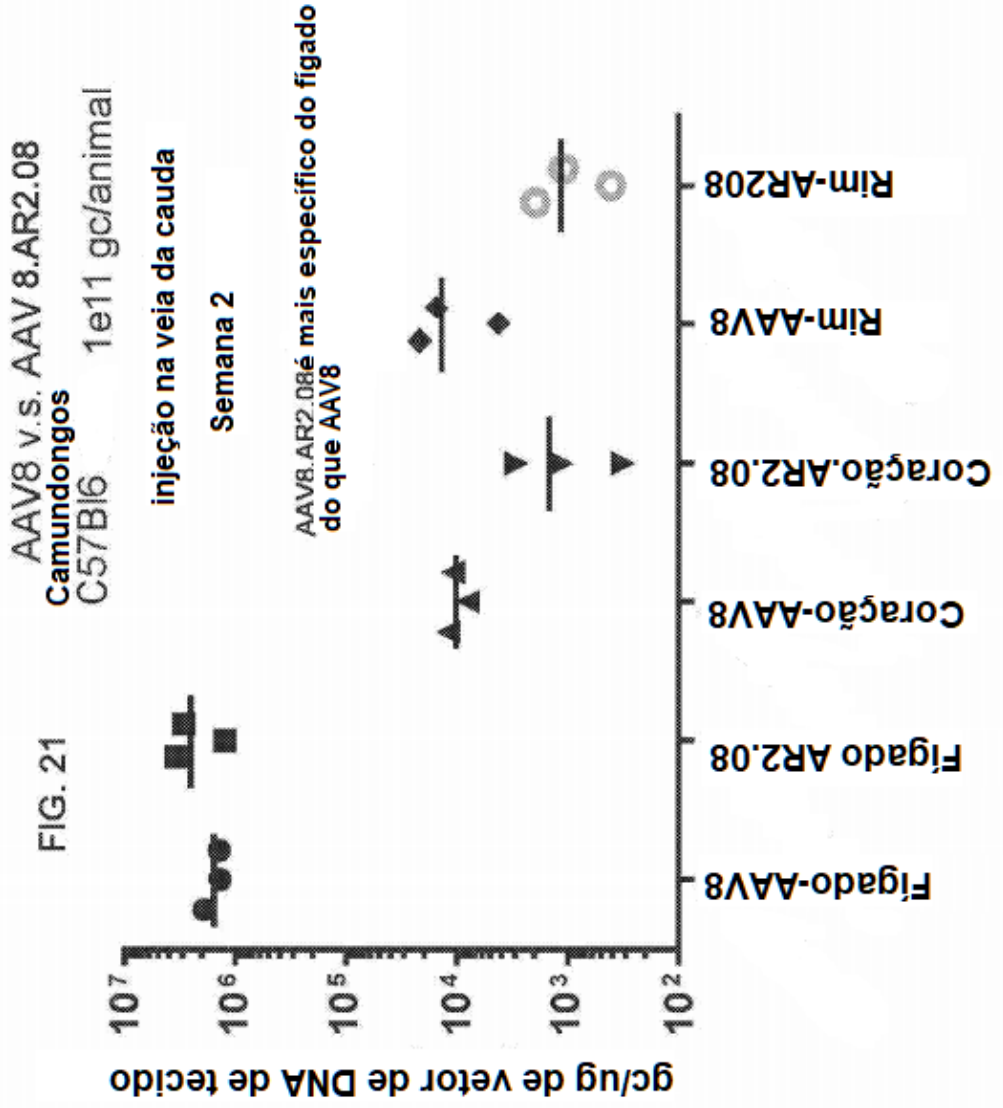
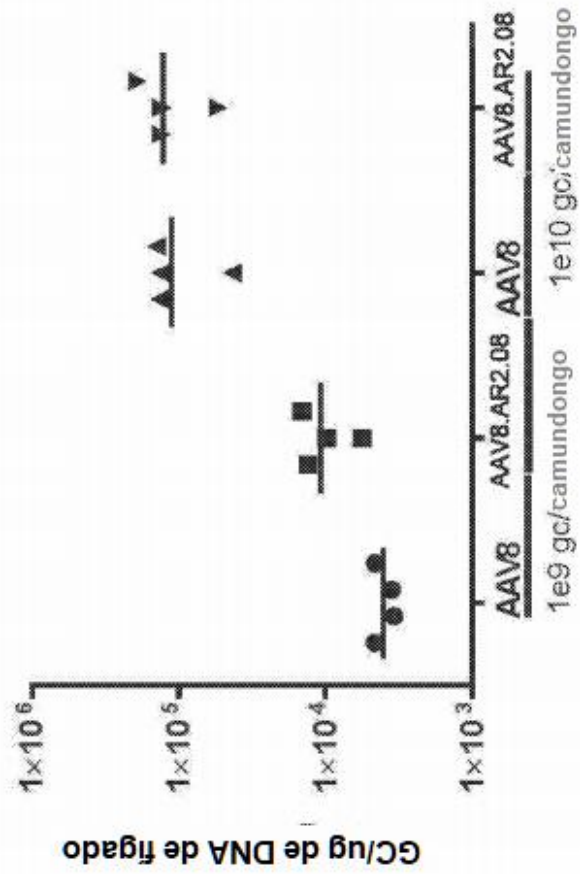


