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

(21) BR 112021000145-0 A2

(22) Data do Depósito: 09/07/2019

(43) Data da Publicação Nacional: 06/04/2021

(54) Título: COMPOSIÇÕES DE FUSOSSOMA E USOS DAS MESMAS

(51) Int. Cl.: A61K 39/12; C12N 7/00; C07K 14/005; C12N 15/88.

(30) Prioridade Unionista: 09/07/2018 US 62/695,537; 09/07/2018 US 62/695,650; 14/11/2018 US 62/767,241; 14/11/2018 US 62/767,261; 15/05/2019 US 62/848,284; (...).

(71) Depositante(es): FLAGSHIP PIONEERING INNOVATIONS V, INC.

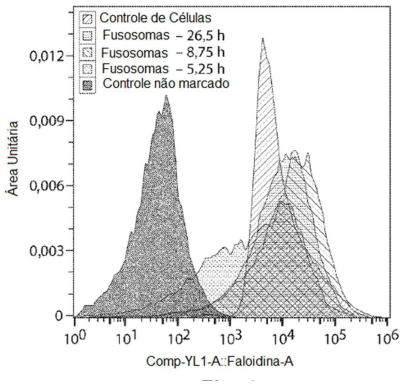
(72) Inventor(es): GEOFFREY A. VON MALTZAHN; JACOB ROSENBLUM RUBENS; JOHN MILES MILWID; MICHAEL TRAVIS MEE; NEAL FRANCIS GORDON; JAGESH VIJAYKUMAR SHAH.

(86) Pedido PCT: PCT US2019040978 de 09/07/2019

(87) Publicação PCT: WO 2020/014209 de 16/01/2020

(85) Data da Fase Nacional: 06/01/2021

(57) Resumo: A presente invenção fornece, pelo menos em parte, métodos e composições para entrega in vivo de fusossoma. Em algumas modalidades, o fusossoma compreende uma combinação de elementos que promovem a especificidade para células-alvo, por exemplo, um ou mais dentre um fusogênio, um elemento regulatório específico de célula-alvo positivo e um elemento regulatório específico de célula não alvo. Em algumas modalidades, o fusossoma compreende uma ou mais modificações que diminuem uma resposta imunológica contra o fusossoma.



Relatório Descritivo da Patente de Invenção para **"COMPOSIÇÕES DE FUSOSSOMA E USOS DAS MESMAS".**

REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

[0001] O presente pedido reivindica prioridade sobre os pedidos provisórios n° U.S. 62/695.537, depositado em 9 de julho de 2018, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF”, n° U.S. 62/767.241, depositado em 14 de novembro de 2018, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF”, n° U.S. 62/848.284, depositado em 15 de maio de 2019, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF”, n° U.S. 62/695.650, depositado em 9 de julho de 2018, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF”, n° U.S.62/767.261, depositado em 14 de novembro 2018, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF” e n° U.S. 62/848.305, depositado em 15 de meio de 2019, intitulado “FUSOSSOME COMPOSITIONS AND USES THEREOF”, cujos conteúdos são incorporados a título de referência em suas totalidades.

INCORPORAÇÃO A TÍTULO DE REFERÊNCIA DA LISTAGEM DE SEQUÊNCIAS

[0002] O presente pedido está sendo depositado junto da Listagem de Sequências em formato eletrônico. A Listagem de Sequências é fornecida como um arquivo denominado V2050-7024WO_SeqList.TXT, criado em 9 de julho de 2019, que tem 2.549.164 bytes de tamanho. As informações no formato eletrônico da Listagem de Sequências são incorporadas a título de referência em sua totalidade.

ANTECEDENTES

[0003] Produtos biológicos complexos são candidatos terapêuticos promissores para uma variedade de doenças. No entanto, é difícil entregar grandes agentes biológicos em uma célula devido ao fato de que a membrana plasmática atua como uma barreira entre a célula e o espaço extracelular. Há uma necessidade na técnica de novos métodos para entregar produtos biológicos complexos em células de um indivíduo.

SUMÁRIO

[0004] A presente divulgação fornece, pelo menos em parte, métodos e composições de fusossoma para entrega *in vivo*. Em algumas modalidades, o fusossoma compreende uma combinação de elementos que promovem a especificidade para células-alvo, por exemplo, um ou mais dentre um fusogênio, um elemento regulatório específico de célula-alvo positivo, e um elemento regulatório específico de célula não alvo. Em algumas modalidades, o fusossoma compreende uma ou mais modificações que diminuem uma resposta imunológica contra o fusossoma.

MODALIDADES ENUMERADAS

[0005] No presente documento são fornecidos fusossomas, incluindo vetores ou partículas retrovirais, tais como vetores ou partículas lentivirais, que geralmente resultam no aumento da expressão de um agente exógeno desejado (por exemplo, um transgene terapêutico) em células-alvo em comparação a células não alvo após a introdução dos fusossomas em células, por exemplo, em um indivíduo. Por exemplo, em alguns casos, o aumento na expressão segue a administração *in vivo* de um fusossoma fornecido (por exemplo, vetores ou partículas retrovirais) a um indivíduo, por exemplo, indivíduo humano. Em particular, um dos principais desafios para a terapia gênica bem-sucedida é a capacidade de manter a expressão estável e a longo prazo de um transgene terapêutico (por exemplo, um agente exógeno) a partir de células geneticamente modificadas *in vivo*. A expressão de transgene em células não alvo, tais como células apresentadoras de antígeno (APCs), pode, em alguns aspectos, resultar na ativação da resposta imunológica adaptativa, o que causa a geração de anticorpos neutralizantes contra o produto do transgene pelas células B e/ou eliminação de células produtoras de transgene pelas células T. Dito isso, a limitação da expressão do transgene para células-alvo pode, em algumas

modalidades, impactar substancialmente a durabilidade da expressão do transgene, evitando a depuração imune. Além disso, a expressão do transgene específico do tipo de célula pode ser muito relevante para a biologia da doença, como a limitação da expressão de genes pró-apoptóticos para células-alvo hepáticas.

[0006] Em particular, são fornecidos no presente documento fusossomas (por exemplo, vetor ou partículas retrovirais) que incluem a expressão de sequências de ácidos nucleicos sob o controle ou regulação de um elemento regulatório específico de célula do fígado positivo (por exemplo, promotor de célula hepática) e/ou um elemento regulatório específico de célula não hepática. Em algumas modalidades, o elemento regulatório específico de célula não hepática ocorre por meio de silenciamento de gene mediado por miRNA, tal como por sequências de ácidos nucleico do complemento às sequências de miRNA em uma célula não hepática. Em algumas modalidades, os fusossomas fornecidos (por exemplo, vetores retrovirais ou partículas) podem acionar especificamente a expressão do transgene (agente exógeno) em uma célula hepática ao mesmo tempo que retringe ou limita a expressão em células não alvo (não hepáticas).

[0007] Entre as modalidades fornecidas estão:

[0008] 1. Um fusossoma que compreende:

[0009] a) uma bicamada lipídica que compreende um fusogênio; e

[0010] b) um ácido nucleico que compreende:

[0011] (i) um gene de carga útil que codifica um agente exógeno, por exemplo, um gene de carga útil que codifica um agente exógeno da Tabela 5, opcionalmente em que o agente exógeno é apresentado em qualquer uma das SEQ ID NOS: 161 a 518 ou é um fragmento funcional ou variante funcional do mesmo composta pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com qualquer uma das SEQ ID NOS: 161 a 518; e

[0012] (ii) um elemento regulatório específico de célula do fígado positivo (por exemplo, um promotor específico da célula hepática) ligado de maneira operacional ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula hepática em relação ao fusossoma, de outro modo semelhante, que carece do elemento regulatório específico de célula hepática positiva.

[0013] 2. O fusossoma da modalidade 1, em que o ácido nucleico compreende adicionalmente um elemento regulatório específico de células não hepáticas (por exemplo, uma sequência de reconhecimento de miRNA específicos para células não hepáticas), ligado operacionalmente ao gene de carga útil, em que o elemento regulatório específico de célula não hepática diminui a expressão do gene de carga útil em uma célula não hepática em relação a um fusossoma de outra forma similar sem o elemento regulatório específico de célula não hepática.

[0014] 3. Um fusossoma que compreende:

[0015] a) uma bicamada lipídica que compreende um fusogênio; e

[0016] b) um ácido nucleico que compreende:

[0017] (i) um gene de carga útil que codifica um agente exógeno, por exemplo, um agente exógeno da Tabela 5, em que, opcionalmente, o agente exógeno é apresentado em qualquer uma das SEQ ID NOS: 161 a 518 ou é um fragmento funcional ou variante funcional do mesmo que compreende pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com qualquer uma das SEQ ID NOS: 161 a 518; e

[0018] (ii) um promotor ligado operacionalmente ao gene de carga útil, em que o promotor é escolhido a partir de um promotor Apoa2, Cyp3a4, LP1B, MIR122, hemopexina, SERPINA1 ou HLP, por exemplo, de acordo com uma sequência da Tabela 3 ou uma sequência que

tem pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com o mesmo, em que, opcionalmente, o promotor compreende a sequência apresentada em qualquer uma das SEQ ID NOS: 133 a 136 ou 519 a 525 ou uma sequência com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com qualquer um de SEQ ID NOS: 133 a 136 ou 519 a 525.

[0019] 4. Um fusossoma que compreende:

[0020] a) uma bicamada lipídica que compreende um fusogênio; e

[0021] b) um ácido nucleico que compreende:

[0022] (i) um gene de carga útil que codifica um agente exógeno, por exemplo, um gene de carga útil que codifica um agente exógeno da Tabela 5, em que, opcionalmente, o agente exógeno é apresentado em qualquer uma das SEQ ID NOS: 161 a 518 ou é um fragmento funcional ou variante funcional do mesmo que compreende pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com qualquer uma dentre as SEQ ID NOS: 161 a 518; e

[0023] (ii) um elemento regulatório específico de célula não alvo (NTCSRE) (por exemplo, uma sequência de reconhecimento de miRNA específico de célula não alvo), ligado operacionalmente ao gene de carga útil, em que o NTCSRE diminui a expressão do gene de carga útil em uma célula ou tecido não alvo em relação a um fusossoma semelhante sem o NTCSRE.

[0024] 5. Um fusossoma que compreende:

[0025] a) uma bicamada lipídica que compreende um fusogênio; e

[0026] b) um ácido nucleico que compreende:

[0027] (i) um gene de carga útil que codifica um agente exógeno, por exemplo, um gene de carga útil que codifica um agente exógeno da Tabela 5, em que, opcionalmente, o agente exógeno é apresentado em qualquer uma das SEQ ID NOS: 161 a 518 ou é um fragmento funcional ou variante funcional do mesmo que compreende pelo me-

nos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com qualquer uma dentre as SEQ ID NOS: 161 a 518; e

[0028] (ii) um elemento regulatório específico de célula-alvo nevativo (TCSRE negativo) (por exemplo, uma sequência de reconhecimento de miRNA tecido-específico), ligado de maneira operacional ao gene de carga útil, em que o TCSRE negativo diminui a expressão do agente exógeno em uma célula não alvo ou tecido em relação a um ácido nucleico, de outro modo semelhante, que carece do TCSRE negativo.

[0029] 6. O fusossoma de qualquer uma das modalidades 4 ou 5, em que o ácido nucleico compreende adicionalmente um elemento regulatório específico de célula do fígado positivo (por exemplo, um promotor específico da célula hepática) ligado de maneira operacional ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula hepática em relação ao fusossoma, de outro modo semelhante, que carece do elemento regulatório específico de célula hepática positiva.

[0030] 7. Um fusossoma que compreende:

[0031] a) uma bicamada lipídica que compreende um fusogênio;

[0032] b) um ácido nucleico que compreende um gene de carga útil que codifica um agente exógeno, por exemplo, um gene de carga útil que codifica um agente exógeno da Tabela 5, em que, opcionalmente, o agente exógeno é apresentado em qualquer uma das SEQ ID NOS: 161 a 518 ou é um fragmento funcional ou variante funcional do mesmo que compreende pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com qualquer uma dentre as SEQ ID NOS: 161 a 518; e

[0033] c) um ou dois dentre:

[0034] (i) uma primeira proteína imunossupressora exógena ou

superexpressa na bicamada lipídica; ou

[0035] ii) uma primeira proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzida em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação a um fusossoma gerado a partir de uma célula de origem não modificada de outra forma semelhante.

[0036] 8. O fusossoma de qualquer uma das modalidades anteriores, em que um ou mais dos seguintes:

[0037] i) o fusossoma é fundido com uma célula-alvo a uma taxa mais alta que com uma célula não alvo, por exemplo, em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

[0038] ii) o fusossoma é fundido a uma célula-alvo a uma taxa mais alta que com outro fusossoma, por exemplo, em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

[0039] iii) o fusossoma é fundido a células-alvo a uma taxa em que um agente no fusossoma é entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% de células-alvo após 24, 48 ou 72 horas;

[0040] iv) o fusossoma entrega o ácido nucleico a uma célula-alvo a uma taxa mais alta que a uma célula não alvo, por exemplo, em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

[0041] v) o fusossoma entrega o ácido nucleico, por exemplo, um ácido nucleico retroviral, a uma célula-alvo a uma taxa mais alta que outro fusossoma, por exemplo, em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10

vezes, 20 vezes, 50 vezes ou 100 vezes; ou

[0042] vi) o fusossoma entrega o ácido nucleico, por exemplo, o ácido nucleico retroviral, a uma célula-alvo a uma taxa de modo que um agente no fusossoma seja entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% das células-alvo após 24, 48 ou 72 horas.

[0043] 9. O fusossoma de qualquer uma das modalidades anteriores, em que um ou mais dentre (por exemplo, 2 ou todos os 3) o seguinte se aplica: o fusossoma é um vetor retroviral, a bicamada lipídica é composta por um envelope, por exemplo, um envelope viral e o ácido nucleico é um ácido nucleico retroviral.

[0044] 10. O fusossoma de qualquer uma das modalidades anteriores, em que o ácido nucleico compreende uma ou mais dentre (por exemplo, todos) as seguintes sequências de ácido nucleicos: LTR 5' (por exemplo, que compreende a U5 e que carece de um domínio da U3 funcional), elemento de empacotamento Psi (Psi), Promotor de trato central de polipurina (cPPT) ligado de maneira operacional ao gene de carga útil, gene de carga útil (que compreende, opcionalmente, um íntron antes do quadro aberto de leitura), sequência de cauda Poli-A, WPRE e LTR 3' (por exemplo, que compreende a U5 e que carece de uma U3 funcional).

[0045] 11. O fusossoma de qualquer uma das modalidades anteriores que compreende uma ou mais dentre (por exemplo, todas) uma polimerase (por exemplo, uma transcriptase reversa, por exemplo, pol ou uma porção da mesma), uma integrase (por exemplo, pol ou uma porção da mesma, por exemplo, uma variante funcional ou não funcional), uma proteína de matriz (por exemplo, gag ou uma porção da mesma), uma proteína capsidial (por exemplo, gag ou uma porção da mesma), proteína nucleocapsidial (por exemplo, gag ou uma porção da mesma) e uma protease (por exemplo, pro).

[0046] 12. O fusossoma da modalidade 7 que compreende (i) e (ii).

[0047] 13. O fusossoma de qualquer uma das modalidades 7 a 12 que compreende adicionalmente uma segunda proteína imunossupressora exógena ou superexpressa na bicamada lipídica.

[0048] 14. O fusossoma de acordo com qualquer uma das modalidades 7 a 13 que compreende adicionalmente uma segunda proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzido em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação a um fusossoma gerado a partir de uma célula de origem não modificada, de outro modo semelhante.

[0049] 15. O fusossoma de qualquer uma das modalidades 7 a 14, em que o ácido nucleico, por exemplo, vetor retroviral, compreende adicionalmente um elemento regulatório específico de célula do fígado positivo (por exemplo, um promotor específico da célula hepática) ligado operacionalmente ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula hepática em relação a um fusossoma semelhante que carece do elemento regulatório específico de célula do fígado positivo.

[0050] 16. O fusossoma de qualquer uma das modalidades 7 a 15, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, compreende adicionalmente um elemento regulatório específico de célula não alvo (NTCSRE) (por exemplo, uma sequência de reconhecimento de miRNA específico de célula não alvo), ligado operacionalmente ao gene de carga útil, em que o NTCSRE diminui a expressão do gene de carga útil em uma célula ou tecido não alvo em relação a um fusossoma semelhante que carece o NTCSRE.

[0051] 17. O fusossoma de qualquer uma das modalidades 7 a 15, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, com-

preende adicionalmente um elemento regulatório específico de célula alvo negativo (TCSRE negativo) (por exemplo, uma sequência de reconhecimento de miRNA específico de tecido), ligado operacionalmente ao gene de carga útil, em que o TCSRE negativo diminui a expressão do agente exógeno em uma célula ou tecido não alvo em relação a um ácido nucléico semelhante, por exemplo, ácido nucléico retroviral, que carece do TCSRE negativo.

[0052] 18. O fusossoma de qualquer uma das modalidades 7 a 17, em que, quando administrado a um indivíduo (por exemplo, um indivíduo humano ou um camundongo), ocorre um ou mais dentre:

[0053] i) o fusossoma não produz uma resposta de anticorpo detectável (por exemplo, após uma única administração ou uma pluralidade de administrações) ou anticorpos contra o fusossoma estão presentes em um nível inferior a 10%, 5%, 4%, 3%, 2%, ou 1% acima de um nível precedente, por exemplo, por um ensaio de detecção de anticorpo FACS, por exemplo, um ensaio do Exemplo 13 ou Exemplo 14;

[0054] ii) o fusossoma não produz uma resposta imunológica celular detectável (por exemplo, resposta de células T, resposta de células NK ou resposta de macrófago) ou uma resposta imunológica celular contra o fusossoma está presente em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima de um nível precedente, por exemplo, por um ensaio de lise de PBMC (por exemplo, um ensaio do Exemplo 5), por um ensaio de lise de células NK (por exemplo, um ensaio do Exemplo 6), por um ensaio de lise de células T exterminadoras CD8 (por exemplo, um ensaio do Exemplo 7), por um ensaio de fagocitose de macrófagos (por exemplo, um ensaio do Exemplo 8);

[0055] iii) o fusossoma não produz uma resposta imunológica inata detectável, por exemplo, ativação do complemento (por exemplo, após uma única administração ou uma pluralidade de administrações), ou a resposta imunológica inata contra o fusossoma está presente em um

nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima de um nível precedente, por exemplo, por um ensaio de atividade do complemento (por exemplo, um ensaio do Exemplo 9);

[0056] iv) menos de 10%, 5%, 4%, 3%, 2% ou 1% de fusossomas são inativados por soro, por exemplo, por um ensaio de inativação do soro, por exemplo, um ensaio do Exemplo 11 ou Exemplo 12;

[0057] v) uma célula-alvo que recebeu o agente exógeno do fusossoma não produz uma resposta de anticorpo detectável (por exemplo, após uma única administração ou uma pluralidade de administrações) ou anticorpos contra a célula-alvo estão presentes em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima de um nível precedente, por exemplo, por um ensaio de detecção de anticorpo FACS, por exemplo, um ensaio do Exemplo 15; ou

[0058] vi) uma célula-alvo que recebeu o agente exógeno do fusossoma não produz uma resposta imunológica celular detectável (por exemplo, resposta de célula T, resposta de célula NK ou resposta de macrófago), ou uma resposta celular contra a célula-alvo está presente em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima de um nível precedente, por exemplo, por um ensaio de fagocitose de macrófagos (por exemplo, um ensaio do Exemplo 16), por um ensaio de lise de PBMC (por exemplo, um ensaio do Exemplo 17), por um ensaio de lise de células NK (por exemplo, um ensaio do Exemplo 18) ou por um ensaio de lise de células T exterminadoras CD8 (por exemplo, um ensaio do Exemplo 19).

[0059] 19. O fusossoma da modalidade 18, em que o nível em segundo plano é o nível correspondente no mesmo indivíduo antes da administração do fusossoma.

[0060] 20. O fusossoma de acordo com qualquer uma das modalidades 7 a 19, em que a proteína imunossupressora (por exemplo, primeira proteína imunossupressora ou segunda proteína imunossupres-

sora) é uma proteína reguladora complementar ou CD47.

[0061] 21 O fusossoma de qualquer uma das modalidades 7 a 20, em que a proteína imunoestimuladora (por exemplo, primeira proteína imunoestimuladora ou segunda proteína imunoestimuladora) é um MHC I (por exemplo, HLA-A, HLA-B, HLA-C, HLA-E ou HLA -G) ou proteína MHC II (por exemplo, HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ ou HLA-DR).

[0062] 22. O fusossoma de qualquer uma das modalidades anteriores, em que o agente exógeno é escolhido a partir de: OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMA-CHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAH, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LMBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTD, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1 ou LDLR.

[0063] 23. O fusossoma de qualquer uma das modalidades anteriores, em que o fusogênio compreende a VSV-G.

[0064] 24. O fusossoma de quaisquer modalidades 1, 2, 6, 15, 22 ou 23, em que o elemento regulatório específico do fígado positivo compreende um promotor específico do fígado, um intensificador específico do fígado, um sítio de splice específico do fígado, um sítio específico do fígado que estende a meia-vida de um RNA ou proteína, um sítio de promoção de exportação nuclear de mRNA específico do fígado, um sítio de intensificação de tradução específico do fígado ou um sítio de modificação pós-tradução específico do fígado.

[0065] 25. O fusossoma de quaisquer modalidades 1, 2, 6, 15 ou 22 a 24, em que o elemento regulatório específico do fígado positivo compreende um promotor específico do hepatócito.

[0066] 26. O fusossoma da modalidade 25, em que o promotor específico dos hepatócitos compreende um motivo da Tabela 3, em que, opcionalmente, o promotor é apresentado em qualquer uma das SEQ ID NOS: 133 a 136 ou 519 a 525 ou uma sequência que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência em relação a qualquer uma das SEQ ID NOS: 133 a 136 ou 519 a 525

[0067] 27. O fusossoma da modalidade 25 ou 26, em que o elemento regulatório específico do fígado positivo compreende um promotor selecionado a partir de um promotor transtirretina (ET), hAAT, Alb, Apoa2, Cyp3a4, LP1B, MIR122, hemopexina, SERPINA1 ou HLP.

[0068] 28. O fusossoma de qualquer uma das modalidades 4 a 6 ou 16 a 21, em que o TCSRE negativo ou NTCSRE compreende uma sequência de reconhecimento de miRNA específico de célula não alvo, sítio de reconhecimento de protease específico de célula não alvo, sítio de ubiquitina ligase específico de célula não alvo, sítio de repressão transcricional específico de célula não alvo ou célula não alvo sítio específico de repressão epigenética.

[0069] 29. O fusossoma de qualquer uma das modalidades 4 a 6, 16 a 21 ou 28, em que o TCSRE negativo ou NTCSRE compreende uma sequência de reconhecimento de miRNA específica de tecido, sítio de reconhecimento de protease específico de tecido, sítio de ligase de ubiquitina específico de tecido, sítio de repressão transcricional específico de tecido, ou sítio de repressão transcricional epigenético específico de tecido.

[0070] 30. O fusossoma de qualquer uma das modalidades 4-6, 16-21, 28 ou 29, em que o TCSRE negativo ou NTCSRE compreende

uma sequência de reconhecimento de miRNA específico de célula não hepática, sítio de reconhecimento de protease específico de célula não hepática, sítio de ubiquitina ligase específico de célula não hepática, sítio de repressão transcricional específico de célula não hepática ou sítio de repressão epigenética específico de célula não hepática.

[0071] 31. O fusossoma de qualquer uma das modalidades 4 a 6, 16 a 21 ou 28 a 30, em que o TCSRE negativo ou NTCSRE compreende uma sequência de reconhecimento de miRNA específico de célula não hepática ligada por um miRNA da Tabela 4, por exemplo, por um ou mais de (por exemplo, dois ou mais de) miR-142, mir-181a-2, mir-181b- 1, mir-181c, mir-181a-1, mir-181b-2, mir-181d, miR-223 ou miR-126.

[0072] 32. O fusossoma de acordo com qualquer uma das modalidades 28 a 31, em que o TCSRE negativo ou NTCSRE é situado ou codificado dentro de uma região transcrita (por exemplo, a região transcrita que codifica o agente exógeno), por exemplo, de modo que um RNA produzido pela região transcrita compreende a sequências de reconhecimento de miRNA dentro de uma UTR ou região de codificação.

[0073] 33. O fusossoma de qualquer uma das modalidades anteriores, em que o ácido nucleico, por exemplo, ácido nucleico retroviral compreende um ou mais elementos de isolamento.

[0074] 34. O fusossoma da modalidade 33 em que o ácido nucleico, por exemplo, ácido nucleico retroviral, compreende dois elementos isolantes, por exemplo, um primeiro elemento isolante a montante do gene de carga útil e um segundo elemento isolante a jusante do gene de carga útil, por exemplo, em que o primeiro elemento isolante e o segundo elemento isolante compreendem as mesmas sequências ou sequências diferentes.

[0075] 35. O fusossoma de qualquer uma das modalidades ante-

cedentes, que não é genotóxico ou não aumenta a taxa de formação tumoral em células-alvo.

[0076] 36. O fusossoma de qualquer uma das modalidades anteriores, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, tem capacidade para se integrar no genoma de uma célula alvo.

[0077] 37. O fusossoma da modalidade 36, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, é um lentivírus competente para integração ou um lentivírus com deficiência de integração.

[0078] 38. O fusossoma de qualquer uma das modalidades anteriores, em que a célula alvo é escolhida a partir de um hepatócito, célula endotelial sinusoidal do fígado, colangiócito, célula estrelada, célula apresentadora de antígeno residente no fígado (por exemplo, Célula de Kupffer), linfócito imune residente no fígado (por exemplo, célula T, célula B ou célula NK) ou fibroblasto portal.

[0079] 39. O fusossoma de qualquer uma das modalidades 4 a 6 e 9 a 38, em que um ou mais dos seguintes:

[0080] i) menos de 10%, 5%, 4%, 3%, 2% ou 1% do agente exógeno presente, de modo detectável, no indivíduo está em células não alvo;

[0081] ii) pelo menos 90%, 95%, 96%, 97%, 98% ou 99% das células do indivíduo que compreendem, de modo detectável, o agente exógeno, são células-alvo (por exemplo, células de um tipo de célula única, por exemplo, células T);

[0082] iii) menos de 1.000.000, 500.000, 200.000, 100.000, 50.000, 20.000 ou 10.000 células dentre as células do indivíduo que compreendem, de modo detectável, o agente exógeno são células não alvo;

[0083] iv) os níveis médios do agente exógeno em todas as células-alvo no indivíduo são pelo menos 100 vezes, 200 vezes, 500 vezes ou 1.000 vezes maiores que os níveis médios do agente exógeno em

todas as células não alvo no indivíduo; ou

[0084] v) o agente exógeno não é detectável em qualquer célula não alvo no indivíduo.

[0085] 40. O fusossoma de qualquer uma das modalidades anteriores, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, codifica um TCSRE positivo e/ou um NTCSRE ou TCSRE negativo.

[0086] 41. O fusossoma de qualquer uma das modalidades anteriores, em que o ácido nucleico, por exemplo, ácido nucleico retroviral, compreende o complemento de um TCSRE positivo e/ou um NTCSRE ou TCSRE negativo.

[0087] 42. O fusossoma de qualquer modalidade 40 ou 41, em que o TCSRE positivo compreende um promotor específico do fígado que é pelo menos 10%, 25%, 50%, 75%, 100%, 150%, 200%, 250%, 300%, 400%, 500%, 750%, 1.000% ou mais ativo em uma célula hepática (por exemplo, hepatócito) do que uma célula não hepática.

[0088] 43. O fusossoma de qualquer uma das modalidades 40 a 42, em que o TCSRE ou NTCSRE negativo compreende uma sequência de reconhecimento de miRNA que diminui a expressão do gene em pelo menos 10%, 25%, 50%, 75% ou 100% em células hematopoiéticas em comparação com hepatócitos.

[0089] 44. O fusossoma de qualquer uma das modalidades anteriores, que não entrega ácido nucleico a uma célula não alvo, por exemplo, uma célula apresentadora de antígeno, uma célula MHC de classe II+, uma célula apresentadora de antígeno profissional, uma célula apresentadora de antígeno atípica, um macrófago, uma célula dendrítica, uma célula dendrítica mieloide, uma célula dendrítica plasmocitoide, uma célula CD11c+, uma célula CD11b+, um esplenócito, uma célula B, um hepatócito, uma célula endotelial ou uma célula não cancerosa.

[0090] 45. O fusossoma de qualquer uma das modalidades anteri-

ores, em que menos de 10%, 5%, 2,5%, 1%, 0,5%, 0,1%, 0,01%, 0,001%, 0,0001%, 0,00001% ou 0,000001% de um tipo de célula não alvo (por exemplo, uma ou mais dentre uma célula apresentadora de antígenos, uma célula MHC de classe II+, uma célula apresentadora de antígeno profissional, uma célula apresentadora de antígeno atípica, um macrófago, uma célula dendrítica, uma célula dendrítica mieloide, uma célula dendrítica plasmocitoide, uma célula CD11c+, uma célula CD11b+, um esplenócito, uma célula B, um hepatócito, uma célula endotelial ou uma célula não cancerosa) compreende o ácido nucleico, por exemplo, ácido nucleico retroviral, por exemplo, usando PCR quantitativa, por exemplo, com o uso de um ensaio do Exemplo 1.

[0091] 46. O fusossoma de qualquer uma das modalidades anteriores, em que as células-alvo compreendem 0,00001 a 10, 0,0001 a 10, 0,001 a 10, 0,01 a 10, 0,1 a 10, 0,5 a 5, 1 a 4, 1 a 3 ou 1 a 2 cópias do ácido nucleico, por exemplo, ácido nucleico retroviral ou uma porção do mesmo, por genoma de célula hospedeira, por exemplo, em que o número de cópias do ácido nucleico é avaliado após administração *in vivo*.

[0092] 47. O fusossoma de qualquer uma das modalidades anteriores, em que:

[0093] menos de 10%, 5%, 2,5%, 1%, 0,5%, 0,1%, 0,01% das células não alvo (por exemplo, uma célula apresentadora de antígeno, uma célula MHC de classe II+, uma célula apresentadora de antígeno profissional, uma célula apresentadora de antígeno atípica, um macrófago, uma célula dendrítica, uma célula dendrítica mieloide, uma célula dendrítica plasmocitoide, uma célula CD11c+, uma célula CD11b+, um esplenócito, uma célula B, um hepatócito, uma célula endotelial ou uma célula não cancerosa) compreendem o agente exógeno; ou

[0094] o agente exógeno (por exemplo, proteína) não está presente de modo detectável em uma célula não alvo, por exemplo, uma cé-

lula apresentadora de antígeno, uma célula MHC de classe II+, uma célula apresentadora de antígeno profissional, uma célula apresentadora de antígeno atípica, um macrófago, uma célula dendrítica, uma célula dendrítica mieloide, uma célula dendrítica plasmocitoide, uma célula CD11c+, uma célula CD11b+, um esplenócito, uma célula B, um hepatócito, uma célula endotelial ou uma célula não cancerosa.

[0095] 48. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma entrega o ácido nucleico, por exemplo, ácido nucleico retroviral, a uma célula-alvo, por exemplo, uma célula T, uma célula T CD3+, uma célula T CD4+, uma célula T CD8+, um hepatócito, uma célula-tronco hematepoiética, uma célula-tronco hematopoiética CD34+, uma célula-tronco hematepoiética CD105+, uma célula-tronco hematepoiética CD117+, uma célula endotelial CD105+, uma célula B, uma célula B CD20+, uma célula B CD19+, uma célula cancerosa, uma célula cancerosa CD133+, uma célula cancerosa EpCAM+, uma célula canceladora CD19+, uma célula de câncer Her2/Neu+, um neurônio GluA2+, um neurônio GluA4+, uma célula exterminadora natural NKG2D+, um astrócito SLC1A3+, um adipócito SLC7A10+ ou uma célula epitelial pulmonar CD30+.

[0096] 49. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,00001%, 0,0001%, 0,001%, 0,001%, 0,01%, 0,1%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% das células-alvo (por exemplo, uma ou mais de uma célula T, uma célula T CD3+, uma célula T CD4+, uma célula T CD8+, um hepatócito, uma célula-tronco hematepoiética, uma célula-tronco hematepoiética CD34+, um célula-tronco hematepoiética CD105+, uma célula-tronco hematepoiética CD117+, uma célula endotelial CD105+, uma célula B, uma célula B CD20+, uma célula B CD19+, uma célula cancerosa, uma célula cancerosa CD133+, uma célula cancerosa EpCAM+, uma célula cancelada CD19+, Célula cancerosa Her2/Neu+,

um neurônio GluA2+, um neurônio GluA4+, uma célula exterminadora natural NKG2D+, um astrócito SLC1A3+, um adipócito SLC7A10+ ou uma célula epitelial pulmonar CD30+) compreendem o ácido nucleico, por exemplo, usando PCR quantitativa, por exemplo, com o uso de um ensaio do Exemplo 3.

[0097] 50. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,00001%, 0,0001%, 0,001%, 0,001%, 0,01%, 0,1%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% das células-alvo (por exemplo, uma célula T CD3+, uma célula T CD4+, uma célula T CD8+, um hepatócito, uma célula-tronco hematepoiética, uma célula-tronco hematepoiética CD34+, um Célula-tronco hematepoiética CD105+, uma célula-tronco hematepoiética CD117+, uma célula endotelial CD105+, uma célula B, uma célula B CD20+, uma célula B CD19+, uma célula cancerosa, uma célula cancerosa CD133+, uma célula cancerosa EpCAM+, uma célula cancelada CD19+, Célula cancerosa Her2/Neu+, um neurônio GluA2+, um neurônio GluA4+, uma célula exterminadora natural NKG2D+, um astrócito SLC1A3+, um adipócito SLC7A10+ ou uma célula epitelial pulmonar CD30+) compreendem o agente exógeno.

[0098] 51. O fusossoma de qualquer uma das modalidades anteriores, em que, mediante administração, a razão entre células-alvo que compreendem o ácido nucleico, por exemplo, ácido nucleico retroviral, e células não alvo que compreendem o ácido nucleico, por exemplo, ácido nucleico retroviral, é pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa, por exemplo, com o uso dos ensaios do Exemplo 1 e Exemplo 3.

[0099] 52. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o número médio de cópia de ácido nucleico, por exemplo, ácido nucleico retroviral, ou uma porção dos mesmos

em células-alvo e o número médio de cópia de ácido nucleico, por exemplo, ácido nucleico retroviral, ou uma porção dos mesmos em células não alvo é pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa, por exemplo, com o uso de ensaios do Exemplo 1 e Exemplo 3.

[00100] 53. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o número mediano de cópia de ácido nucleico, por exemplo, ácido nucleico retroviral, ou uma porção dos mesmos em células-alvo e o número mediano de cópia de ácido nucleico, por exemplo, ácido nucleico retroviral, ou uma porção dos mesmos em células não alvo é pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa, por exemplo, com o uso de ensaios do Exemplo 1 e Exemplo 3.

[00101] 54. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre células-alvo que compreendem o agente de RNA exógeno e células não alvo que compreendem o agente de RNA exógeno é de pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa de transcrição reversa.

[00102] 55. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o nível médio de agente de RNA exógeno de células-alvo e o nível médio de agente de RNA exógeno de células não alvo é de pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa de transcrição reversa.

[00103] 56. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o nível mediano de agente de RNA exógeno de células-alvo e o nível mediano de agente de RNA exógeno de células não alvo é de pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500,

1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de PCR quantitativa de transcrição reversa.

[00104] 57. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre células-alvo que compreendem o agente de proteína exógeno e as células não-alvo que compreendem o agente de proteína exógeno é pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio de FACS, por exemplo, com o uso dos ensaios do Exemplo 2 e do Exemplo 4.

[00105] 58. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o nível médio de agente de proteína exógena de células-alvo e o nível médio de agente de proteína exógena de células não alvo é de pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio FACS, por exemplo, usando os ensaios do Exemplo 2 e/ou do Exemplo 4.

[00106] 59. O fusossoma de qualquer uma das modalidades anteriores, em que a razão entre o nível mediano do agente de proteína exógena de células-alvo e o nível mediano do agente de proteína exógena de células não alvo é de pelo menos 1,5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1.000, 5.000, 10.000, por exemplo, de acordo com um ensaio FACS, por exemplo, usando os ensaios do Exemplo 2 e/ou Exemplo 4.

[00107] 60. O fusossoma de qualquer uma das modalidades anteriores que compreende um ou ambos:

[00108] i) uma proteína imunossupressora exógena ou superexpressa na bicamada lipídica, por exemplo, envelope; e

[00109] ii) uma proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzida em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação

a um fusossoma gerado a partir de uma célula de origem não modificada de outra forma semelhante.

[00110] 61. O fusossoma de qualquer uma das modalidades anteriores que compreende um ou mais dentre:

[00111] i) uma primeira proteína imunossupressora exógena ou superexpressa na bicamada lipídica, por exemplo, envelope, e uma segunda proteína imunossupressora exógena ou superexpressa na bicamada lipídica, por exemplo, envelope;

[00112] ii) uma primeira proteína imunossupressora exógena ou superexpressa na bicamada lipídica, por exemplo, envelope, e uma segunda proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzida em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação a um fusossoma gerado a partir de uma célula de origem não modificada de outra forma semelhante; ou

[00113] iii) uma primeira proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzida em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação a um fusossoma gerado a partir de uma célula de origem não modificada de outra modo semelhante e uma segunda proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzido em pelo menos 10%, 20%, 30%, 40%, 50%, 60% 70%, 80% ou 90%) em comparação a um fusossoma gerado a partir de uma célula de origem não modificada semelhante de outra forma.

[00114] 62. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma permanece em circulação pelo menos 0,5, 1, 2, 3, 4, 6, 12, 18, 24, 36 ou 48 horas após a administração ao indivíduo.

[00115] 63. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%,

40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 30 minutos após a administração.

[00116] 64. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 1 hora após a administração.

[00117] 65. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 2 horas após a administração.

[00118] 66. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 4 horas após a administração.

[00119] 67. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 8 horas após a administração.

[00120] 68. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 12 horas após a administração.

[00121] 69. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 18 horas após a administração.

[00122] 70. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 24 horas após a administração.

[00123] 71. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 36 horas após a administração.

[00124] 72. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas permanecem em circulação 48 horas após a administração.

[00125] 73. O fusossoma de qualquer uma das modalidades anteriores que tem uma redução na imunogenicidade medida por uma redução na resposta humoral após uma ou mais administrações do fusossoma a um modelo animal apropriado, por exemplo, um modelo animal descrito no presente documento, em comparação a um retrovírus de referência, por exemplo, um fusossoma não modificado, de outro modo semelhante, ao fusossoma.

[00126] 74. O fusossoma da modalidade 73, em que a redução em resposta humoral é medida em uma amostra sérica por uma titulação de anticorpo anticelular, por exemplo, titulação de anticorpo antirretroviral, por exemplo, por ELISA.

[00127] 75. O fusossoma de qualquer uma das modalidades anteriores, em que uma amostra sérica de animais que receberam o fusossoma tem uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de uma titulação de anticorpo antifusossoma em comparação à amostra sérica de um indivíduo que recebeu uma célula não modificada.

[00128] 76. O fusossoma de qualquer uma das modalidades anteriores, em que uma amostra de soro de um indivíduo ao qual foi administrado o fusossoma tem uma titulação de anticorpo anticélula aumentado, por exemplo, aumentado em 1%, 2%, 5%, 10%, 20%, 30% ou 40% da linha basal, por exemplo, em que a linha basal se refere à

amostra de soro do mesmo indivíduo antes administração do fusossoma.