FIG. 22

- **AAV8 VS AAV8.AR2.08**

- CB7.CI.eGFP.W  
PRE.rBG
- ddPCR(PK),  
1e9 e 1e10  
gc/animal
- B6, macho, i.V.,  
semana 2, fígado



Tratamentos

FIG. 23

Lote #	Nome de vetor	Título de genoma ddPCR-GC/ml	ddPCR-Rendimento total	Pureza
WL184C S	AAVG3.TBG.PI.eGFP.WPRE.bGH(p0080;p0146)	1.141e14 GC/ml	6.104e14 GC	100.000
WL1549C S	AAV8.TBG.PI.eGFP.WPRE.bGH	1.11e14	5.578e14	98.9
WL1956C S	AAVrh79.TBG.PI.eGFP.WPRE.bGH(p3889;p0146)	1.609e14	8.568e14	

**FIG. 24**  
**Biodistribuição de vetor de AAV8.AR2.08 (17-07)**

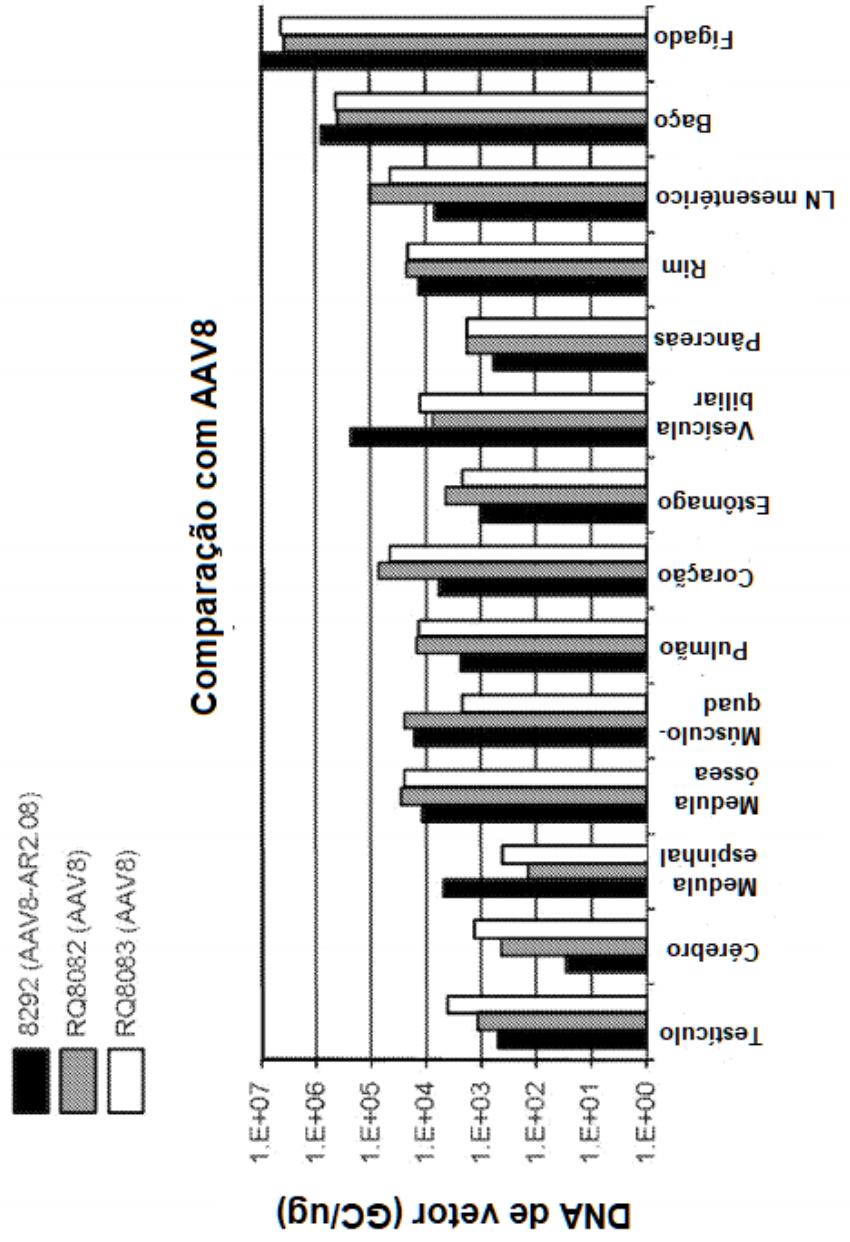


FIG. 25

**17-07: Transdução de fígado por novos sorotipos**AAV.TBG.EGFP.WPRE ( $1 \times 10^{13}$  ddGC/kg), dia 7, necrópsia

Sorotipo	ID	Espécies	Sexo	Peso (Kg)
AAV8	RA1475	rhesus	F	4.60
AAVrh79	RA2362	cyno	F	4.10
AAV8.AR2.08	8292	cyno	M	8.55
AAV8.AR2.08	RQ9471	rhesus	F	9.26