[00129] 77. O fusossoma de qualquer uma das modalidades anteriores em que:

[00130] o indivíduo que receberá o fusossoma ou uma composição farmacêutica que compreende o fusossoma tem, ou é conhecido por ter, ou é testado quanto a um anticorpo pré-existente (por exemplo, IgG ou IgM) reativo com o fusossoma;

[00131] o indivíduo ao qual será administrado o fusossoma não tem níveis detectáveis de um anticorpo pré-existente reativo com o fusossoma;

[00132] um indivíduo que recebeu o fusossoma ou uma composição farmacêutica que compreende o fusossoma tem, ou é conhecido por ter, ou é testado quanto a um anticorpo (por exemplo, IgG ou IgM) reativo com o fusossoma;

[00133] o indivíduo que recebeu o fusossoma ou uma composição farmacêutica que compreende o fusossoma (por exemplo, pelo menos uma, duas, três, quatro, cinco vezes ou mais) não tem níveis detectáveis de anticorpo reativo com o fusossoma; ou

[00134] os níveis de anticorpo não aumentam mais do que 1%, 2%, 5%, 10%, 20% ou 50% entre dois instantes do tempo, em que o primeiro instante do tempo ocorre antes da primeira administração do fusossoma, e o segundo momento ocorre após uma ou mais administrações do fusossoma.

[00135] 78. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma é produzido pelos métodos do Exemplo 5, 6 ou 7, por exemplo, de células transfectadas com HLA-G ou HLA-E cDNA.

[00136] 79. O fusossoma de qualquer uma das modalidades anteriores, em que fusossomas gerados das células NMC-HLA-G têm uma

porcentagem diminuída de lise, por exemplo, lise mediada por PBMC, lise mediada por célula NK e/ou lise mediada por célula T CD8+, em instantes específicos em comparação a fusossomas gerados por vetor vazio de NMCs ou NMC.

[00137] 80. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossomamodificado evita a fagocitose por macrófagos.

[00138] 81. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma é produzido pelos métodos do Exemplo 8, por exemplo, de células transfectadas com CD47 cDNA.

[00139] 82. O fusossoma de qualquer uma das modalidades anteriores, em que o índice fagocítico é reduzido quando macrófagos são incubados com fusossomas derivados de NMC-CD47, contra aqueles derivados de vetor de NMC ou vetor sem NMC.

[00140] 83. O fusossoma de qualquer uma das modalidades anteriores, que tem uma redução na fagocitose de macrófagos, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais em macrófagos fagocitose em comparação a um fusossoma de referência, por exemplo, um fusossoma não modificado de outro modo semelhante ao fusossoma, em que a redução na fagocitose de macrófagos é determinada pelo ensaio do índice de fagocitose *in vitro*, por exemplo, conforme descrito no Exemplo 8.

[00141] 84. O fusossoma de qualquer uma das modalidades anteriores, em que a composição de fusossoma tem um índice de fagocitose de 0, 1, 10, 100 ou mais, por exemplo, conforme medido por um ensaio do Exemplo 8, quando incubado com macrofágos em um ensaio *in vitro* de fagocitose de macrófago.

[00142] 85. O fusossoma de qualquer uma das modalidades anteriores, que é modificado e tem atividade do complemento reduzida em comparação a um fusossoma não modificado.

[00143] 86. O fusossoma de qualquer uma das modalidades anteriores que é produzido pelos métodos do Exemplo 9, por exemplo, a partir de células submetidas à transfecção com um cDNA que codifica para proteínas regulatórias do complemento, por exemplo, DAF.

[00144] 87. O fusossoma de qualquer uma das modalidades anteriores, em que a dose de fusossoma na qual 200 pg/ml de C3a está presente para o fusossoma modificado (por exemplo, HEK293-DAF) incubado com soros correspondentes de camundongo (por exemplo, soros de camundongo HEK-293 DAF) é maior que o fusossoma de referência (por exemplo, vetor retroviral HEK293) incubado com soros correspondentes de camundongo (por exemplo, soros de camundongo HEK293).

[00145] 88. O fusossoma de qualquer uma das modalidades anteriores, em que a dose do fusossoma na qual 200 pg/ml do C3a está presente é maior para o fusossoma modificado (por exemplo, HEK293-DAF) incubado com soros de camundongo naïve do que para o fusossoma de referência (por exemplo, vetor retroviral de HEK293) incubados com soros de camundongo naïve.

[00146] 89. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma é resistente a inativação mediada por complemento no soro do paciente 30 minutos após a administração de acordo com um ensaio do Exemplo 9.

[00147] 90. O fusossoma de qualquer uma das modalidades anteriores, em que pelo menos 0,001%, 0,01%, 0,1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% de fusossomas são resistentes à inativação mediada pelo complemento.

[00148] 91. O fusossoma de qualquer uma das modalidades 86 a 90, em que a proteína reguladora do complemento compreende um ou mais dentre as proteínas que se ligam ao fator acelerador de decaimento (DAF, CD55), por exemplo, proteína-1 semelhante ao fator H

(FH) (FHL-1), por exemplo, proteína de ligação a C4b (C4BP), por exemplo, receptor de complemento 1 (CD35), por exemplo, proteína cofator de membrana (MCP, CD46), por exemplo, Profectina (CD59), por exemplo, proteínas que inibem as enzimas CD/C5 convertase da via clássica e alternativa do complemento, por exemplo, proteínas que regulam a montagem do MAC.

[00149] 92. O fusossoma de qualquer uma das modalidades anteriores, que é produzido pelos métodos do Exemplo 10, por exemplo, a partir de células submetidas à transfecção com um DNA que codifica para um shRNA direcionado a classe I de MHC, por exemplo, em que vetores retrovirais derivados de classe I de NMC-shMHC têm expressão mais baixa de classe I de MHC em comparação a NMCs e controle de vetor de NMC.

[00150] 93. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida de imunogenicidade é inativação do soro, por exemplo, inativação do soro medida, conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 11.

[00151] 94. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno não é diferente entre amostras de fusossoma que foram incubadas com soro e soro inativado por calor de camundongos naïve.

[00152] 95. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno não é diferente entre amostras de fusossoma que foram incubadas com soro de camundongos naïve com fusossoma e incubações de controle sem soro.

[00153] 96. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno em amostras de fusossoma que foram incubadas com soro de controle positivo é menor que em amostras de fusossoma que foram incu-

badas com soro de camundongos naïve com fusossoma.

[00154] 97. O fusossoma de qualquer uma das modalidades anteriores, em que um fusossoma modificada, por exemplo, modificada por um método descrito no presente documento, tem uma inativação do soro reduzido (por exemplo, reduzido em comparação à administração de um fusossoma não modificado) após múltiplas (por exemplo, mais de uma, por exemplo, 2 ou mais), administrações do fusossoma modificada.

[00155] 98. O fusossoma de qualquer uma das modalidades anteriores, em que um fusossoma descrito no presente documento não é inativado pelo soro após múltiplas administrações.

[00156] 99. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida de imunogenicidade para o fusossoma é a inativação do soro, por exemplo, após múltiplas administrações, por exemplo, inativação do soro após múltiplas administrações, medida conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 12.

[00157] 100. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno não é diferente entre amostras de fusossoma que foram incubadas com soro e soro inativadas por calor de camundongos tratados com (por exemplo, HEK293-HLA-G) fusossoma modificado.

[00158] 101. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno não é diferente entre amostras de fusossoma que foram incubadas dos camundongos tratados 1, 2, 3, 5 ou 10 vezes com fusossomas modificados (por exemplo, HEK293-HLA-G).

[00159] 102. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno não é diferente entre amostras de fusossoma que foram incuba-

das com soro de camundongos tratados com veículo e de camundongos tratados com fusossomas modificados (por exemplo, HEK293-HLA-G).

[00160] 103. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células que recebem o agente exógeno para fusossomas derivados de uma célula de referência (por exemplo, HEK293) é menor do que para fusossomas modificados (por exemplo, HEK293-HLA-G).

[00161] 104. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida da imunogenicidade para um fusossoma é a resposta de anticorpo.

[00162] 105. O fusossoma de qualquer uma das modalidades anteriores, em que um indivíduo que recebe um fusossoma descrito no presente documento tem anticorpos pré-existentes que se ligam ao fusossoma e reconhecem o mesmo, por exemplo, medido, conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 13.

[00163] 106. O fusossoma de qualquer uma das modalidades anteriores, em que soro de camundongos naïve de fusossoma mostra mais sinal (por exemplo, fluorescência) do que o controle negativo, por exemplo, soro de um camundongo que sofreu depleção de IgM e IgG, por exemplo, indicando que imunogenicidade ocorreu.

[00164] 107. O fusossoma de qualquer uma das modalidades anteriores, em que o soro de camundongos naïve com fusossoma mostram sinal semelhante (por exemplo, fluorescência) em comparação ao controle negativo, por exemplo, o que indica que a imunogenicidade não ocorreu de forma detectável.

[00165] 108. O fusossoma de qualquer uma das modalidades anteriores, que é um fusossoma modificado, por exemplo, modificado por um método descrito no presente documento, e que foi reduzido (por

exemplo, reduzido em comparação à administração de um fusossoma não modificado) resposta humoral após múltiplas (por exemplo, mais de uma, por exemplo, 2 ou mais), administração do fusossoma modificado, por exemplo, medido, conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 14.

[00166] 109. O fusossoma de qualquer uma das modalidades anteriores, em que o fusossoma é produzido pelos métodos do Exemplo 5, 6, 7 ou 14, por exemplo, de células submetidas à transfecção com HLA-G ou HLA-E cDNA.

[00167] 110. O fusossoma de qualquer uma das modalidades anteriores, em que a resposta humoral é avaliada determinando-se um valor para o nível de anticorpos antifusossoma (por exemplo, anticorpos de IgM, IgG1 e/ou IgG2).

[00168] 111. O fusossoma de qualquer uma das modalidades anteriores, em que fusossomas modificados (por exemplo, NMC-HLA-G) têm titulações diminuídas de anticorpo IgM ou IgG1/2 antiviral (por exemplo, conforme medido por intensidade de fluorescência em FACS) após injeções, em comparação a um controle, por exemplo, fusossomas de NMC ou fusossomas vazios de NMC.

[00169] 112. O fusossoma de qualquer uma das modalidades anteriores, em que as células receptoras não são realvejadas por uma resposta de anticorpo, ou uma resposta de anticorpo estará abaixo de um nível de referência, por exemplo, medido conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 15.

[00170] 113. O fusossoma de qualquer uma das modalidades anteriores, em que o sinal (por exemplo, intensidade média de fluorescência) é semelhante para células receptoras de camundongos de camundongos tratados com fusossomas e camundongos tratados com PBS.

[00171] 114. O fusossoma de qualquer uma das modalidades ante-

riores, em que uma medida da imunogenicidade das células receptoras é a resposta do macrófago.

[00172] 115. O fusossoma de qualquer uma das modalidades anteriores, em que as células receptoras não são realvejadas por macrófagos ou são alvejadas abaixo de um nível de referência.

[00173] 116. O fusossoma de qualquer uma das modalidades anteriores, em que o índice fagocítico, por exemplo, medido conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 16, é semelhante para células receptoras derivadas de camundongos tratados com fusossomas e camundongos tratados com PBS.

[00174] 117. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida da imunogenicidade das células receptoras é a resposta de PBMC.

[00175] 118. O fusossoma de qualquer uma das modalidades anteriores, em que as células receptoras não elicitam uma resposta de PBMC.

[00176] 119. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células CD3+/CMG+ é semelhante para células receptoras derivadas de camundongos tratados com fusossoma e camundongos tratados com PBS, por exemplo, conforme medido conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 17.

[00177] 120. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida da imunogenicidade das células receptoras é a resposta natural da célula exterminadora.

[00178] 121. O fusossoma de qualquer uma das modalidades anteriores, em que as células receptoras não elicitam uma resposta natural da célula exterminadora ou induzem uma resposta natural inferior da célula exterminadora, por exemplo, inferior a um valor de referência.

[00179] 122. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células CD3+/CMG+ é semelhante para células receptoras derivadas de camundongos tratados com fusossoma e camundongos tratados com PBS, por exemplo, conforme medido conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 18.

[00180] 123. O fusossoma de qualquer uma das modalidades anteriores, em que uma medida da imunogenicidade das células receptoras é a resposta da célula T CD8+.

[00181] 124. O fusossoma de qualquer uma das modalidades anteriores das modalidades anteriores, em que as células receptoras não elicitam uma resposta da célula T CD8+ ou não elicitam uma resposta inferior da célula T CD8+, por exemplo, inferior a um valor de referência.

[00182] 125. O fusossoma de qualquer uma das modalidades anteriores, em que a porcentagem de células CD3+/CMG+ é semelhante para células receptoras derivadas de camundongos tratados com fusossoma e camundongos tratados com PBS, por exemplo, conforme medido conforme descrito no presente documento, por exemplo, conforme descrito no Exemplo 19.

[00183] 126. O fusossoma de qualquer uma das modalidades anteriores, em que o fusogênio é um fusogênio realvejado.

[00184] 127. O fusossoma de qualquer uma das modalidades anteriores que compreende um ácido nucleico, por exemplo, ácido nucleico retroviral, que codifica um ou dois dentre: (i) um elemento regulatório específico de célula-alvo positivo ligado operacionalmente a um ácido nucleico que codifica um agente exógeno; ou (ii) um elemento regulatório específico de célula não alvo ou TCSRE negativo ligado operacionalmente ao ácido nucleico que codifica o agente exógeno.

[00185] 128. Uma composição farmacêutica que compreende o fu-

sossoma de qualquer uma das modalidades anteriores e um carreador, diluente ou excipiente farmaceuticamente aceitável.

[00186] 129. Um método para entregar um agente exógeno a um indivíduo (por exemplo, um indivíduo humano) que compreende administrar ao indivíduo um fusossoma de acordo com qualquer uma das modalidades 1 a 127 ou uma composição farmacêutica da modalidade 128, desse modo, entregando o agente exógeno ao indivíduo.

[00187] 130. Um método para modular uma função em um indivíduo (por exemplo, um indivíduo humano), tecido-alvo (por exemplo, do fígado) ou célula-alvo (por exemplo, célula hepática, por exemplo, hepatócito), que compreende colocar, por exemplo, administrar, o indivíduo, o tecido-alvo ou a célula-alvo em contato com um fusossoma de qualquer uma das modalidades 1 a 127, ou a composição farmacêutica da modalidade 128.

[00188] 131. O método da modalidade 130, em que o tecido-alvo ou a célula tecidual está presente em um indivíduo.

[00189] 132. Um método de tratamento de uma deficiência genética em um indivíduo (por exemplo, um indivíduo humano) que compreende administrar ao indivíduo um fusossoma de qualquer uma das modalidades 1 a 127 ou a composição farmacêutica da modalidade 128.

[00190] 133. O método da modalidade 132, em que a deficiência genética é uma deficiência genética da Tabela 5.

[00191] 134. O método da modalidade 132 ou 133, em que a deficiência genética é uma deficiência genética que pode ser tratada pelo gene de carga útil que codifica o agente exógeno.

[00192] 135. Um fusossoma de qualquer uma das modalidades 1 a 127 ou composição farmacêutica da modalidade 128 para uso no tratamento de um indivíduo (por exemplo, um indivíduo humano) com uma deficiência genética.

[00193] 136. Uso de um fusossoma de qualquer uma das modali-

dades 1 a 127 ou composição farmacêutica da modalidade 128 para a fabricação de um medicamento para uso no tratamento de um indivíduo (por exemplo, um indivíduo humano) com uma deficiência genética.

[00194] 137. O fusossoma ou composição farmacêutica para uso da modalidade 135 ou o uso das modalidades 136, em que o fusossoma compreende um gene de carga útil que codifica um agente exógeno para tratar a deficiência genética.

[00195] 138. Um método para produzir um fusossoma de qualquer uma das modalidades 1 a 127 que compreende:

[00196] a) fornecer uma célula que compreende o ácido nucleico, por exemplo, ácido nucleico retroviral, e o fusogênio;

[00197] b) cultivar a célula em condições que permitam a produção do fusossoma, e

[00198] c) separar, enriquecer ou purificar o fusossoma a partir da célula, produzindo, assim, o fusossoma.

[00199] Outras características, objetos e vantagens da invenção ficarão evidentes a partir da descrição e dos desenhos, e das reivindicações.

[00200] Salvo quando definido de outro modo, todos os termos técnicos e científicos usados no presente documento têm significado igual ao comumente entendido por uma pessoa versada na técnica à qual esta invenção pertence. Todas as publicações, pedidos de patentes, patentes e outras referências mencionados no presente documento são incorporadas a título de referência em sua totalidade. Por exemplo, todas as sequências de GenBank, Unigene e Entrez citadas no presente documento, por exemplo, em qualquer Tabela no presente documento, são incorporadas a título de referência. Salvo quando especificado de outro modo, os números de acesso de sequência especificados no presente documento, incluindo em qualquer Tabela no pre-

sente documento, se referem às entradas do banco de dados atuais em 15 de maio de 2018. Quando um gene ou proteína faz referência a uma pluralidade de números de acesso de sequência, todas as variantes de sequência são abrangidas. Além disso, os materiais, métodos e exemplos servem apenas a título de ilustração e não devem ser limitativos.

BREVE DESCRIÇÃO DOS DESENHOS

[00201] A descrição detalhada a seguir da invenção será mais bem compreendida quando lida em combinação com os desenhos anexos. A fim de ilustrar a invenção, são mostradas nos desenhos descritos no presente documento algumas modalidades que são exemplificativas aqui. No entanto, deve-se entender que a invenção não se limita à exata disposição e instrumentalidades das modalidades mostradas nos desenhos.

[00202] A Figura 1 quantifica a marcação de fusossomas com um corante para F-actina.

[00203] A Figura 2 é um gráfico que mostra a capacidade que fusossomas e células parentais têm para polimerizar a actina durante um período de 3, 5 e 24 horas.

[00204] A Figura 3 é uma tabela que mostra estatísticas de distribuição de tamanho de fusossomas e células parentais, conforme medido por NTA e microscopia.

[00205] A Figura 4 é uma tabela que mostra o tamanho e volume médios dos fusossomas e células parentais.

[00206] A Figura 5 é uma série de diagramas que mostram a razão solúvel:insolúvel observada para fusossomas ou uma preparação de células.

[00207] A Figura 6 é uma série de diagramas que mostram fusão do fusossoma MvH(CD8)+ F a células-alvo ou não alvo e a quantidade absoluta de fusão alvejada.

[00208] A Figura 7 é um diagrama que mostra intensidade de fluorescência média 2-NBDG em fusossomas de VSV-G.

[00209] A Figura 8 é um diagrama que mostra a atividade esterase no citosol de fusossomas VSV-G.

[00210] As Figuras 9A a 9B são uma série de diagramas que mostram a entrega de Cre recombinase por meio dos fusossomas, conforme detectado por imagem por bioluminescência em camundongos. (A) Imagem ventral e sobreposição de sinal luminescente do fígado e do baço expostos de camundongos tratados com fusossoma IV (concentração de 1x e 3x). A parte inferior é apenas o sinal luminescente. (B) Sinal de fluxo total de baço e fígado alvejados por fusossoma; a escala y está na escala log10. Os camundongos tratados com uma concentração de 3x tratamento com fusossoma tiveram um sinal significativamente maior no baço ($p = 0,0004$) do que o segundo plano 72 horas após o tratamento.

[00211] As Figuras 10A a 10B são uma série de diagramas que mostram a Cre recombinase para fígado e baço de murganho por meio dos fusossomas, conforme detectado por imageamento por bioluminescência. (A) Da esquerda para a direita; a sobreposição de imagem dorsal e sinal luminescente do fígado, coração, pulmões, rim, intestino delgado, pâncreas e baço submetido à excisão e coletados e imageados 5 minutos após a eutanásia. A parte inferior é apenas o sinal luminescente. (B) Sinal de fluxo total de fusossoma alvejou o baço e fígado e outros tecidos; a escala y está na escala log10. Os camundongos tratados com uma concentração de tratamento com fusossoma 3x tiveram um sinal significativamente maior no baço ($p <0,0001$) em comparação ao tecido com o sinal mais baixo (coração).

[00212] A Figura 11 é uma tabela que mostra a entrega de carga Cre por fusossomas NivG+F por meio de uma via não endocítica.

[00213] A Figura 12 é um gráfico que mostra as razões GA-

PDH:Total de proteína medidas por ensaio de ácido bicinconínico em fusossomas e células parentais.

[00214] A Figura 13 é um gráfico que mostra as razões lipídio:proteína medidas por ensaio de ácido bicinconínico em fusossomas e células parentais.

[00215] A Figura 14 é um gráfico que mostra as razões proteína:DNA medidas por ensaio de ácido bicinconínico em fusossomas e células parentais.

[00216] A Figura 15 é um gráfico que mostra as razões lipídios:DNA medidas por ensaio de ácido bicinconínico em fusossomas e células parentais.

[00217] A Figura 16 é um gráfico que mostra os níveis de proteína do marcador de exossoma CD63 em exossomas e fusossomas.

[00218] A Figura 17 é um gráfico que mostra a intensidade do sinal de calnexina detectado em fusossomas e células parentais.

[00219] A Figura 18 é um gráfico que mostra as razões lipídio:DNA determinadas para fusossomas e células parentais.

[00220] As Figuras 19A a 19B são uma série de gráficos que mostram a proporção de espécies de lipídios como uma porcentagem dos lipídios totais nas células parentais, exossomas e fusossomas.

[00221] A Figura 20 é uma série de gráficos que mostram o teor de proteína de células parentais, exossomas e fusossomas com relação às proteínas associadas a compartimentos específicos, conforme indicado.

[00222] A Figura 21 é uma série de gráficos que mostram o nível de ARRDC1 (painel esquerdo) ou TSG101 (painel direito) como uma porcentagem do teor total da proteína em células parentais, exossomas e fusossomas.

[00223] As Figuras 22A a 22C mostram resultados para linhagens celulares, incluindo linhagens celulares de heptoma humano alvo

(HepG2) e linhagens celulares (não hepáticas) não alvo, submetidos à transdução com lentivírus (LV) que codificam construtos de ácido nucleico que contêm TCSREs positivos ou NTCSREs. A Figura 22A mostra a expressão de GFP na linhagem de células de hepatoma humano (HepG2), linhagem de células renais embrionárias humanas (293LX), linhagem de células T humanas de origem hematopoiética (Molt4.8) e linhagem de células endoteliais derivadas de cérebro de camundongo (bEND.3) submetidas à transdução com LV gerado com sequências miRT (hPGK-eGFP+miRT) ou sem sequências miRT (hPGK-eGFP), sob o controle do promotor PGK. A Figura 22B mostra a expressão de GFP em células HepG2 e 293LX submetidas à transdução com LV gerado sob o controle do promotor PGK (hPGK-eGFP) ou LVs que contêm sequências mirT e GFP sob o controle do promotor específico de hepatócitos ApoE (hApoE-eGFP+miRT). A Figura 22C mostra a quantificação de fenilalanina (Phe) no sobrenadante de células HepG2 e 293LX submetidas à transdução com LVs que contêm o transgene fenilalanina amônia-liase (PAL) sob o controle do promotor SFFV (SFFV-PAL) ou LVs que contêm sequências mirT e sob o controle de o promotor hApoE (hApoE-PAL+miRT).

DESCRIÇÃO DETALHADA

A presente divulgação fornece, pelo menos em parte, métodos e composições de fusossoma para entrega *in vivo*. Em algumas modalidades, o fusossoma compreende uma combinação de elementos que promovem a especificidade para células-alvo, por exemplo, um ou mais dentre um fusogênio realvejado, um elemento regulatório específico de célula-alvo positivo e um elemento regulatório específico de célula não alvo. Em algumas modalidades, o fusossoma compreende uma ou mais modificações que diminuem uma resposta imunológica contra o fusossoma.

DEFINIÇÕES

[00224] Os termos usados nas reivindicações e no relatório descritivo são definidos conforme estabelecido a seguir, salvo quando especificado de outro modo.

[00225] Conforme usado no presente documento, o “presente de modo detectável”, quando usado no contexto de um agente exógeno estando presente de forma detectável, significa que o próprio agente exógeno está presente de forma detectável. Por exemplo, caso o agente exógeno seja uma proteína, o agente de proteína exógena pode estar presente de forma detectável, independentemente da possibilidade de um ácido nucleico que o codifica estar presente ou não de forma detectável.

[00226] Conforme utilizado no presente documento, “fusossoma” se refere a uma bicamada de lipídios anfipáticos que envolvem um lúmen ou cavidade e um fusogênio que interage com a bicamada lipídica anfipática. Nas modalidades, o fusossoma compreende um ácido nucleico. Em algumas modalidades, o fusossoma é uma preparação fechada por membrana. Em algumas modalidades, o fusossoma é derivado de uma célula de origem.

[00227] Conforme utilizado no presente documento, “composição de fusossoma” se refere a uma composição que compreende um ou mais fusossomas.

[00228] Conforme utilizado no presente documento, “fusogênio” se refere a um agente ou molécula que cria uma interação entre dois lúmens fechados por membrana. Nas modalidades, o fusogênio facilita a fusão das membranas. Em outras modalidades, o fusogênio cria uma conexão, por exemplo, um poro, entre dois lúmens (por exemplo, um lúmen de um vetor retroviral e um citoplasma de uma célula-alvo). Em algumas modalidades, o fusogênio compreende um complexo de duas ou mais proteínas, por exemplo, em que nenhuma proteína tem ativi-

dade fusogênica sozinha. Em algumas modalidades, o fusogênio compreende um domínio de alvejamento.

[00229] Conforme usado no presente documento, um “elemento isolante” se refere a uma sequência de nucleotídeos que bloqueia intensificadores ou impede que a heterocromatina se espalhe. Um elemento isolante pode ser do tipo selvagem ou mutante.

[00230] O termo “quantidade eficaz”, conforme usado no presente documento, significa uma quantidade de uma composição farmacêutica que é suficiente para modificar significativa e positivamente os sintomas e/ou condições a serem tratadas (por exemplo, fornecer uma resposta clínica positiva). A quantidade eficaz de um ingrediente ativo para uso em uma composição farmacêutica variará conforme a afecção particular a ser tratada, a gravidade da afecção, a duração do tratamento, a natureza da terapia simultânea, o ingrediente (ou ingredientes) ativo particular que é usue, o excipiente (ou excipientes) e/ou carreador (ou carreadores) específico farmaceuticamente aceitáveis usados, e fatores semelhantes com o conhecimento e experiência do médico assistente.

[00231] Um “agente exógeno”, conforme usado no presente documento com referência a um vírus, VLP ou fusossoma, se refere a um agente que não é comprometido tampouco codificado no vírus do tipo selvagem correspondente ou fusogênio produzido a partir de uma célula de origem do tipo selvagem. Em algumas modalidades, o agente exógeno não existe naturalmente, como uma proteína ou ácido nucleico que tem uma sequência que é alterada (por exemplo, por inserção, deleção ou substituição) em relação a uma proteína de ocorrência natural. Em algumas modalidades, o agente exógeno não existe naturalmente na célula de origem. Em algumas modalidades, o agente exógeno existe naturalmente na célula de origem, porém é exógeno ao vírus. Em algumas modalidades, o agente exógeno não existe natu-

ralmente na célula receptora. Em algumas modalidades, o agente exógeno existe naturalmente na célula receptora, porém não está presente em um nível desejado ou em um momento desejado. Em algumas modalidades, o agente exógeno compreende RNA ou proteína.

[00232] O termo “farmaceuticamente aceitável”, conforme usado no presente documento, se refere a excipientes, composições e/ou formas de dosagem que são, dentro do escopo do julgamento médico preciso, adequados para uso em contato com os tecidos de seres humanos e animais sem toxicidade excessiva, irritação, resposta alérgica ou outro problema ou complicação, proporcional a uma relação benefício/risco razoável.

[00233] Conforme usado no presente documento, um “promotor” se refere a uma sequência de DNA cis-regulatória que, quando operacionalmente ligada a uma sequência de codificação de gene, conduz a transcrição do gene. O promotor pode compreender locais de ligação a um fator de transcrição. Em algumas modalidades, um promotor trabalha em combinação com um ou mais intensificadores que são distais em relação ao gene.

[00234] Conforme usado no presente documento, um “elemento regulatório específico de célula-alvo positivo” (ou TCSRE positivo) se refere a uma sequência de ácidos nucleicos que aumenta o nível de um agente exógeno em uma célula-alvo em comparação a uma célula não alvo, em que o ácido nucleico que codifica o agente exógeno está ligado operacionalmente ao TCSRE positivo. Em algumas modalidades, o TCSRE positivo é uma sequência de ácidos nucleicos funcional, por exemplo, o TCSRE positivo pode compreender um promotor ou intensificador. Em algumas modalidades, o TCSRE positivo codifica uma sequência de RNA funcional, por exemplo, o TCSRE positivo pode codificar um sítio de splice que promove o splicing correto do RNA na célula-alvo. Em algumas modalidades, o TCSRE positivo codifica

uma sequência de proteína funcional, ou o TCSRE positivo pode codificar uma sequência de proteína que promove a modificação pós-tradução correta da proteína. Em algumas modalidades, o TCSRE positivo diminui o nível ou atividade de um regulatório descendente ou inibidor do agente exógeno. Em algumas modalidades, a célula-alvo é uma célula hepática e o elemento regulatório específico de célula-alvo positivo é um elemento regulatório específico de célula do fígado positivo.

[00235] Conforme usado no presente documento, um “elemento regulatório específico de célula-alvo negativo” (ou TCSRE negativo) se refere a uma sequência de ácidos nucleicos que diminui o nível de um agente exógeno em uma célula não alvo em comparação com uma célula-alvo, em que o ácido nucleico que codifica o agente exógeno está ligado operacionalmente ao TCSRE negativo. Em algumas modalidades, o TCSRE negativo é uma sequência de ácidos nucleicos funcional, por exemplo, um sítio de reconhecimento de miRNA que causa degradação ou inibição do ácido nucleico retroviral em uma célula não alvo. Em algumas modalidades, a sequência de ácidos nucleicos codifica uma sequência de RNA funcional, por exemplo, o ácido nucleico codifica uma sequência de miRNA presente em um mRNA que codifica um agente de proteína exógena, para que o mRNA seja degradado ou inibido em uma célula não alvo. Em algumas modalidades, o TCSRE negativo aumenta o nível ou atividade de um regulatório descendente ou inibidor do agente exógeno. Em algumas modalidades, a célula não alvo é uma célula não hepática.

[00236] Conforme usado no presente documento, um “elemento regulatório específico de célula não alvo” (ou NTCSRE) se refere a uma sequência de ácidos nucleicos que diminui o nível de um agente exógeno em uma célula não alvo em comparação a uma célula-alvo, em que o ácido nucleico que codifica o agente exógeno está ligado

operacionalmente ao NTCSRE. Em algumas modalidades, o NTCSRE negativo é uma sequência de ácidos nucleicos funcional, por exemplo, um sítio de reconhecimento de miRNA que causa degradação ou inibição do ácido nucleico retroviral em uma célula não alvo. Em algumas modalidades, a sequência de ácidos nucleicos codifica uma sequência de RNA funcional, por exemplo, o ácido nucleico codifica uma sequência de miRNA presente em um mRNA que codifica um agente de proteína exógena, para que o mRNA seja degradado ou inibido em uma célula não alvo. Em algumas modalidades, o NTCSRE aumenta o nível ou atividade de um regulatório descendente ou inibidor do agente exógeno. Em algumas modalidades, a célula não alvo é uma célula não hepática e o elemento regulatório específico de célula não alvo é um elemento regulatório específico de célula não hepática. Os termos “TCSRE negativo” e “NTCSRE” são usados de maneira intercambiável no presente documento.

[00237] Conforme usado no presente documento, um “elemento regulatório específico de célula não hepática” se refere a um elemento regulatório específico de célula não alvo (NTCSRE), em que a célula-alvo é uma célula hepática. Dito isso, um elemento regulatório específico de célula não hepática se refere a uma sequência de ácidos nucleicos que diminui o nível de um agente exógeno em uma célula não hepática (por exemplo, em uma célula imune) ou tecido em comparação a uma célula hepática, em que o ácido que codifica o agente exógeno está ligado operacionalmente ao elemento regulatório específico de célula não hepática.

[00238] Conforme usado no presente documento, um “fusogênio realvejado” se refere a um fusogênio que compreende uma porção química de alvejamento que tem uma sequência que não faz parte da forma de ocorrência natural do fusogênio. Nas modalidades, o fusogênio compreende uma porção química de alvejamento diferente em re-

lação à porção química de alvejamento na forma de ocorrência natural do fusogênio. Nas modalidades, a forma de ocorrência natural do fusogênio carece de um domínio de alvejamento, e o fusogênio realvejado compreende uma porção química de alvejamento que está ausente da forma de ocorrência natural do fusogênio. Nas modalidades, o fusogênio é modificado para compreender uma porção química de alvejamento. Nas modalidades, o fusogênio compreende uma ou mais alterações de sequência fora da porção química de alvejamento em relação à forma de ocorrência natural do fusogênio, por exemplo, em um domínio transmembranar, domínio fusogenicamente ativo ou domínio citoplasmático.

[00239] Conforme usado no presente documento, um “ácido nucleico retroviral” se refere a um ácido nucleico que contém pelo menos as exigências mínimas de sequência para empacotamento em um retrovírus ou vetor retroviral isoladamente ou em combinação com uma célula auxiliar, vírus auxiliar ou plasmídeo auxiliar. Em algumas modalidades, o ácido nucleico retroviral compreende adicionalmente ou codifica um agente exógeno, um elemento regulatório específico de célula-alvo positivo, um elemento regulatório específico de célula não alvo ou um TCSRE negativo. Em algumas modalidades, o ácido nucleico retroviral compreende uma ou mais dentre (por exemplo, todos) uma LTR 5’ (por exemplo, para promover integração), uma U3 (por exemplo, para ativar transcrição de RNA genômico viral), R (por exemplo, uma região de ligação Tat), U5, 3’ (por exemplo, para promover a integração), um sítio de empacotamento (por exemplo, psi (Ψ)), RRE (por exemplo, para se ligar a Rev e promover exportação nuclear). O ácido nucleico retroviral pode compreender RNA (por exemplo, quando parte de um vírion) ou DNA (por exemplo, quando é introduzido em uma célula de origem ou após a transcrição reversa em uma célula receptora). Em algumas modalidades, o ácido nucleico retroviral é empacotado com o

uso de uma célula auxiliar, vírus auxiliar ou plasmídeo auxiliar que compreende um ou mais dentre (por exemplo, todos) gag, pol e env.

[00240] Conforme usado no presente documento, uma “célula-alvo” se refere a uma célula de um tipo à qual é desejado que um fusossoma (por exemplo, vetor lentiviral) entregue um agente exógeno. Nas modalidades, uma célula-alvo é uma célula de um tipo ou classe de tecido específico, por exemplo, uma célula efetora imunológica, por exemplo, uma célula T. Em algumas modalidades, uma célula-alvo é uma célula doente, por exemplo, uma célula cancerosa. Em algumas modalidades, o fusogênio, por exemplo, fusogênio realvejado (isoladamente ou em combinação com o TCSRE positivo, NTCSRE, TCSRE negativo ou qualquer combinação dos mesmos) causa a entrega preferencial do agente exógeno a uma célula-alvo em comparação a uma célula não alvo.

[00241] Conforme usado no presente documento uma “célula não alvo” se refere a uma célula de um tipo ao qual não é desejável que um vetor lentiviral entregue um agente exógeno. Em algumas modalidades, uma célula não alvo é uma célula de um tipo ou classe de tecido específico. Em algumas modalidades, uma célula não alvo é uma célula não doente, por exemplo, uma célula não cancerosa. Em algumas modalidades, o fusogênio, por exemplo, fusogênio realvejado (sozinho ou em combinação com o TCSRE positivo, NTCSRE, TCSRE negativo ou qualquer combinação dos mesmos) causa a entrega mais baixa do agente exógeno a uma célula não alvo em comparação a uma célula-alvo.

[00242] Conforme usado no presente documento, os termos “tratar”, “que trata” ou “tratamento” se referem a atenuar uma doença ou distúrbio, por exemplo, retardar ou interromper ou reduzir o desenvolvimento da doença ou distúrbio, por exemplo, uma causa-raiz do distúrbio ou pelo menos um dos seus sintomas clínicos.

[00243] Conforme usado no presente documento, “citobiológico” se refere a uma porção de uma célula que compreende um lúmen e uma membrana celular ou uma célula com inativação nuclear parcial ou completa. Em algumas modalidades, o citobiológico compreende um ou mais dentre um componente do citoesqueleto, uma organela e um ribossoma. Em modalidades, o citobiológico é uma célula enucleada, uma microvesícula ou um fantasma de célula.

FUSOSSOMAS, POR EXEMPLO, FUSOSSOMAS DERIVADOS DE CÉLULAS

[00244] Os fusossomas podem assumir várias formas. Por exemplo, em algumas modalidades, um fusossoma descrito no presente documento é derivado de uma célula de origem. Um fusossoma pode compreender, por exemplo, uma vesícula extracelular, uma microvesícula, uma nanovesícula, um exossomo, um corpo apoptótico (de células apoptóticas), uma micropartícula (que pode ser derivada de, por exemplo, plaquetas), um ectossoma (derivável de, por exemplo, neutrófilos e monócitos no soro), um prostatossoma (obtido a partir de células de câncer de próstata), um cardiossoma (derivado de células cardíacas) ou qualquer combinação dos mesmos. Em algumas modalidades, um fusossoma é liberado naturalmente de uma célula de origem e, em algumas modalidades, a célula de origem é tratada para aumentar a formação de fusossomas. Em algumas modalidades, o fusossoma tem entre de aproximadamente 10 a 10.000 nm de diâmetro, por exemplo, de aproximadamente 30 a 100 nm de diâmetro. Em algumas modalidades, o fusossoma compreende um ou mais lipídios sintéticos.

[00245] Em algumas modalidades, o fusossoma é ou compreende um vírus, por exemplo, um retrovírus, por exemplo, um lentivírus. De acordo com uma modalidade da invenção, um fusossoma que compreende uma bicamada lipídica compreende um vetor retroviral que

compreende um envelope. Por exemplo, em algumas modalidades, a bicamada do fusossoma de lipídios anfipáticos é ou compreende o envelope viral. O envelope viral pode compreender um fusogênio, por exemplo, um fusogênio que é endógeno ao vírus ou um fusogênio pseudotipado. Em algumas modalidades, o lúmen ou a cavidade do fusossoma compreende um ácido nucleico viral, por exemplo, um ácido nucleico retroviral, por exemplo, um ácido nucleico lentiviral. O ácido nucleico viral pode ser um genoma viral. Em algumas modalidades, o fusossoma compreende adicionalmente uma ou mais proteínas não estruturais virais, por exemplo, em sua cavidade ou lúmen.

[00246] Os fusossomas podem ter várias propriedades que facilitem a entrega de uma carga útil, tal como um transgene desejado, ou codificação de um agente exógeno, para uma célula-alvo. Por exemplo, em algumas modalidades, o fusossoma e a célula de origem juntos compreendem ácido nucleico (ou ácidos nucleicos) suficiente para produzir uma partícula que pode se fundir com uma célula-alvo. Nas modalidades, esse ácido nucleico (ou ácidos nucleicos) codifica proteínas com uma ou mais dentre (por exemplo, todas) as seguintes atividades: atividade de poliproteína gag, atividade de polimerase, atividade de integrase, atividade de protease e atividade de fusogênio.

[00247] Os fusossomas também podem compreender várias estruturas que facilitam a entrega de uma carga útil a uma célula-alvo. Por exemplo, em algumas modalidades, o fusossoma (por exemplo, vírus, por exemplo, retrovírus, por exemplo, lentivírus) compreende um ou mais dentre (por exemplo, todas) as seguintes proteínas: poliproteína gag, polimerase (por exemplo, pol), integrase (por exemplo, uma variante funcional ou não funcional), protease e um fusogênio. Em algumas modalidades, o fusossoma compreende adicionalmente rev. Em algumas modalidades, uma ou mais das proteínas supracitadas são codificadas no genoma retroviral e, em algumas modalidades, uma ou

mais das proteínas supracitadas são fornecidas em células trans, por exemplo, por uma célula auxiliar, vírus auxiliar ou plasmídeo auxiliar. Em algumas modalidades, o ácido nucleico do fusossoma (por exemplo, ácido nucleico retroviral) compreende uma ou mais dentre (por exemplo, todas) as seguintes sequências de ácidos nucleicos: LTR 5' (por exemplo, que compreende a U5 e sem um domínio de U3 funcional), elemento de empacotamento Psi (Psi), Promotor do trato polipurina central (cPPT) ligado operacionalmente ao gene de carga útil, gene de carga útil (que compreende opcionalmente um íntron antes da estrutura de leitura aberta), sequência de cauda Poli-A, WPRE e LTR 3' (por exemplo, que compreende a U5 e sem uma U3 funcional). Em algumas modalidades, o ácido nucleico do fusossoma (por exemplo, ácido nucleico retroviral) compreende adicionalmente um ou mais elementos isolantes. Em algumas modalidades, o ácido nucleico do fusossoma (por exemplo, ácido nucleico retroviral) compreende adicionalmente um ou mais locais de reconhecimento de miRNA. Em algumas modalidades, um ou mais dentre os locais de reconhecimento de miRNA estão situados a jusante da sequência da cauda poli-A, por exemplo, entre a sequência da cauda poli-A e o WPRE.