## 17-07: Transdução de fígado por novos sorotipos

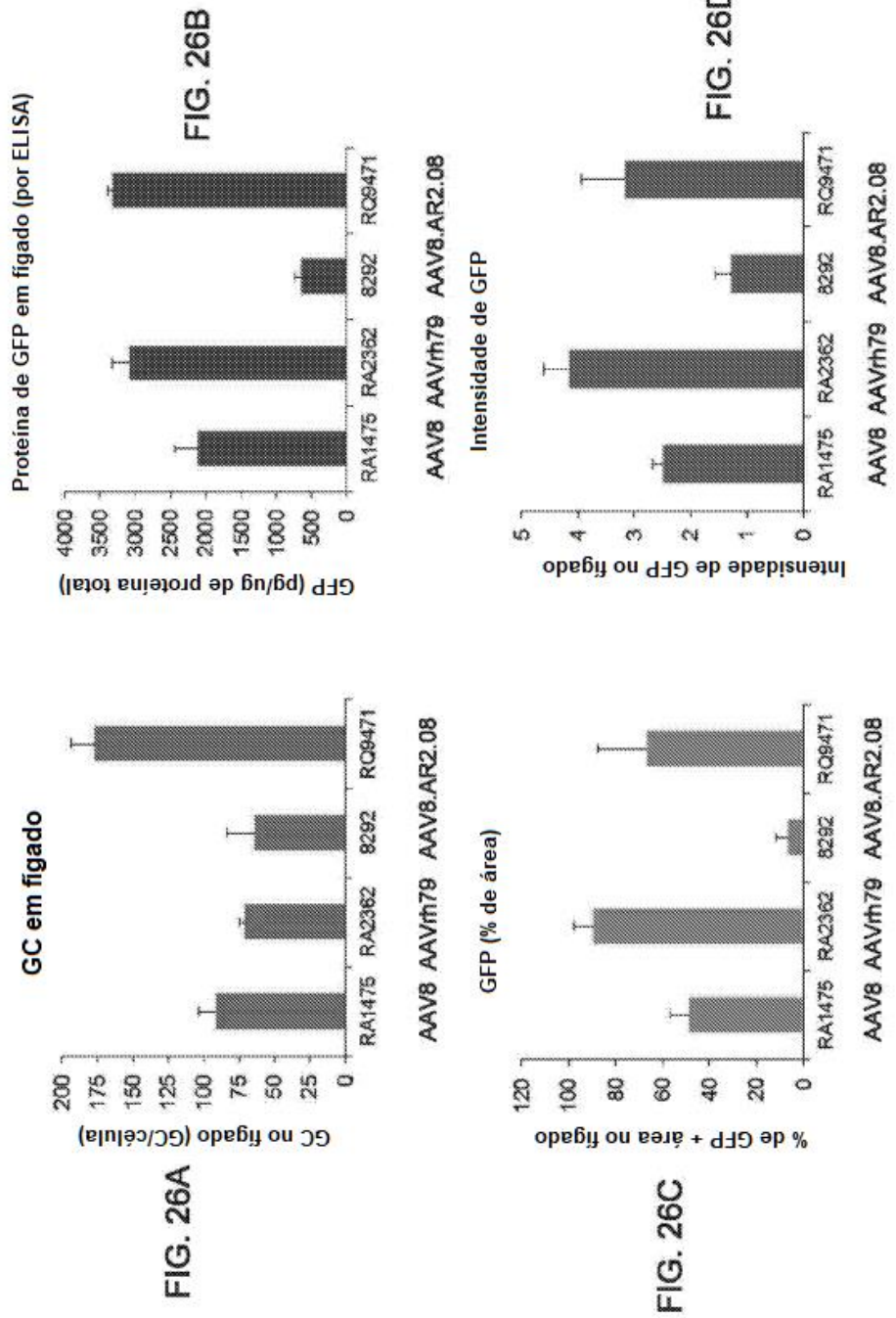
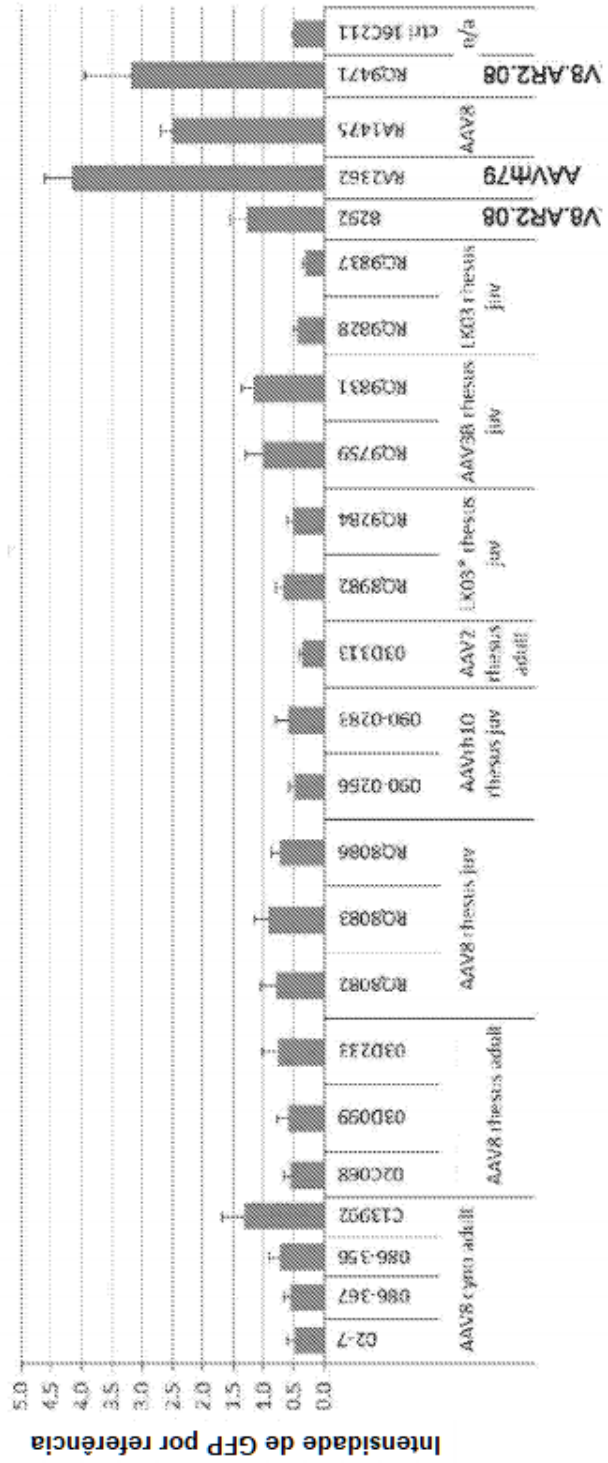