[00248] Em algumas modalidades, um fusossoma fornecido no presente documento é administrado a um indivíduo, por exemplo, um mamífero, por exemplo, um ser humano. Em tais modalidades, o indivíduo pode estar correndo o risco, pode ter um sintoma ou pode ter sido diagnosticado ou identificado como tendo uma doença ou afecção específica (por exemplo, uma doença ou afecção descrita no presente documento). Em uma modalidade, o indivíduo tem uma deficiência genética, como qualquer uma listada na Tabela 5. Em algumas modalidades, o fusossoma contém sequências de ácido nucleico que codificam um agente exógeno para tratar a doença ou afecção, tal como para tratar a deficiência genética.

COMPONENTES LENTIVIRAIS E CÉLULAS AUXILIARES

[00249] Em algumas modalidades, o ácido nucleico retroviral compreende um ou mais dentre (por exemplo, todos): um promotor 5’ (por exemplo, para controlar a expressão de todo o RNA empacotado), uma LTR 5' (por exemplo, que inclui R (sinal de cauda de poliadenilação) e/ou uma U5 que inclui um sinal de ativação do iniciador), um sítio de ligação ao iniciador, um sinal de empacotamento psi, um elemento de RRE para exportação nuclear, um promotor diretamente a montante do transgene para controlar a expressão do transgene, um transgene (ou outro elemento de agente exógeno), um trato de polipurina e uma LTR 3' (por exemplo, que inclui uma U3 mutante, uma R e um U5). Em algumas modalidades, o ácido nucleico retroviral compreende ainda um ou mais dentre um cPPT, um WPRE e/ou um elemento isolante.

[00250] Um retrovírus normalmente se replica por transcrição reversa de seu RNA genômico em uma cópia linear de DNA de dupla fita e, subsequentemente, integra de maneira covalente seu DNA genômico em um genoma hospedeiro. Retrovírus ilustrativos adequados para uso em modalidades particulares incluem, porém sem limitação: vírus da leucemia murina de Moloney (M-MuLV), vírus do sarcoma de murganho Moloney (MoMSV), vírus do sarcoma de murganho Harvey (HaMuSV), vírus do tumor mamário de murganho (MuMTV), vírus da leucemia do macaco gibão (GaLV), vírus da leucemia felina (FLV), spumavírus, vírus da leucemia murina Friend, vírus da célula-tronco murina (MSCV) e vírus do sarcoma de Rous (RSV)) e lentivírus.

[00251] Em algumas modalidades, o retrovírus é um Gammaretrovírus. Em algumas modalidades, o retrovírus é um Epsilonretrovírus. Em algumas modalidades, o retrovírus é um Alpharetrovirus. Em algumas modalidades, o retrovírus é um Betaretrovírus. Em algumas modalidades, o retrovírus é um Deltaretrovírus. Em algumas modalidades, o

retrovírus é um Lentivírus. Em algumas modalidades, o retrovírus é um Spumaretrovírus. Em algumas modalidades, o retrovírus é um retrovírus endógeno.

[00252] Lentivírus ilustrativos incluem, porém sem limitação: HIV (vírus da imunodeficiência humana; incluindo HIV tipo 1 e HIV tipo 2); vírus visna-maedi (VMV); o vírus da artrite-encefalite caprina (CAEV); vírus da anemia infecciosa equina (EIAV); vírus da imunodeficiência felina (FIV); vírus da imunodeficiência bovina (BIV); e vírus da imunodeficiência símia (SIV). Em algumas modalidades, são usadas cadeias principais de vetores com base em HIV (isto é, elementos de sequência de ação cis do HIV).

[00253] Em algumas modalidades, um vetor no presente documento é uma molécula de ácido nucleico com capacidade para transferir ou transportar outra molécula de ácido nucleico. O ácido nucleico transferido está geralmente ligado, por exemplo, inserido na molécula de ácido nucleico do vetor. Um vetor pode incluir sequências que direcionam a replicação autônoma em uma célula ou pode incluir sequências suficientes para permitir a integração no DNA da célula hospedeira. Os vetores úteis incluem, por exemplo, plasmídeos (por exemplo, plasmídeos de DNA ou plasmídeos de RNA), transposons, cosmídeos, cromossomos artificiais bacterianos e vetores virais. Os vetores virais úteis incluem, por exemplo, retrovírus e lentivírus com defeito de replicação.

[00254] Um vetor viral pode compreender, por exemplo, uma molécula de ácido nucleico (por exemplo, um plasmídeo de transferência) que inclui elementos de ácido nucleico derivados de vírus que normalmente facilitam a transferência da molécula de ácido nucleico ou integração no genoma de uma célula ou a uma partícula viral que medeia a transferência de ácido nucleico. As partículas virais incluirão tipicamente vários componentes virais e, algumas vezes, também

componentes da célula hospedeira além do ácido nucleico (ou ácidos nucleicos). Um vetor viral pode compreender, por exemplo, um vírus ou partícula viral com capacidade para transferir um ácido nucleico para uma célula ou para o ácido nucleico transferido (por exemplo, como DNA nu). Os vetores virais e plasmídeos de transferência podem compreender elementos genéticos estruturais e/ou funcionais que são principalmente derivados de um vírus. Um vetor retroviral pode compreender um vetor viral ou plasmídeo que contém elementos genéticos estruturais e funcionais, ou porções dos mesmos, que são principalmente derivados de um retrovírus. Um vetor lentiviral pode compreender um vetor viral ou plasmídeo que contém elementos genéticos estruturais e funcionais, ou porções dos mesmos, incluindo LTRs que são derivadas principalmente de um lentivírus.

[00255] Nas modalidades, um vetor lentiviral (por exemplo, vetor de expressão lentiviral) pode compreender um plasmídeo de transferência lentiviral (por exemplo, como DNA nu) ou uma partícula lentiviral infecciosa. Em relação a elementos como locais de clonagem, promotores, elementos regulatórios, ácidos nucleicos heterólogos etc., deve ser entendido que as sequências desses elementos podem estar presentes na forma de RNA em partículas lentivirais e podem estar presentes na forma de DNA nos plasmídeos de DNA.

[00256] Em alguns vetores descritos no presente documento, pelo menos parte de uma ou mais regiões de codificação de proteína que contribuem ou são essenciais para a replicação podem estar ausentes em comparação ao vírus de tipo selvagem correspondente. Isso torna a replicação do vetor viral deficiente. Em algumas modalidades, o vetor tem capacidade para submeter uma célula-alvo hospedeira à transdução sem divisão e/ou integrar seu genoma em um genoma hospedeiro.

[00257] A estrutura de um genoma de retrovírus de tipo selvagem

compreende frequentemente uma repetição terminal longa (LTR) 5’ e uma LTR 3', entre ou dentro das quais estão localizados um sinal de empacotamento para permitir que o genoma seja empacotado, um sítio de ligação de iniciador, locais de integração para permitir a integração no genoma de uma célula hospedeira e os genes gag, pol e env que codificam os componentes de empacotamento que promovem a montagem de partículas virais. Retrovírus mais complexos têm características adicionais, tais como sequências rev e RRE no HIV, que permitem a exportação eficiente de transcrições de RNA do pró-vírus integrado do núcleo para o citoplasma de uma célula-alvo infectada. No provírus, os genes virais são flanqueados em ambas as extremidades por regiões chamadas de repetições terminais longas (LTRs). As LTRs estão envolvidas na integração e transcrição proviral. LTRs também servem como sequências intensificadoras-promotoras e podem controlar a expressão dos genes virais. A encapsidação dos RNAs retrovirais ocorre em virtude de uma sequência psi localizada na extremidade 5' do genoma viral.

[00258] As próprias LTRs são tipicamente sequências semelhantes (por exemplo, idênticas) que podem ser divididas em três elementos, que são denominados de U3, R e U5. U3 é derivado da sequência única da extremidade 3' do RNA. R é derivado de uma sequência repetida em ambas as extremidades do RNA, e U5 é derivado da sequência única para a extremidade 5' do RNA. Os tamanhos dos três elementos podem variar consideravelmente entre os diferentes retrovírus.

[00259] Para o genoma viral, o sítio de iniciação da transcrição está tipicamente na fronteira entre U3 e R em uma LTR e o sítio de adição (terminação) de poli (A) está na fronteira entre R e U5 na outra LTR. U3 contém a maioria dos elementos de controle da transcrição do provírus que incluem o promotor e múltiplas sequências potenciadoras responsivas às proteínas ativadoras da transcrição celular e, em al-

guns casos, viral. Alguns retrovírus compreendem um ou mais dentre os seguintes genes que codificam para proteínas que estão envolvidas na regulação da expressão gênica: tot, rev, tax e rex.

[00260] Em relação aos próprios genes estruturais gag, pol e env, gag codifica a proteína estrutural interna do vírus. A proteína Gag é processada de maneira proteolítica nas proteínas maduras MA (matriz), CA (capsídeo) e NC (nucleocapsídeo). O gene pol codifica a transcriptase reversa (RT), que contém DNA polimerase, associada a RNase H e integrase (IN), que medeia a replicação do genoma. O gene env codifica a glicoproteína de superfície (SU) e a proteína transmembranar (TM) do vírion que formam um complexo que interage especificamente com proteínas receptoras celulares. Essa interação promove infecção, por exemplo, pela fusão da membrana viral com a membrana celular.

[00261] Em um genoma de vetor retroviral deficiente para replicação, gag, pol e env podem estar ausentes ou podem ser não funcionais. As regiões R em ambas as extremidades do RNA são sequências tipicamente repetidas. U5 e U3 representam sequências únicas nas extremidades 5' e 3' do genoma do RNA, respectivamente.

[00262] Os retrovírus também podem conter genes adicionais que codificam para proteínas diferentes de gag, pol e env. Exemplos de genes adicionais incluem (no HIV), um ou mais dentre vif, vpr, vpx, vpu, tat, rev e nef. EIAV tem (entre outros) o gene adicional S2. As proteínas codificadas por genes adicionais têm várias funções, dentre as quais algumas podem ser uma duplicação de uma função fornecida por uma proteína celular. Em EIAV, por exemplo, a tat atua como um ativador transcricional da LTR viral (Derse e Newbold 1993 Virology 194: 530 a 536; Maury et al. 1994 Virology 200: 632 a 642). Esta se liga a uma estrutura secundária de RNA haste-alça estável denominada de TAR. Rev regula e coordena a expressão de genes virais atra-

vés de elementos de resposta a rev (RRE) (Martarano et al. 1994 J. Virol. 68: 3.102 a 3.111). Os mecanismos de ação dessas duas proteínas são considerados amplamente semelhantes aos mecanismos análogos dos vírus primatas. Além disso, foi identificada uma proteína EIAV, Ttm, que é codificada pelo primeiro éxon de tat unido à sequência de codificação env no início da proteína transmembranar.

[00263] Além da protease, da transcriptase reversa e da integrase, os lentivírus não primatas contêm um quarto produto do gene pol que codifica uma dUTPase. Isso pode exercer uma função na capacidade que esses lentivírus têm para infectar determinados tipos de células que não se dividem ou se dividem lentamente.

[00264] Nas modalidades, um vetor lentiviral recombinante (RLV) é um vetor com informação genética retroviral suficiente para permitir o empacotamento de um genoma de RNA, na presença de componentes de empacotamento, em uma partícula viral com capacidade para infectar uma célula-alvo. A infecção da célula-alvo pode compreender a transcrição reversa e integração no genoma da célula-alvo. O RLV tipicamente porta sequências de codificação não virais que devem ser entregues pelo vetor à célula-alvo. Nas modalidades, um RLV não tem capacidade para replicação independente para produzir partículas retrovirais infecciosas dentro da célula-alvo. Normalmente, o RLV carece de um gene gag-pol e/ou env e/ou outros genes envolvidos na replicação. O vetor pode ser configurado como um vetor de íntron dividido, por exemplo, conforme descrito no pedido de patente PCT WO 99/15683, que é incorporado no presente documento a título de referência em sua totalidade.

[00265] Em algumas modalidades, o vetor lentiviral compreende um genoma viral mínimo, por exemplo, o vetor viral foi manipulado de modo a remover os elementos não essenciais e reter os elementos essenciais a fim de fornecer a funcionalidade necessária para infectar,

submeter à transdução e entregar uma sequência de nucleotídeos de interesse para uma célula-alvo hospedeira, por exemplo, conforme descrito no documento n° WO 98/17815, que é incorporado no presente documento a título de referência em sua totalidade.

[00266] Um genoma lentiviral mínimo pode compreender, por exemplo, (5')R-U5-uma ou mais primeiras sequências de nucleotídeo-U3-R(3'). No entanto, o vetor de plasmídeo usado para produzir o genoma lentiviral dentro de uma célula de origem também pode incluir sequências de controle regulatório da transcrição operacionalmente ligadas ao genoma lentiviral para direcionar a transcrição do genoma em uma célula de origem. Essas sequências regulatórias podem compreender as sequências naturais associadas à sequência retroviral transcrita, por exemplo, a região 5' U3, ou podem compreender um promotor heterólogo, tal como outro promotor viral, por exemplo, o promotor CMV. Alguns genomas lentivirais compreendem sequências adicionais para promover a produção eficiente de vírus. Por exemplo, no caso do HIV, as sequências rev e RRE podem ser incluídas. Alternativamente ou em combinação, a otimização de códons pode ser usada, por exemplo, o gene que codifica o agente exógeno pode ser otimizado por códons, por exemplo, conforme descrito no documento n° WO 01/79518, que é incorporado no presente documento a título de referência em sua totalidade. Também podem ser utilizadas sequências alternativas que desempenham uma função semelhante ou igual à do sistema rev/RRE. Por exemplo, um análogo funcional do sistema rev/RRE é encontrado no vírus do macaco Mason Pfizer. Este é conhecido como CTE e compreende uma sequência do tipo RRE no genoma que interage com um fator na célula infectada. O fator celular pode ser considerado um análogo do rev. Assim, o CTE pode ser usado como uma alternativa ao sistema rev/RRE. Além disso, a proteína Rex do HTLV-I pode substituir funcionalmente a proteína Rev do HIV-I.

Rev e Rex têm efeitos semelhantes ao IRE-BP.

[00267] Em algumas modalidades, um ácido nucleico retroviral (por exemplo, um ácido nucleico lentiviral, por exemplo, um ácido nucleico lentiviral de primata ou não primata) (1) compreende um gene gag submetido à deleção em que a deleção em gag remove um ou mais nucleotídeos a jusante aproximadamente dentre o nucleotídeo 350 ou 354 da sequência de codificação gag; (2) tem um ou mais genes acessórios ausentes do ácido nucleico retroviral; (3) não tem o gene tat, porém inclui a sequência líder entre a extremidade da LTR 5' e o ATG de gag; e (4) combinações de (1), (2) e (3). Em uma modalidade, o vetor lentiviral compreende todas as características (1) e (2) e (3). Essa estratégia é descrita mais detalhadamente no documento n° WO 99/32646, que é incorporado no presente documento a título de referência em sua totalidade.

[00268] Em algumas modalidades, um sistema mínimo de lentivírus de primata não precisa de nenhum dos genes adicionais HIV/SIV vif, vpr, vpx, vpu, tat, rev e nef para produção de vetor ou para transdução de células em divisão e não em divisão. Em algumas modalidades, um sistema de vetor mínimo EIAV não precisa S2 para produção de vetor ou para transdução de células em divisão e não divisão.

[00269] A deleção de genes adicionais pode permitir que os vetores sejam produzidos sem os genes associados à doença em infecções lentivirais (por exemplo, HIV). Em particular, a tat está associada a doenças. Em segundo lugar, a deleção de genes adicionais permite que o vetor empacote mais DNA heterólogo. Em terceiro lugar, genes cuja função é desconhecida, como S2, podem ser omitidos, reduzindo assim o risco de causar efeitos indesejáveis. Exemplos de vetores lentivirais mínimos são divulgados no documento n° WO 99/32646 e no documento n° WO 98/17815.

[00270] Em algumas modalidades, o ácido nucleico retroviral é des-

provido de pelo menos tat e S2 (se for um sistema de vetor EIAV) e possivelmente também vif, vpr, vpx, vpu e nef. Em algumas modalidades, o ácido nucleico retroviral também é desprovido de rev, RRE ou ambos.

[00271] Em algumas modalidades, o ácido nucleico retroviral compreende vpx. O polipeptídeo Vpx se liga ao fator de restrição SAMHD1 e induz a degradação do mesmo, que degrada dNTPs livres no citoplasma. Assim, a concentração de dNTPs livres no citoplasma aumenta à medida que Vpx degrada SAMHD1 e a atividade de transcrição reversa é aumentada, facilitando assim a transcrição reversa do genoma retroviral e a integração no genoma da célula-alvo.

[00272] Diferentes células diferem no uso de códons específicos. Essa tendência de códon corresponde a uma tendência na abundância relativa de tRNAs particulares no tipo de célula. Alterando-se os códons na sequência de modo que sejam adaptados para corresponder à abundância relativa de tRNAs correspondentes, é possível aumentar a expressão. Da mesma forma, é possível diminuir a expressão escolhendo-se deliberadamente códons para os quais os tRNAs correspondentes são conhecidos por serem raros no tipo de célula particular. Assim, um grau adicional de controle translacional está disponível. Uma descrição adicional de otimização de códon é encontrada, por exemplo, no documento n° WO 99/41397, que é incorporado no presente documento a título de referência em sua totalidade.

[00273] Muitos vírus, incluindo HIV e outros lentivírus, usam um grande número de códons raros e, alterando-os para corresponder aos códons comumente usados em mamíferos, pode ser alcançada uma expressão aumentada dos componentes de empacotamento em células produtoras de mamíferos.

[00274] A otimização de códons tem várias outras vantagens. Em virtude de alterações em suas sequências, as sequências de nucleotí-

deos que codificam os componentes de empacotamento podem ter sequências de instabilidade de RNA (INS) reduzidas ou eliminadas das mesmas. Ao mesmo tempo, a sequência de codificação da sequência de aminoácidos para os componentes de embalagem é retida de modo que os componentes virais codificados pelas sequências permaneçam iguais ou pelo menos suficientemente semelhantes para que a função dos componentes de embalagem não seja comprometida. Em algumas modalidades, a otimização de códons também supera a exigências de Rev/RRE para exportação, tornando sequências otimizadas independentes de Rev. Em algumas modalidades, a otimização de códons também reduz a recombinação homóloga entre diferentes construtos dentro do sistema de vetor (por exemplo, entre as regiões de sobreposição nas estruturas de leitura aberta gag-pol e env). Em algumas modalidades, a otimização de códons causa um aumento no título viral e/ou segurança aprimorada.

[00275] Em algumas modalidades, apenas os códons relacionados a INS são otimizados por códons. Em outras modalidades, as sequências são otimizadas por códons em sua totalidade, com exceção da sequência que abrange o sítio de deslocamento de quadro de gag-pol.

[00276] O gene gag-pol compreende dois quadros de leitura sobrepostos que codificam as proteínas gag-pol. A expressão de ambas as proteínas depende de uma mudança de quadro durante a tradução. Essa mudança de quadro ocorre como resultado do “deslizamento” do ribossoma durante a tradução. Acredita-se que esse deslizamento seja causado, pelo menos em parte, por estruturas secundárias de RNA que paralisam o ribossoma. Essas estruturas secundárias existem a jusante do sítio de deslocamento de quadro no gene gag-pol. Para o HIV, a região de sobreposição se estende do nucleotídeo 1222 a jusante do início do gag (em que o nucleotídeo 1 é o A do ATG do gag) até o final do gag (nt 1503). Consequentemente, um fragmento de 281

bp que abrange o sítio de deslocamento de quadro e a região de sobreposição dos dois quadros de leitura é, de preferência, não otimizado por códon. Em algumas modalidades, a retenção desse fragmento permitirá uma expressão mais eficiente das proteínas gag-pol. Para EIAV, o início da sobreposição é no nt 1262 (em que o nucleotídeo 1 é o A do ATG gag). O fim da sobreposição está no nt 1461. A fim de garantir que o sítio de deslocamento de quadro e a sobreposição de gag-pol sejam preservados, a sequência de tipo selvagem pode ser retida de nt 1156 a 1465.

[00277] Podem ser feitas derivações da utilização ideal de códons, por exemplo, a fim de acomodar locais de restrição convenientes, e mudanças conservativas de aminoácidos podem ser introduzidas nas proteínas gag-pol.

[00278] Em algumas modalidades, a otimização de códons se baseia em códons com uso insuficiente de códons em sistemas de mamíferos. A terceira e, às vezes, a segunda e a terceira bases podem ser alteradas.

[00279] Devido à natureza degenerada do código genético, será observado que inúmeras sequências gag-pol podem ser obtidas por um trabalhador especializado. Além disso, há muitas variantes retrovirais descritas que podem ser usadas como um ponto de partida para gerar uma sequência gag-pol otimizada por códons. Os genomas lentivirais podem ser bastante variáveis. Por exemplo, existem muitas quasispecies de HIV-I que ainda são funcionais. Esse também é o caso do EIAV. Essas variantes podem ser usadas para melhorar partes específicas do processo de transdução. Exemplos de variantes do HIV-I podem ser encontrados nos bancos de dados de HIV mantidos pelo Laboratório Nacional de Los Alamos. Detalhes dos clones de EIAV podem ser encontrados no banco de dados NCBI mantido pelo National Institutes of Health.

[00280] A estratégia para sequências gag-pol otimizados por códon pode ser usada em relação a qualquer retrovírus, por exemplo, EIAV, FIV, BIV, CAEV, VMR, SIV, HIV-I e HIV-2. Além disso, esse método pode ser usado para aumentar a expressão de genes de HTLV-I, HTLV-2, HFV, HSRV e retrovírus endógenos humanos (HERV), MLV e outros retrovírus.

[00281] Conforme descrito acima, os componentes de empacotamento para um vetor retroviral podem incluir produtos de expressão dos genes gag, pol e env. Além disso, o empacotamento pode utilizar uma sequência curta de 4 hastes-alças seguidas por uma sequência parcial de gag e env como um sinal de empacotamento. Desse modo, a inclusão de uma sequência gag submetida à deleção no genoma do vetor retroviral (além da sequência gag completa no construto de empacotamento) pode ser usada. Nas modalidades, o vetor retroviral compreende um sinal de empacotamento que compreende de 255 a 360 nucleotídeos de gag em vetores que ainda retêm sequências env ou aproximadamente 40 nucleotídeos de gag em uma combinação particular de mutação de doador de splice, deleções gag e env. Em algumas modalidades, o vetor retroviral inclui uma sequência gag que compreende uma ou mais deleções, por exemplo, a sequência gag compreende aproximadamente 360 nucleotídeos deriváveis da extremidade N.

[00282] O vetor retroviral, célula auxiliar, vírus auxiliar ou plasmídeo auxiliar pode compreender proteínas retrovirais estruturais e acessórias, por exemplo, proteínas gag, pol, env, tat, rev, vif, vpr, vpu, vpx ou nef ou outras proteínas retrovirais. Em algumas modalidades, as proteínas retrovirais são derivadas do mesmo retrovírus. Em algumas modalidades, as proteínas retrovirais são derivadas de mais de um retrovírus, por exemplo, 2, 3, 4 ou mais retrovírus.

[00283] As sequências de codificação gag e pol são geralmente or-

ganizadas como o precursor Gag-Pol no lentivírus nativo. A sequência gag codifica uma proteína precursora de Gag de 55 kD, também denominada de p55. O p55 é clivado pela protease4 codificada por vírus (um produto do gene pol) durante o processo de maturação em quatro proteínas menores designadas como MA (matriz [p17]), CA (capsídeo [p24]), NC (nucleocapsídeo [p9]) e p6. A proteína precursora de pol é removida por clivagem da Gag por uma protease codificada por vírus e posteriormente digerida para separar as atividades de protease (p10), RT (p50), RNase H (p15) e integrase (p31).

[00284] As sequências nativas de Gag-Pol podem ser utilizadas em um vetor auxiliar (por exemplo, plasmídeo auxiliar ou vírus auxiliar) ou podem ser feitas modificações. Essas modificações incluem Gag-Pol quimérico, em que as sequências Gag e Pol são obtidas de diferentes vírus (por exemplo, diferentes espécies, subespécies, cepas, clados, etc.) e/ou em que as sequências foram modificadas para melhorar a transcrição e/ou translação e/ou reduzir a recombinação.

[00285] Em vários exemplos, o ácido nucleico retroviral inclui um polinucleotídeo que codifica uma porção de nucleotídeo 150 a 250 (por exemplo, 168) de uma proteína gag que (i) inclui uma sequência inibidora de INS1 mutada que reduz a restrição de exportação nuclear de RNA em relação ao tipo selvagem INS1, (ii) contém a inserção de dois nucleotídeos que resulta no deslocamento de quadro e no término prematuro e/ou (iii) não inclui as sequências inibidoras de INS2, INS3 e INS4 de gag.

[00286] Em algumas modalidades, um vetor descrito no presente documento é um vetor híbrido que compreende sequências retrovirais (por exemplo, lentivirais) e sequências virais não lentivirais. Em algumas modalidades, um vetor híbrido compreende retrovirais, por exemplo, lentivirais, sequências para transcrição reversa, replicação, integração e/ou empacotamento.

[00287] De acordo com determinadas modalidades específicas, a maioria ou todas as sequências da cadeia principal do vetor viral são derivadas de um lentivírus, por exemplo, HIV-1. No entanto, deve ser entendido que muitas fontes diferentes de sequências retrovirais e/ou lentivirais podem ser usadas ou combinadas, e inúmeras substituições e alterações em algumas das sequências lentivirais podem ser feitas sem prejudicar a capacidade que um vetor de transferência tem para exercer as funções descritas no presente documento. Uma variedade de vetores lentivirais é descrita em Naldini et al., (1996a, 1996b e 1998); Zufferey et al., (1997); Dull et al., 1998, na Patente n° U.S. 6.013.516; e 5.994.136, dentre os quais muitos podem ser adaptados para produzir um ácido nucleico retroviral.

[00288] Em cada extremidade do provírus, repetições terminais longas (LTRs) são normalmente encontradas. Uma LTR compreende tipicamente um domínio localizado nas extremidades do ácido nucleico retroviral que, em seu contexto de sequência natural, são repetições diretas e contêm regiões U3, R e U5. As LTRs geralmente promovem a expressão de genes retrovirais (por exemplo, promoção, iniciação e poliadenilação de transcritos de genes) e replicação viral. A LTR pode compreender vários sinais regulatórios, incluindo elementos de controle da transcrição, sinais de poliadenilação e sequências para replicação e integração do genoma viral. A LTR viral é normalmente dividida em três regiões denominadas de U3, R e U5. A região U3 contém tipicamente os elementos intensificadores e promotores. A região U5 é tipicamente a sequência entre o sítio de ligação do iniciador e a região R e pode conter a sequência de poliadenilação. A região R (repetição) pode ser flanqueada pelas regiões U3 e U5. A LTR é tipicamente composta de regiões U3, R e U5 e pode aparecer nas extremidades 5' e 3' do genoma viral. Em algumas modalidades, adjacentes à LTR 5' estão sequências para a transcrição reversa do genoma (o sítio de li-

gação do iniciador tRNA) e para empacotamento eficiente de RNA viral em partículas (o sítio de Psi).

[00289] Um sinal de empacotamento pode compreender uma sequência localizada dentro do genoma retroviral que medeia a inserção do RNA viral no capsídeo ou partícula viral, consultar, por exemplo, Clever et al., 1995. J. of Virology, Volume 69, n° 4; páginas 2.101 a 2.109. Vários vetores retrovirais usam um sinal de empacotamento mínimo (uma sequência psi [Ψ]) para encapsidação do genoma viral.

[00290] Em várias modalidades, os ácidos nucleicos retrovirais compreendem a LTR 5' e/ou a LTR 3' modificadas. Uma dentre ou ambas as LTRs podem compreender uma ou mais modificações incluindo, porém sem limitação, uma ou mais deleções, inserções ou substituições. As modificações da LTR 3' são frequentemente feitas para melhorar a segurança dos sistemas lentivirais ou retrovirais, tornando os vírus defeituosos na replicação, por exemplo, vírus que não tem capacidade para replicação completa e eficaz, de modo que vírions infecciosos não sejam produzidos (por exemplo, progênie lentiviral de replicação defeituosa).

[00291] Em algumas modalidades, um vetor é um vetor de autoinativação (SIN), por exemplo, vetor de replicação defeituosa, por exemplo, vetor retroviral ou lentiviral, em que a região promotora-intensificadora de LTR direita (3'), conhecida como região U3, foi modificada (por exemplo, por deleção ou substituição) para prevenir a transcrição viral além da primeira rodada de replicação viral. Isso ocorre devido ao fato de que a região direita (3') LTR U3 pode ser usada como um molde para a região esquerda (5') LTR U3 durante a replicação viral e, portanto, a ausência do promotor de intensificação da U3 inibe a replicação viral. Nas modalidades, a LTR 3' é modificada de modo que a região U5 seja removida, alterada ou substituída, por exemplo, por uma sequência exógena de Poli(A). A LTR 3', a LTR 5'

ou tanto a LTRs 3’ quanto a 5’ podem ser LTRs modificadas.

[00292] Em algumas modalidades, a região U3 da LTR 5' é substituída por um promotor heterólogo para conduzir a transcrição do genoma viral durante a produção de partículas virais. Exemplos de promotores heterólogos que podem ser usados incluem, por exemplo, vírus símio viral 40 (SV40) (por exemplo, precoce ou tardio), citomegalovírus (CMV) (por exemplo, início imediato), vírus da leucemia murina Moloney (MoMLV), vírus do sarcoma de Rous (RSV) e promotores do vírus herpes simplex (HSV) (timidina quinase). Em algumas modalidades, os promotores são capazes de conduzir altos níveis de transcrição de uma maneira independente de Tat. Em determinadas modalidades, o promotor heterólogo tem vantagens adicionais em controlar a maneira em que o genoma viral é transcrito. Por exemplo, o promotor heterólogo pode ser induzível, de modo que a transcrição de todo ou de parte do genoma viral ocorra apenas quando os fatores de indução estiverem presentes. Os fatores de indução incluem, porém sem limitação, um ou mais compostos químicos ou as condições fisiológicas, tais como temperatura ou pH, nas quais as células hospedeiras são cultivadas.

[00293] Em algumas modalidades, os vetores virais compreendem um elemento TAR (resposta de transativação), por exemplo, localizado na região R de LTRs lentivirais (por exemplo, HIV). Esse elemento interage com o elemento genético transativador lentiviral (tat) para aumentar a replicação viral. No entanto, esse elemento não é necessário, por exemplo, em modalidades em que a região U3 da LTR 5’ é substituída por um promotor heterólogo.

[00294] A região R, por exemplo, a região dentro de LTRs retrovirais que começam no início do grupo de cobertura (isto é, o início da transcrição) e terminando imediatamente antes do início do trato poli A, pode ser flanqueada pelas regiões U3 e U5. A região R exerce uma

função durante a transcrição reversa na transferência de DNA nascente de uma extremidade do genoma para a outra.

[00295] O ácido nucleico retroviral também pode compreender um elemento FLAP, por exemplo, um ácido nucleico cuja sequência inclui o trato central de polipurina e sequências de terminação central (cPPT e CTS) de um retrovírus, por exemplo, HIV-1 ou HIV-2. Elementos FLAP adequados são descritos na Patente n° U.S. 6.682.907 e no documento Zennou, et al., 2000, Cell, 101: 173, que são incorporados no presente documento a título de referência na sua totalidade. Durante a transcrição reversa do HIV-1, a iniciação central do DNA de fita positiva no trato central de polipurina (cPPT) e a terminação central na sequência de terminação central (CTS) pode causar a formação de uma estrutura de DNA tripla fita: a aba de DNA central do HIV-1. Em algumas modalidades, a estrutura do vetor retroviral ou lentiviral compreende um ou mais elementos FLAP a montante ou a jusante do gene que codifica o agente exógeno. Por exemplo, em algumas modalidades, um plasmídeo de transferência inclui um elemento FLAP, por exemplo, um elemento FLAP derivado ou isolado de HIV-1.

[00296] Nas modalidades, um ácido nucleico retroviral ou lentiviral compreende um ou mais elementos de exportação, por exemplo, um elemento regulatório pós-transcricional de ação cis que regula o transporte de uma transcrição de RNA do núcleo para o citoplasma de uma célula. Exemplos de elementos de exportação de RNA incluem, porém sem limitação, o elemento de resposta rev do vírus da imunodeficiência humana (HIV) (RRE) (consultar, por exemplo, Cullen et al., 1991. J. Virol. 65: 1.053; e Cullen et al., 1991. Cell 58: 423), e o elemento regulatório pós-transcricional do vírus da hepatite B (HPRE), que são incorporados no presente documento a título de referência em sua totalidade. Geralmente, o elemento de exportação de RNA é colocado dentro da UTR 3' de um gene e pode ser inserido como uma ou várias

cópias.

[00297] Em algumas modalidades, a expressão de sequências heterólogas em vetores virais é aumentada pela incorporação de um ou mais dentre, por exemplo, todos os elementos regulatórios pós-transcricionais, locais de poliadenilação e sinais de terminação da transcrição nos vetores. Uma variedade de elementos regulatórios pós-transcricionais pode aumentar a expressão de um ácido nucleico heterólogo na proteína, por exemplo, elemento regulatório pós-transcricional do vírus da hepatite da marmota (WPRE; Zufferey et al., 1999, J. Virol., 73: 2886); o elemento regulatório pós-transcricional presente no vírus da hepatite B (HPRE) (Huang et al., Mol. Cell. Biol., 5: 3864); e semelhantes (Liu et al., 1995, Genes Dev., 9: 1.766), dentre os quais cada um é incorporado no presente documento a título de referência na sua totalidade. Em algumas modalidades, um ácido nucleico retroviral descrito no presente documento compreende um elemento regulatório pós-transcricional, como um WPRE ou HPRE

[00298] Em algumas modalidades, um ácido nucleico retroviral descrito no presente documento carece ou não compreende um elemento regulatório pós-transcricional, como um WPRE ou HPRE.

[00299] Elementos que direcionam a terminação e poliadenilação dos transcritos de ácido nucleico heterólogos podem ser incluídos, por exemplo, para aumentar a expressão do agente exógeno. Os sinais de terminação da transcrição podem ser encontrados a jusante do sinal de poliadenilação. Em algumas modalidades, os vetores compreendem uma sequência de poliadenilação 3' de um polinucleotídeo que codifica o agente exógeno. Um sítio Poli-A pode compreender uma sequência de DNA que direciona tanto a terminação quanto a poliadenilação da transcrição de RNA nascente pela RNA polimerase II. As sequências de poliadenilação podem promover a estabilidade do mRNA pela adição de uma cauda Poli-A à extremidade 3' da sequên-

cia de codificação e, assim, contribuir para o aumento da eficiência da tradução. Exemplos ilustrativos de sinais de Poli-A que podem ser usados em um ácido nucleico retroviral incluem AATAAA, ATTAAA, AGTAAA, uma sequência de Poli-A do hormônio de crescimento bovino (BGHpA), uma sequência de Poli-A de β-globina de coelho (rβgpA) ou outra sequência de Poli-A heteróloga ou endógena adequada.

[00300] Em algumas modalidades, um vetor retroviral ou lentiviral compreende adicionalmente um ou mais elementos isolantes, por exemplo, um elemento isolante descrito no presente documento.

[00301] Em várias modalidades, os vetores compreendem um promotor ligado operacionalmente a um polinucleotídeo que codifica um agente exógeno. Os vetores podem ter uma ou mais LTRs, em que qualquer LTR compreende uma ou mais modificações, tais como uma ou mais substituições, adições ou deleções de nucleotídeos. Os vetores podem compreender adicionalmente um ou mais elementos acessórios para aumentar a eficiência de transdução (por exemplo, um cPPT/FLAP), empacotamento viral (por exemplo, um sinal de empacotamento Psi (Ψ), RRE) e/ou outros elementos que aumentam a expressão do gene exógeno (por exemplo, sequências poli (A)) e pode opcionalmente compreender um WPRE ou HPRE.

[00302] Em algumas modalidades, ácido nucleico lentiviral compreende um ou mais, por exemplo, todas, por exemplo, de 5’ a 3’, dentre um promotor (por exemplo, CMV), uma sequência R (por exemplo, que compreende TAR), uma sequência U5 (por exemplo, para integração), uma sequência PBS (por exemplo, para transcrição reversa), uma sequência DIS (por exemplo, para dimerização de genoma), um sinal de empacotamento psi, uma sequência de gag parcial, uma sequência de RRE (por exemplo, para exportação nuclear), uma sequência cPPT (por exemplo, para importação nuclear), um promotor para conduzir a expressão do agente exógeno, um gene que codifica o agente exóge-

no, uma sequência WPRE (por exemplo, para expressão transgênica eficiente), uma sequência PPT (por exemplo, para transcrição reversa), uma sequência R (por exemplo, para poliadenilação e terminação) e um sinal U5 (por exemplo, para integração).

VETORES PROJETADOS PARA REMOVER LOCAIS DE SPLICE

[00303] Alguns vetores lentivirais se integram dentro de genes ativos e possuem fortes sinais de splicing e poliadenilação que podem levar à formação de transcritos aberrantes e possivelmente truncados.

[00304] Os mecanismos de ativação do proto-oncogene podem envolver a geração de transcritos quiméricos originados da interação de elementos promotores ou locais de splice contidos no genoma do agente mutagênico de inserção com a unidade de transcrição celular direcionada pela integração (Gabriel et al. 2009. Nat Med 15: 1.431 a 1.436; Bokhoven, et al. J Virol 83:283 a 229). Transcritos de fusão quimérica que compreendem sequências de vetor e mRNAs celulares podem ser gerados por transcrição de leitura que começam das sequências de vetor e prosseguem para os genes celulares flanqueadores ou vice-versa.

[00305] Em algumas modalidades, um ácido nucleico lentiviral descrito no presente documento compreende uma cadeia principal lentiviral em que pelo menos dois dos locais de splice foram eliminados, por exemplo, para melhorar o perfil de segurança do vetor lentiviral. As espécies de tais locais de splice e métodos de identificação são descritos no documento n° WO2012156839A2, cuja totalidade é incluída a título de referência.

MÉTODOS DE PRODUÇÃO RETROVIRAL

[00306] A produção de partículas virais em grande escala é frequentemente útil para atingir um título viral desejado. As partículas virais podem ser produzidas por transfecção de um vetor de transferência em uma linhagem celular de empacotamento que compreende ge-

nes estruturais e/ou acessórios virais, por exemplo, genes gag, pol, env, tat, rev, vif, vpr, vpu, vpx ou nef ou outros genes retrovirais.

[00307] Nas modalidades, o vetor de empacotamento é um vetor de expressão ou vetor viral que carece de um sinal de empacotamento e compreende um polinucleotídeo que codifica um, dois, três, quatro ou mais genes estruturais e/ou acessórios virais. Normalmente, os vetores de empacotamento são incluídos em uma célula de empacotamento e são introduzidos na célula por meio de transfecção, transdução ou infecção. Um vetor de transferência retroviral, por exemplo, lentiviral, pode ser introduzido em uma linhagem celular de empacotamento, por meio de transfecção, transdução ou infecção para gerar uma célula ou linhagem celular de origem. Os vetores de empacotamento podem ser introduzidos em células ou linhagens de células humanas por métodos padrão incluindo, por exemplo, transfecção com fosfato de cálcio, lipofecção ou eletroporação. Em algumas modalidades, os vetores de empacotamento são introduzidos nas células juntamente com um marcador selecionável dominante, tal como neomicina, higromicina, puromicina, blastocidina, zeocina, timidina quinase, DHFR, Gln sintetase ou ADA, seguido por seleção na presença do fármaco apropriado e isolamento de clones. Um gene marcador selecionável pode ser ligado fisicamente a genes que codificam pelo vetor de empacotamento, por exemplo, por IRES ou peptídeos virais de autoclivagem.