FIG. 27A

# Expressão de GFP em fígado NHP

Intensidade de GFP



# Expressão de GFP em fígado NHP

GFP+ área (fígado)

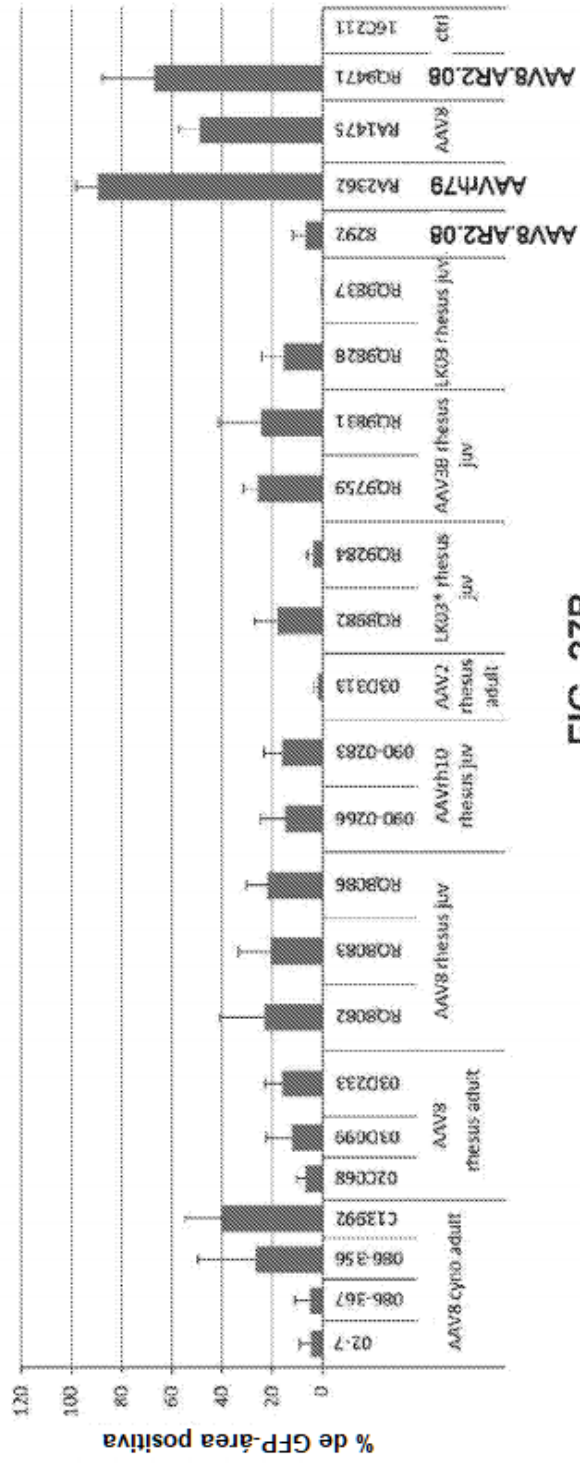
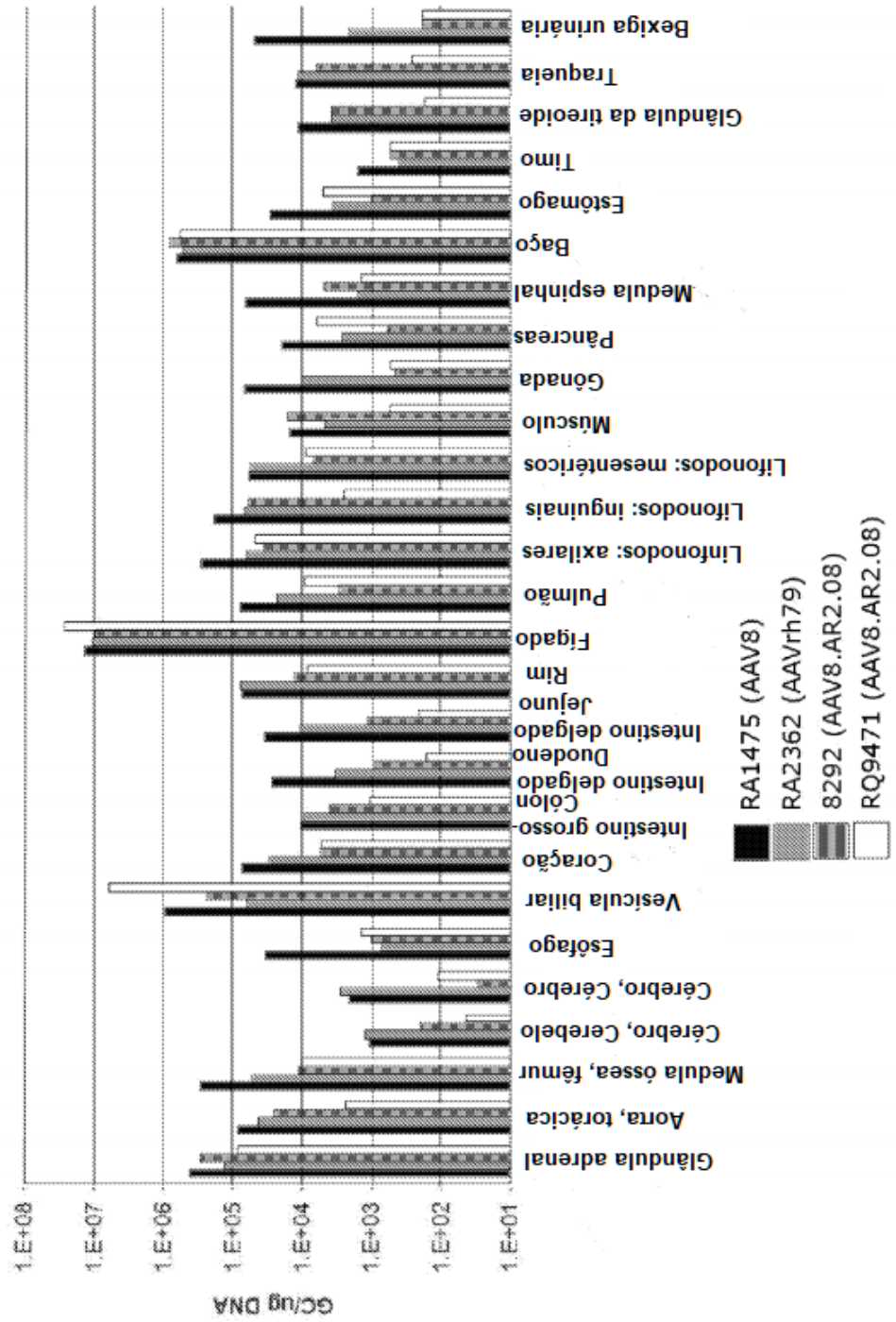
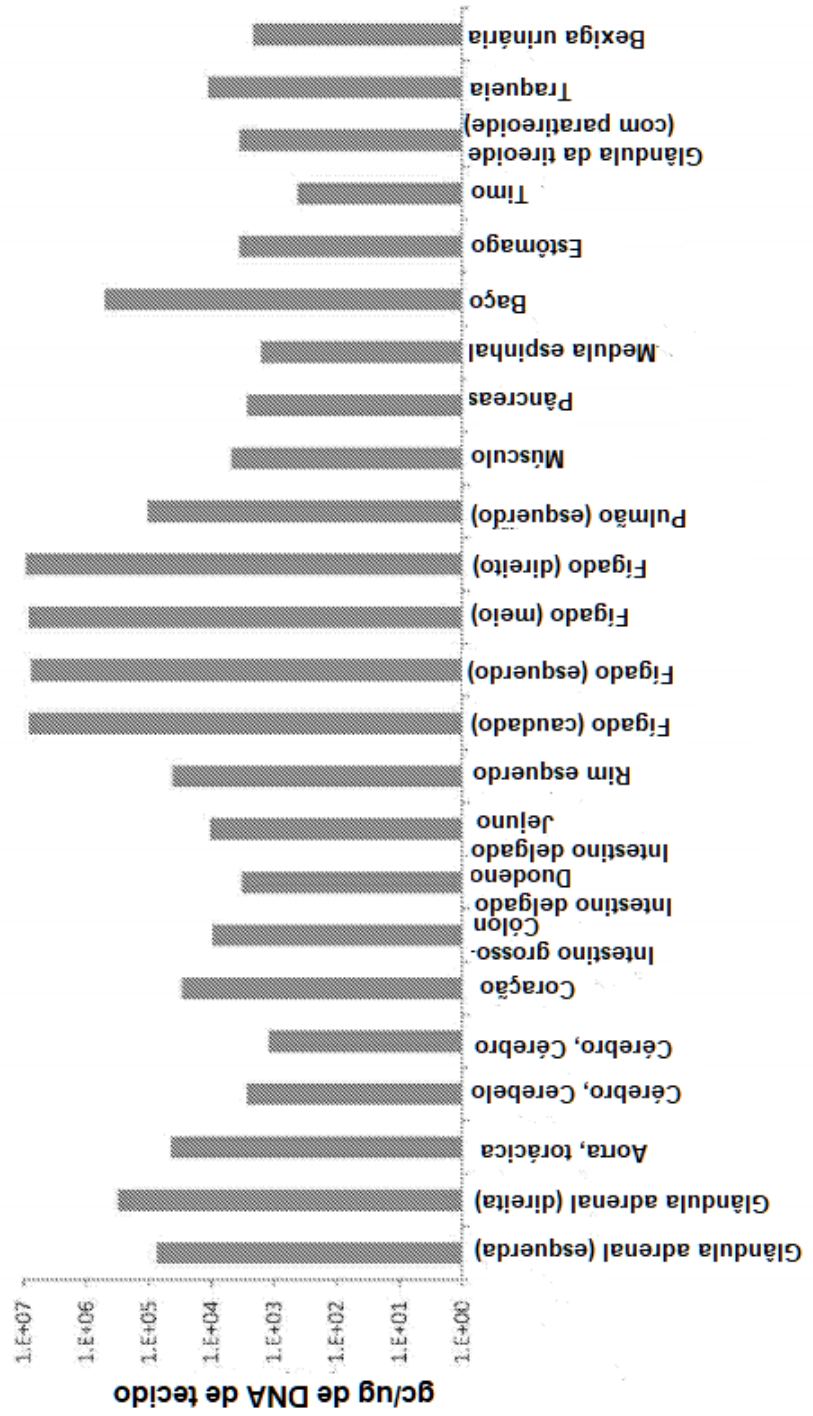