[00308] As linhagens de células de empacotamento incluem linhagens celulares que não contêm um sinal de empacotamento, porém expressam de forma estável ou transiente as proteínas estruturais virais e enzimas de replicação (por exemplo, gag, pol e env) que podem empacotar partículas virais. Qualquer linhagem celular adequada pode ser usada, por exemplo, células de mamíferos, por exemplo, células humanas. As linhagens celulares que podem ser usadas incluem, por exemplo, células CHO, células BHK, células MDCK, células C3H

10T1/2, células FLY, células Psi-2, células BOSC 23, células PA317, células WEHI, células COS, células BSC 1, células BSC 40, células BMT 10, células VERO, células W138, células MRC5, células A549, células HT1080, células 293, células 293T, células B-50, células 3T3, células NIH3T3, células HepG2, células Saos-2, células Huh7, células HeLa, células W163, células 211 e células 211A. Nas modalidades, as células de empacotamento são células 293, células 293T ou células A549.

[00309] Uma linhagem celular de origem inclui uma linhagem celular que tem capacidade para produzir partículas retrovirais recombinantes, que compreende uma linhagem celular de empacotamento e um construto de vetor de transferência que compreende um sinal de empacotamento. Os métodos de preparação de soluções-mãe virais são ilustrados pelos documentos, por exemplo, Y. Soneoka et al. (1995) *Nucl. Acids Res.* 23:628 a 633, e N. R. Landau et al. (1992) *J. Virol.* 66:5.110 a 5.113, que são incorporados no presente documento a título de referência. Partículas de vírus infecciosas podem ser coletadas das células de empacotamento, por exemplo, por lise celular ou coleta do sobrenadante da cultura de células. Opcionalmente, as partículas de vírus coletadas podem ser enriquecidas ou purificadas.

PLASMÍDEOS DE EMPACOTAMENTO E LINHAGENS CELULARES

[00310] Em algumas modalidades, a célula de origem compreende um ou mais plasmídeos que codificam para proteínas estruturais virais e enzimas de replicação (por exemplo, gag, pol e env) que podem empacotar partículas virais. Em algumas modalidades, as sequências que codificam para pelo menos dois dos precursores gag, pol e env estão no mesmo plasmídeo. Em algumas modalidades, as sequências que codificam para os precursores gag, pol e env estão em plasmídeos diferentes. Em algumas modalidades, as sequências que codificam para os precursores gag, pol e env têm o mesmo sinal de expressão,

por exemplo, promotor. Em algumas modalidades, as sequências que codificam para os precursores gag, pol e env têm um sinal de expressão diferente, por exemplo, promotores diferentes. Em algumas modalidades, a expressão dos precursores gag, pol e env é induzível. Em algumas modalidades, os plasmídeos que codificam para proteínas estruturais virais e enzimas de replicação são submetidos à transecção ao mesmo tempo ou em momentos diferentes. Em algumas modalidades, os plasmídeos que codificam para proteínas estruturais virais e enzimas de replicação são submetidos à transecção ao mesmo tempo ou em um momento diferente do vetor de empacotamento.

[00311] Em algumas modalidades, a linhagem celular de origem compreende um ou mais genes estruturais virais integrados estavelmente. Em algumas modalidades, a expressão dos genes estruturais virais integrados estavelmente é induzível.

[00312] Em algumas modalidades, a expressão dos genes estruturais virais é regulada no nível transcricional. Em algumas modalidades, a expressão dos genes estruturais virais é regulada no nível de tradução. Em algumas modalidades, a expressão dos genes estruturais virais é regulada no nível de pós-tradução.

[00313] Em algumas modalidades, a expressão dos genes estruturais virais é regulada por um sistema dependente de tetraciclina (Tet), em que um repressor transcricional regulado por Tet (Tet-R) se liga a sequências de DNA incluídas em um promotor e reprime a transcrição por impedimento estérico (Yao et al, 1998; Jones et al, 2005). Após a adição de doxiciclina (dox), o Tet-R é liberado, permitindo a transcrição. Vários outros promotores regulatórios da transcrição adequados, fatores de transcrição e indutores de moléculas pequenas são adequados para regular a transcrição de genes estruturais virais.

[00314] Em algumas modalidades, os componentes de lentivírus de terceira geração, vírus da imunodeficiência humana tipo 1 (HIV) Rev,

Gag/Pol e um envelope sob o controle de promotores regulados por Tet e acoplados a cassetes de resistência a antibióticos são integrados separadamente no genoma da célula de origem. Em algumas modalidades, a célula de origem tem apenas uma cópia de cada Rev, Gag/Pol e uma proteína de envelope integrada no genoma.

[00315] Em algumas modalidades, um ácido nucleico que codifica o agente exógeno (por exemplo, um ácido nucleico retroviral que codifica o agente exógeno) também é integrado no genoma da célula de origem. Em algumas modalidades, um ácido nucleico que codifica o agente exógeno é mantido pelo episssoma. Em algumas modalidades, um ácido nucleico que codifica o agente exógeno é submetido à transfecção na célula de origem que integrou de forma estável Rev, Gag/Pol e uma proteína de envelope no genoma. Consultar, por exemplo, Milani et al. *EMBO Molecular Medicine*, 2017, que é incorporado no presente documento incorporado a título de referência em sua totalidade.

[00316] Em algumas modalidades, um ácido nucleico retroviral descrito no presente documento não pode ser submetido à transcrição reversa. Tal ácido nucleico, nas modalidades, pode expressar de maneira transitória um agente exógeno. O retrovírus ou VLP pode compreender uma proteína de transcriptase reversa desativada ou pode não compreender uma proteína de transcriptase reversa. Nas modalidades, o ácido nucleico retroviral compreende um sítio de ligação do iniciador desativado (PBS) e/ou no local. Nas modalidades, um ou mais genes acessórios virais, incluindo rev, tat, vif, nef, vpr, vpu, vpx e S2 ou equivalentes funcionais dos mesmos, estão desativados ou ausentes do ácido nucleico retroviral. Nas modalidades, um ou mais genes acessórios selecionados a partir de S2, rev e tat estão desativados ou ausentes do ácido nucleico retroviral.

ESTRATÉGIAS PARA EMPACOTAR UM ÁCIDO NUCLEICO RETROVIRAL

[00317] Tipicamente, os sistemas de vetores retrovirais modernos consistem em genomas virais portadores de sequências de vetores de ação cis para transcrição, transcrição reversa, integração, tradução e empacotamento de RNA viral nas partículas virais e (2) linhagens celulares produtoras que expressam as sequências de genes retrovirais de ação trans (por exemplo, gag, pol e env) necessárias para a produção de partículas de vírus. Ao separar completamente as sequências do vetor de ação cis e trans, o vírus é incapaz de manter a replicação por mais de um ciclo de infecção. A geração de vírus vivos pode ser evitada por uma série de estratégias, por exemplo, minimizando-se a sobreposição entre as sequências de ação cis e trans para evitar a recombinação.

[00318] Uma partícula de vetor viral que compreende uma sequência que é desprovida ou sem RNA viral pode ser o resultado da remoção ou eliminação do RNA viral da sequência. Em uma modalidade, isso pode ser alcançado com o uso de um sítio de ligação de sinal de empacotamento endógeno na gag. Alternativamente, o sítio de ligação do sinal de empacotamento endógeno está no pol. Nessa modalidade, o RNA que deve ser entregue conterá um sinal de empacotamento cognato. Em outra modalidade, um domínio de ligação heterólogo (que é heterólogo para gag) localizado no RNA a ser entregue, e um sítio de ligação cognato localizado em gag ou pol, podem ser usados para garantir o empacotamento do RNA a ser entregue. A sequência heteróloga pode ser não viral ou pode ser viral, nesse caso pode ser derivada de um vírus diferente. As partículas do vetor podem ser usadas para entregar RNA terapêutico, caso em que a integrase funcional e/ou transcriptase reversa não é necessária. Essas partículas de vetor também podem ser usadas para entregar um gene terapêutico de inte-

resse, nesse caso o pol é tipicamente incluído.

[00319] Em uma modalidade, gag-pol são alterados e o sinal de empacotamento é substituído por um sinal de empacotamento correspondente. Nessa modalidade, a partícula pode empacotar o RNA com o novo sinal de empacotamento. A vantagem dessa abordagem é o fato de que é possível empacotar uma sequência de RNA que é desprovida de sequência viral, por exemplo, RNAi.

[00320] Uma abordagem alternativa é contar com a superexpressão do RNA a ser empacotado. Em uma modalidade, o RNA a ser empacotado é superexpresso na ausência de qualquer RNA que contém um sinal de empacotamento. Isso pode resultar em um nível significativo de RNA terapêutico que é empacotado e nessa quantidade sendo suficiente para submeter uma célula à transdução e ter um efeito biológico.

[00321] Em algumas modalidades, um polinucleotídeo compreende uma sequência de nucleotídeos que codifica uma proteína gag viral ou gag retroviral e proteínas pol, em que a proteína gag ou proteína pol compreende um domínio de ligação de RNA heterólogo com capacidade para reconhecer uma sequência correspondente em uma sequência de RNA para facilitar o empacotamento da sequência de RNA em uma partícula de vetor viral.

[00322] Em algumas modalidades, o domínio de ligação de RNA heterólogo compreende um domínio de ligação de RNA derivado de uma proteína de revestimento de bacteriófago, uma proteína Rev, uma proteína da partícula de ribonucleoproteína nuclear pequena U1, uma proteína Nova, uma proteína TF111A, uma proteína TIS11, uma proteína de atenuação de ligação de RNA trp (TRAP) ou uma pseudouridina sintase.

[00323] Em algumas modalidades, um método compreende no presente documento detectar ou confirmar a ausência de retrovírus com-

petente para replicação. Os métodos podem incluir a avaliação dos níveis de RNA de um ou mais genes alvo, tais como genes virais, por exemplo, genes estruturais ou de empacotamento, a partir dos quais os produtos gênicos são expressos em determinadas células infectadas com um retrovírus competente para replicação, como um gamma-retrovírus ou lentivírus, porém não está presente em um vetor viral usado para submeter células à transdução com um ácido nucleico heterólogo e não, ou cuja presença não seja esperada, presente e/ou expresso em células que não contêm retrovírus competente para replicação. Retrovírus competente para replicação pode ser determinado como estando presente caso os níveis de RNA de um ou mais genes alvo sejam maiores que um valor de referência que pode ser medido direta ou indiretamente, por exemplo, de uma amostra de controle positivo que contém o gene-alvo. Para divulgações adicionais, consultar o documento n° WO2018023094A1.

REPRESSÃO DE UM GENE QUE CODIFICA UM AGENTE EXÓGENO EM UMA CÉLULA DE ORIGEM

[00324] A proteína (super)expressa na célula de origem pode ter um efeito indireto ou direto na montagem e/ou na infectividade do vírion do vetor. A incorporação do agente exógeno em vírions de vetor também pode impactar o processamento a jusante de partículas de vetor.

[00325] Em algumas modalidades, um promotor específico de tecido é usado para limitar a expressão do agente exógeno nas células de origem. Em algumas modalidades, um sistema de controle de tradução heterólogo é usado em culturas de células eucarióticas para reprimir a tradução do agente exógeno em células de origem. Mais especificamente, o ácido nucleico retroviral pode compreender um sítio de ligação ligado operacionalmente ao gene que codifica o agente exógeno, em que o sítio de ligação tem capacidade para interagir com uma pro-

teína de ligação a RNA de modo que a tradução do agente exógeno seja reprimida ou evitada na célula de origem.

[00326] Em algumas modalidades, a proteína de ligação a RNA é a proteína de atenuação de ligação a RNA de triptofano (TRAP), por exemplo, proteína de atenuação de ligação a RNA de triptofano bacteriano. O uso de uma proteína de ligação a RNA (por exemplo, a proteína reguladora de operon de trp bacteriano, proteína de atenuação de ligação a RNA de triptofano, TRAP), e alvos de RNA aos quais a mesma se liga, reprimirá ou impedirá a tradução de transgene dentro de uma célula de origem. Esse sistema é denominado de sistema de células de produção de vetores de repressão no transgene ou sistema TRIP.

[00327] Nas modalidades, a colocação de um sítio de ligação para uma proteína de ligação de RNA (por exemplo, uma sequência de ligação TRAP, tbs) a montante do códon de iniciação da tradução NOI permite a repressão específica da tradução de mRNA derivado do cassete de expressão interno, embora não tenha nenhum efeito danoso na produção ou estabilidade do RNA vetor. O número de nucleotídeos entre a tbs e o códon de iniciação da tradução do gene que codifica o agente exógeno pode variar de 0 a 12 nucleotídeos. A tbs pode ser colocada a jusante de um sítio de entrada do ribossoma interno (IRES) para reprimir a tradução do gene que codifica o agente exógeno em um mRNA multicistrônico.

SISTEMAS DE INTERRUPTOR E AMPLIFICAÇÃO

[00328] Em algumas modalidades, um polinucleotídeo ou célula que abriga o gene que codifica o agente exógeno utiliza um gene suicida, por exemplo, um gene suicida induzível, para reduzir o risco de toxicidade direta e/ou proliferação descontrolada. Em aspectos específicos, o gene suicida não é imunogênico para a célula hospedeira que abriga o agente exógeno. Exemplos de genes suicidas incluem caspase-9,

caspase-8 ou citosina desaminase. A caspase-9 pode ser ativada com o uso de um indutor químico específico de dimerização (CID).

[00329] Em determinadas modalidades, os vetores compreendem segmentos gênicos que fazem com que as células-alvo, por exemplo, células efetoras imunes, por exemplo, células T, sejam suscetíveis à seleção negativa *in vivo*. Por exemplo, a célula submetida à transdução pode ser eliminada como resultado de uma mudança na condição *in vivo* do indivíduo. O fenótipo selecionável negativo pode resultar da inserção de um gene que confere sensibilidade a um agente administrado, por exemplo, um composto. Genes selecionáveis negativos são conhecidos na técnica e incluem, entre outros, os seguintes: o gene da timidina quinase do vírus Herpes simplex tipo I (TK de HSV-I) (Wigler et al., Cell 11: 223, 1977) que confere sensibilidade ao ganciclovir; o gene celular hipoxantina fosfribosiltransferase (HPRT), o gene celular adenina fosforibosiltransferase (APRT) e citosina desaminase bacteriana, (Mullen et al., Proc. Natl. Acad. Sei. EUA. 89:33 (1992)).

[00330] Em algumas modalidades, células submetidas à transdução, por exemplo, células efetoras imunes, como células T, compreendem um polinucleotídeo que compreende ainda um marcador positivo que permite a seleção de células do fenótipo selecionável negativo *in vitro*. O marcador selecionável positivo pode ser um gene que, ao ser introduzido na célula-alvo, expressa um fenótipo dominante que permite a seleção positiva de células portadoras do gene. Os genes desse tipo incluem, entre outros, o gene da higromicina-B fosfotransferase (hph) que confere resistência à higromicina B, o gene da fosfotransferase do amino glicosídeo (neo ou aph) de Tn5 que codifica a resistência ao antibiótico G418, o gene da di-hidrofolato redutase (DHFR), o gene da adenosina desaminase (ADA) e o gene de resistência a múltiplos fármacos (MDR).

[00331] Em algumas modalidades, o marcador selecionável positivo

e o elemento selecionável negativo estão ligados de modo que a perda do elemento selecionável negativo também seja necessariamente acompanhada pela perda do marcador selecionável positivo. Por exemplo, os marcadores selecionáveis positivos e negativos podem ser fundidos de forma que a perda de um cause obrigatoriamente a perda do outro. Um exemplo de um polinucleotídeo fundido que produz como um produto de expressão um polipeptídeo que confere ambas as características de seleção positiva e negativa desejadas descritas acima é um gene de fusão de higromicina fosfotransferase timidina quinase (HyTK). A expressão desse gene produz um polipeptídeo que confere resistência à higromicina B para seleção positiva *in vitro* e sensibilidade ao ganciclovir para seleção negativa *in vivo*. Consultar Lupton S. D., et al, Mol. and Cell. Biology 1 1:3.374 a 3.378, 1991. Além disso, nas modalidades, os polinucleotídeos que codificam os receptores quiméricos estão em vetores retrovirais que contêm o gene fundido, particularmente aqueles que conferem resistência à higromicina B para seleção positiva *in vitro* e sensibilidade ao ganciclovir para seleção negativa *in vivo*, por exemplo, o vetor retroviral HyTK descrito em Lupton, SD et al. (1991), supra. Consultar, também, as publicações dos documentos PCT U591/08442 e PCT/U594/05601, que descrevem o uso de genes de fusão selecionáveis bifuncionais derivados de marcadores selecionáveis positivos dominantes de fusão com marcadores selecionáveis negativos.

[00332] Marcadores selecionáveis positivos adequados podem ser derivados de genes selecionados do grupo que consiste em hph, nco e gpt e marcadores selecionáveis negativos adequados podem ser derivados de genes selecionados do grupo que consiste em citosina desaminase, HSV-I TK, VZV TK, HPRT, APRT e gpt. Outros marcadores adequados são genes de fusão selecionáveis bifuncionais em que o marcador selecionável positivo é derivado de hph ou neo, e o marca-

dor selecionável negativo é derivado de citosina desaminase ou um gene TK ou marcador selecionável.

ESTRATÉGIAS PARA REGULAR A INTEGRAÇÃO LENTIVIRAL

[00333] São divulgados ácidos nucleicos retrovirais e lentivirais que estão ausentes ou desativados em proteínas/sequências-chave de modo a evitar a integração do genoma retroviral ou lentiviral no genoma da célula-alvo. Por exemplo, ácidos nucleicos virais que não têm cada um dos aminoácidos que constituem o motivo DDE altamente conservado (Engelman and Craigie (1992) *J. Virol.* 66:6.361 a 6.369; Johnson et al. (1986) *Proc. Natl. Acad. Sci. EUA*83:7.648 a 7.652; Khan et al. (1991) *Nucleic Acids Res.* 19:851 a 860) de integrase retroviral possibilitam a produção de ácidos nucleicos retrovirais com integração defeituosa.

[00334] Por exemplo, em algumas modalidades, um ácido nucleico retroviral no presente documento compreende uma integrase lentiviral que compreende uma mutação que faz com que a dita integrase seja incapaz de catalisar a integração do genoma viral em um genoma celular. Em algumas modalidades, as ditas mutações são mutações do tipo I que afetam diretamente a integração, ou mutações do tipo II que desencadeiam defeitos pleiotrópicos que afetam a morfogênese do vírion e/ou a transcrição reversa. Exemplos ilustrativos não limitantes de mutações do tipo I são aquelas mutações que afetam qualquer um dos três resíduos que participam no domínio do núcleo catalítico da integrase: $DX_{39-58}DX_{35}E$ (resíduos D64, D116 e E152 da integrase do HIV-1). Em uma modalidade particular, a mutação que faz com que a dita integrase não tenha capacidade para catalisar a integração do genoma viral em um genoma celular é a substituição de um ou mais resíduos de aminoácidos do motivo DDE do domínio do núcleo catalítico da integrase, de preferência, a substituição do primeiro resíduo aspártico do dito motivo DEE por um resíduo asparagina. Em algumas mo-

dalidades, o vetor retroviral não compreende uma proteína integrase.

[00335] Em algumas modalidades, o retrovírus se integra em unidades de transcrição ativas. Em algumas modalidades, o retrovírus não se integra perto dos locais de início da transcrição, a extremidade 5' dos genes ou os locais de clivagem de DNAse1. Em algumas modalidades, a integração do retrovírus não ativa proto-oncogenes ou genes supressores de tumor inativos. Em algumas modalidades, o retrovírus não é genotóxico. Em algumas modalidades, o lentivírus se integra aos íntrons.

[00336] Em algumas modalidades, o ácido nucleico retroviral se integra ao genoma de uma célula-alvo com um número de cópias específico. O número médio de cópias pode ser determinado a partir de células individuais, uma população de células ou colônias de células individuais. Métodos exemplificativos para determinar o número de cópias incluem reação em cadeia da polimerase (PCR) e citometria de fluxo.

[00337] Em algumas modalidades, o DNA que codifica o agente exógeno é integrado no genoma. Em algumas modalidades, o DNA que codifica o agente exógeno é integrado no episssoma. Em algumas modalidades, a razão entre DNA integrado e DNA episssomal que codifica o agente exógeno é de pelo menos 0,01, 0,1, 0,5, 1,0, 2, 5, 10, 100.