FIG. 27B

**FIG. 28**  
**17-07: Biodistribuição de vetor**



**FIG. 29**  
**Biodistribuição de**  
**17-07c RA2362**



**RESUMO**

Patente de Invenção: **“VETORES DE VÍRUS ADENOASSOCIADO (AAV), VETORES DE AAV TENDO DESAMIDAÇÃO DE CAPSÍDEOS REDUZIDA E USOS PARA OS MESMOS”**.

A presente invenção refere-se a um vetor de vírus adenoassociado recombinante (rAAV) compreendendo um capsídeo de AAV tendo uma população heterogênea de proteínas vp1, uma população heterogênea de proteína vp2 e uma população heterogênea de proteínas vp3. O capsídeo contém aminoácidos modificados em comparação com a sequência de aminoácidos VP1 codificada, o capsídeo contendo resíduos de asparagina altamente desamidados no par asparagina-glicina, e compreendendo ainda vários outros resíduos de asparagina menos desamidada e opcionalmente de glutamina.

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

### Código de Controle

Campo 1



Campo 2



### Outras Informações:

- Nome do Arquivo: LISTAGEM DE SEQUÊNCIA DE TRADUÇÃO 30 OU 60
- Data de Geração do Código: 29/09/2020
- Hora de Geração do Código: 17:06:19
- Código de Controle:
  - Campo 1: 796FD5891B6EB3FC
  - Campo 2: 60451D5A9ECE7833