[00338] Em algumas modalidades, o DNA que codifica o agente exógeno é linear. Em algumas modalidades, o DNA que codifica o agente exógeno é circular. Em algumas modalidades, a razão entre cópias lineares e cópias circulares de DNA que codificam o agente exógeno é de pelo menos 0,01, 0,1, 0,5, 1,0, 2, 5, 10, 100.

[00339] Nas modalidades, o DNA que codifica o agente exógeno é circular com 1 LTR. Em algumas modalidades, o DNA que codifica o agente exógeno é circular com 2 LTRs. Em algumas modalidades, a razão entre DNA circular que compreende 1 LTR que codifica o agente exógeno e DNA circular que compreende 2 LTRs que codifica o agen-

te exógeno é de pelo menos 0,1, 0,5, 1,0, 2, 5, 10, 20, 50, 100.

MANUTENÇÃO DE UM VÍRUS EPISSOMAL

[00340] Nos retrovírus com deficiência de integração, produtos derivados de cDNA circular da retrotranscrição (por exemplo, de LTR 1 e de LTR 2) podem se acumular no núcleo da célula sem se integrar ao genoma do hospedeiro (consultar Yáñez-Muñoz RJ et al., Nat. Med. 2006, 12: 348 a 353). Como outro DNA exógeno, esses intermediários podem então se integrar no DNA celular em frequências iguais (por exemplo, 10^3 a 10^5/célula).

[00341] Em algumas modalidades, o ácido nucleico retroviral episssomal não se replica. O DNA do vírus episssomal pode ser modificado para ser mantido em células em replicação através da inclusão de origem eucariótica de replicação e uma região de fixação de estrutura/matriz (S/MAR) para associação à matriz nuclear.

[00342] Desse modo, em algumas modalidades, um ácido nucleico retroviral descrito no presente documento compreende uma origem eucariótica de replicação ou uma variante da mesma. Exemplos de origens eucarióticas de replicação de interesse são a origem de replicação do gene β-globina, conforme descrito por Aladjem et al (Science, 1995, 270: 815 a 819), uma sequência consenso a partir das sequências de replicação autônoma associadas a sequências de alfa-satélite isoladas anteriormente de células CV-1 de macaco e fibroblastos de pele humana, conforme descrito por Price et al Journal of Biological Chemistry, 2003, 278 (22): 19.649 a 19.659, a origem de replicação da região do promotor c-myc humano foi descrita por McWinney e Leffak (McWinney C. e Leffak M., Nucleic Acid Research 1990, 18 (5): 1.233 a 1.242). Nas modalidades, a variante mantém substancialmente a capacidade de iniciar a replicação em eucariotos. A capacidade de uma sequência particular de iniciar a replicação pode ser determinada por qualquer método adequado, por exemplo, o ensaio de replicação

autônoma com base na incorporação de bromodeoxiuridina e mudança de densidade (Araujo F. D. et al., supra; Frappier L. et al., supra).

[00343] Em algumas modalidades, o ácido nucleico retroviral compreende uma região de fixação de arcabouço/matriz (S/MAR) ou uma variante da mesma, por exemplo, um elemento de DNA rico em AT não semelhante a consenso com várias centenas de pares de bases de comprimento que organiza o DNA nuclear do genoma eucariótico em domínios de cromatina, por fixação periódica à estrutura da proteína ou matriz do núcleo da célula. São tipicamente encontradas em regiões não codificantes, como regiões flanqueadoras, regiões de fronteira da cromatina e íntrons. Os exemplos de regiões S/MAR são 1,8 kbp de S/MAR do gene IFN-γ humano (hIFN-γ^{grande}), conforme descrito por Bode et al (Bode J. et al., Science, 1992, 255: 195 a 197), a região mínima de 0,7 Kbp da S/MAR di gene IFN-γ humano (hIFN-γ^{curto}) conforme foi descrito por Ramezani (Ramezani A. et al., Blood 2003, 101: 4.717 a 4724), a região mínima de 0,2 Kbp da S/MAR do gene humano de desidrofolato redutase (hDHFR), conforme foi descrito por Mesner L. D. et al., Proc Natl Acad Sci EUA, 2003, 100: 3.281 a 3.286). Nas modalidades, a variante funcionalmente equivalente da S/MAR é uma sequência selecionada com base no conjunto de seis regras que, juntas ou sozinhas, foram sugeridas para contribuir para a função S/MAR (Kramer et al (1996) *Genomics* 33, 305; Singh et al (1997) *Nucl. Acids Res* 25, 1.419). Essas regras foram incorporadas ao programa de computador MAR-Wiz, disponível gratuitamente em genomecluster.secs.oakland.edu/MAR-Wiz. Nas modalidades, a variante mantém substancialmente as mesmas funções da S/MAR da qual deriva, em particular, a capacidade de se ligar especificamente à matriz nuclear. A pessoa versada na técnica pode determinar se uma variante particular tem capacidade para se ligar especificamente à matriz nuclear, por exemplo, pelos ensaios MAR *in vitro* ou *in vivo* descritos por

Mesner et al. (Mesner L. D. et al, supra). Em algumas modalidades, uma sequência específica é uma variante de uma S/MAR, caso a variante particular mostre propensão para a separação da fita do DNA. Essa propriedade pode ser determinada com o uso de um programa específico com base em métodos da mecânica estatística de equilíbrio. A técnica de análise de desestabilização duplex induzida por estresse (SIDD) "[...] calcula até que ponto o nível imposto de estresse superélico diminui a energia livre necessária para abrir o duplex em cada posição ao longo de uma sequência de DNA. Os resultados são exibidos como um perfil de SIDD, no qual os locais de forte desestabilização aparecem como profundos mínimos [...]" Conforme definido em Bode et al (2005) *J. Mol. Biol.* 358.597. O algoritmo de SIDD e a base matemática (Bi e Benham (2004) Bioinformatics 20, 1.477) e a análise do perfil de SIDD podem ser realizados utilizando o recurso de internet disponível gratuitamente em WebSIDD (www.genomecenter.ucdavis.edu/benham). Por conseguinte, em algumas modalidades, o polinucleotídeo é considerado uma variante da sequência de S/MAR, caso o mesmo mostre um perfil de SIDD semelhante à S/MAR.

FUSOGÊNIOS E PSEUDOTIPAGEM

[00344] Os fusossomas (por exemplo, vetores retrovirais) descritos no presente documento podem compreender um fusogênio, por exemplo, um fusogênio endógeno ou um fusogênio pseudotipado.

[00345] Em algumas modalidades, o fusogênio compreende uma proteína (por exemplo, glicoproteína), lipídio ou molécula pequena. Um fusogênio pode ser, por exemplo, um fusogênio de mamífero ou um fusogênio viral. Em algumas modalidades, o fusogênio é um fusogênio de proteína, por exemplo, uma proteína de mamífero ou um homólogo de uma proteína de mamífero (por exemplo, que tem 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% de identidade ou mais), uma proteína de não mamífero, tal como uma proteína viral ou um

homólogo de uma proteína viral (por exemplo, que tem 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% de identidade ou mais), uma proteína nativa ou um derivado de uma proteína nativa, uma proteína sintética, um fragmento da mesma, uma variante da mesma, uma fusão de proteína que compreende um ou mais dentre os fusogênios ou fragmentos e qualquer combinação dos mesmos. Em algumas modalidades, um fusogênio viral é uma proteína de fusão de membrana viral de Classe I, uma proteína de membrana viral de Classe II, uma proteína de fusão de membrana viral Classe III, uma glicoproteína de membrana viral ou outras proteínas de fusão viral ou um homólogo das mesmas, um fragmento da mesma, uma variante da mesma, ou uma fusão de proteína que compreende um ou mais proteínas ou fragmentos das mesmas.

[00346] Fusogênios, que incluem proteínas de envelope viral (env), determinam geralmente a faixa de células hospedeiras que podem ser infectadas e transformadas por fusossomas. No caso de lentivírus, como HIV-1, HIV-2, SIV, FIV e EIV, as proteínas env nativas incluem gp41 e gp120. Em algumas modalidades, as proteínas env virais expressas por células de origem descritas no presente documento são codificadas em um vetor separado dos genes gag e pol virais, conforme descrito anteriormente.

[00347] Exemplos ilustrativos de genes env derivados de retrovirais que podem ser usados incluem, porém sem limitação: envelopes MLV, envelope 10A1, BAEV, FeLV-B, RD114, SSAV, Ebola, Sendai, FPV (vírus da peste aviária) e envelopes de vírus da gripe. De modo similar, podem ser utilizados envelopes que codificam genes de vírus de RNA (por exemplo, famílias de vírus de RNA de Picornaviridae, Calciviridae, Astroviridae, Togaviridae, Flaviviridae, Coronaviridae, Paramyxoviridae, Rhabdoviridae, Filoviridae, Orthomyxoviridae, Bunyaviridae, Arenaviridae, Reoviridae, Birnaviridae, Retroviridae) bem como

os vírus de DNA (famílias de Hepadnaviridae, Circoviridae, Parvoviridae, Papovaviridae, Adenoviridae, Herpesviridae, Poxyiridae e Iridoviridae). Exemplos representativos incluem FeLV, VEE, HFVW, WDSV, SFV, Raiva, ALV, BIV, BLV, EBV, CAEV, SNV, ChTLV, STLV, MPMV, SMRV, RAV, FuSV, MH2, AEV, AMV, CT10 e EIAV.

[00348] Em algumas modalidades, as proteínas de envelope para exibição em um fusossoma incluem, porém sem limitação, qualquer uma dentre seguintes fontes: Influenza A, tais como H1N1, H1N2, H3N2 e H5N1 (gripe aviária), vírus da Influenza B, Influenza C, vírus da Hepatite A, vírus da Hepatite B, vírus da Hepatite C, vírus da Hepatite D, vírus da Hepatite E, Rotavírus, qualquer vírus dentre grupo de vírus Norwalk, adenovírus entéricos, parvovírus, vírus da febre do Dengue, Monkeypox, Mononegavirales, Lyssavírus, tais como vírus da raiva, vírus de morcego de Lagos, vírus Mokola, vírus Duvenhage, vírus do morcego Europeu 1 & 2 e vírus do morcego australiano, Ephemerovírus, Vesiculovírus, Vírus da Estomatite Vesicular (VSV), Vírus do Herpes, tal como vírus do Herpes Simples tipos 1 e 2, varicela-zóster, citomegalovírus, vírus Epstein-Bar (EBV), vírus do herpes humano (HHV), vírus do herpes humano tipo 6 e 8, vírus da imunodeficiência humana (HIV), papilomavírus, vírus do gama-herpes de murganho, Arenavírus, tais como vírus da febre hemorrágica argentina, vírus da febre hemorrágica boliviana, vírus da febre hemorrágica associada ao Sabiá, vírus da febre hemorrágica venezuelana, vírus da febre Lassa, vírus Machupo, vírus da Coriomeningite linfocitária (LCMV), Bunyaviridiae, tais como Crimean-Congo vírus da febre hemorrágica, Hantavírus, vírus causador da febre hemorrágica com síndrome renal, vírus da febre Rift Valley, Filoviridae (filovírus) incluindo febre hemorrágica do Ebola e febre hemorrágica de Marburg, Flaviviridae incluindo vírus da doença da Floresta de Kaysanur, vírus da febre hemorrágica de Omsk, vírus causador da encefalite transportada por carrapato e

Paramyxoviridae, tais como vírus Hendra e vírus Nipah, variola major e variola minor (catapora), alfavírus, tais como vírus da encefalite equina venezuelana, vírus da encefalite equina do leste, vírus da encefalite equina ocidental, coronavírus associado à SARS (SARS-CoV), Vírus do Nilo do Oeste, quaisquer vírus causadores da encefalite.

[00349] Em algumas modalidades, uma célula de origem descrita no presente documento produz um fusossoma, por exemplo, retrovírus recombinante, por exemplo, lentivírus, pseudotipado com a glicoproteína VSV-G.

[00350] Um fusossoma ou vírus pseudotipado geralmente tem uma modificação em uma ou mais de suas proteínas de envelope, por exemplo, uma proteína de envelope é substituída por uma proteína de envelope de outro vírus. Por exemplo, o HIV pode ser pseudotipado com uma proteína de fusão do rabdovírus, por exemplo, proteínas de envelope do tipo proteína G do vírus da estomatite vesicular (VSV-G), o que permite que o HIV infecte uma faixa maior de células devido ao fato de que as proteínas de envelope de HIV (codificadas pelo gene env) normalmente alvejam o vírus para células apresentadoras de CD4+. Em algumas modalidades, as proteínas do envelope lentiviral são pseudotipadas com VSV-G. Em uma modalidade, as células de origem produzem retrovírus recombinantes, por exemplo, lentivírus, pseudotipados com a glicoproteína do envelope VSV-G.

[00351] Além disso, um fusogênio ou proteína de envelope viral pode ser modificado ou projetado para conter sequências polipeptídicas que permitem que o vetor de transdução alveje e infecte células hospedeiras fora de sua faixa normal ou, mais especificamente, limite a transdução para uma célula ou tipo de tecido. Por exemplo, o fusogênio ou a proteína de envelope podem ser unidos no quadro com sequências de alvejamento, tais como ligantes de receptor, anticorpos (com o uso de uma porção de ligação ao antígeno de um anticorpo ou

uma molécula do tipo de anticorpo recombinante, como um anticorpo de cadeia única) e porções químicas polipeptídicas ou modificações das mesmas (por exemplo, em que um sítio de glicosilação está presente na sequência de alvejamento) que, quando exibido no revestimento do vetor de transdução, facilita a entrega dirigida da partícula de vírion a uma célula-alvo de interesse. Além disso, as proteínas de envelope podem compreender ainda sequências que modulam a função celular. A modulação da função celular com um vetor de transdução pode aumentar ou diminuir a eficiência de transdução para certos tipos de células em uma população mista de células. Por exemplo, as células-tronco podem ser submetidas à transdução mais especificamente com sequências de envelope que contêm ligantes ou parceiros de ligação que se ligam especificamente às células-tronco, em vez de outros tipos de células que são encontrados no sangue ou na medula óssea. Exemplos não limitantes são o fator de células-tronco (SCF) e o ligante Flt-3. Outros exemplos incluem, por exemplo, anticorpos (por exemplo, anticorpos de cadeia única que são específicos para um tipo de célula) e, essencialmente, qualquer antígeno (incluindo receptores) que se liga a tecidos como câncer de pulmão, fígado, pâncreas, coração, endotelial, liso, mama, próstata, epitelial, vascular etc.

FUSOGÊNIOS EXEMPLIFICATIVOS

[00352] Em algumas modalidades, o fusossoma inclui um ou mais fusogênios, por exemplo, para facilitar a fusão do fusossoma a uma membrana, por exemplo, uma membrana celular.

[00353] Em algumas modalidades, o fusossoma compreende um ou mais fusogênios em seu envelope para ter um como alvo uma célula ou tipo de tecido específico. Os fusogênios incluem, sem limitação, fusogênios baseados em proteínas, lipídios e produtos químicos. Em algumas modalidades, o fusossoma inclui um primeiro fusogênio que é um fusogênio de proteína e um segundo fusogênio que é um fusogê-

nio de lipídio ou um fusogênio químico. O fusogênio pode se ligar a um parceiro de ligação a fusogênio em uma superfície de célula-alvo. Em algumas modalidades, o fusossoma que compreende o fusogênio integrará a membrana em uma bicamada de lipídio de uma célula-alvo.

[00354] Em algumas modalidades, um ou mais dentre os fusogênios descritos no presente documento podem ser incluídos no fusossoma.

PROTEÍNA FUSOGÊNICA

[00355] Em algumas modalidades, o fusogênio é um fusogênio de proteína, por exemplo, uma proteína de mamífero ou um homólogo de uma proteína de mamífero (por exemplo, que tem 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% de identidade ou mais), uma proteína de não mamífero, tal como uma proteína viral ou um homólogo de uma proteína viral (por exemplo, que tem 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% de identidade ou mais), uma proteína nativa ou um derivado de uma proteína nativa, uma proteína sintética, um fragmento da mesma, uma variante da mesma, uma fusão de proteína que compreende um ou mais dentre os fusogênios ou fragmentos e qualquer combinação dos mesmos.

[00356] Em algumas modalidades, o fusogênio resulta na mistura entre os lipídios no fusossoma e lipídios na célula-alvo. Em algumas modalidades, o fusogênio resulta na formação de um ou mais poros entre o interior do fusossoma e o citosol da célula-alvo.

PROTEÍNAS DE MAMÍFEROS

[00357] Em algumas modalidades, o fusogênio pode incluir uma proteína de mamífero, consultar a Tabela 1A. Exemplos de fusogênios de mamíferos podem incluir, porém sem limitação, uma proteína da família SNARE, tal como vSNAREs e tSNAREs, uma proteína sincitina, tal como Sincytina-1 (DOI: 10.1128/JVI.76.13.6442–6452.2002), e Sincitina-2, myomaker (biorxiv.org/content/early/2017/04/02/123158, doi.org/10.1101/123158, doi: 10.1096/fj.201600945R, doi:10.1038/ na-

ture12343), myomixer (www.nature.com/nature/journal/v499/n7458/full/nature12343.html, doi:10.1038/nature12343), myomerger (science.sciencemag.org/content/early/2017/04/05/science.aam9361, DOI: 10.1126/science.aam9361), FGFRL1 (fator de crescimento semelhante ao receptor de fibroblastos 1), Minion (doi.org/10.1101/122697), uma isoforma de gliceraldeído-3-fosfato desidrogenase (GAPDH) (por exemplo, conforme divulgado no documento U.S. 6.099.857A), uma proteína de junção de lacuna, tal como conexina 43, conexina 40, conexina 45, conexina 32 ou conexina 37 (por exemplo, conforme divulgado no documento n° U.S. 2007/0224176, Hap2, qualquer proteína com capacidade para induzir a formação de sincício entre células heterólogas (consultar a Tabela 2), qualquer proteína com propriedades de fusogênio (consultar a Tabela 3), um homólogo da mesma, um fragmento da mesma, uma variante da mesma e uma fusão de proteína que compreende uma ou mais proteínas ou fragmentos das mesmas. Em algumas modalidades, o fusogênio é codificado por um elemento retroviral endógeno humano (hERV) encontrado no genoma humano. Fusogênios exemplificativos adicionais são divulgados nos documentos n° U.S. 6.099.857A e n° U.S. 2007/0224176, cujo conteúdo é incorporado no presente documento a título de referência em sua totalidade.

TABELA 1A: EXEMPLOS NÃO LIMITATIVOS DE FUSOGÊNIOS NÃO HUMANOS E HUMANOS.

Classes de fusogênio humano e não humano		
Classe de Fusogênio	**ID da família de proteínas Uniprot**	**n° de sequências**
EFF-AFF	PF14884	191
SNARE	PF05739	5.977
DC-STAMP	PF07782	633
ENV	PF00429	312

TABELA 1B: GENES QUE CODIFICAM PROTEÍNAS COM PROPRIEDADES DE FUSOGÊNIO.

Genes humanos com a anotação de ontologia genética de:	
Formação de sincício por proteínas de fusão da membrana plasmática	
ID	**Símbolo**
A0A024R0I0	DYRK1B
A0A024R1N1	MYH9
A0A024R2D8	CAV3
A0A096LNV2	FER1L5
A0A096LPA8	FER1L5
A0A096LPB1	FER1L5
A0AVI2	FER1L5
A6NI61	TMEM8C (myomaker)
B3KSL7	-
B7ZLI3	FER1L5
H0YD14	MYOF
O43184	ADAM12
O60242	ADGRB3
O60500	NPHS1
O95180	CACNA1H
O95259	KCNH1
P04628	WNT1
P15172	MYOD1
P17655	CAPN2
P29475	NOS1
P35579	MYH9
P56539	CAV3
Q2NNQ7	FER1L5

Q4KMG0	CDON
Q53GL0	PLEKHO1
Q5TCZ1	SH3PXD2A
Q6YHK3	CD109
Q86V25	VASH2
Q99697	PITX2
Q9C0D5	TANC1
Q9H295	DCSTAMP
Q9NZM1	MYOF
Q9Y463	DYRK1B

TABELA 1C: CANDIDATOS A FUSOGÊNIO HUMANO

Classe de Fusogênio	**ID de Gene**
SNARE	O15400
	Q16623
	K7EQB1
	Q86Y82
	E9PN33
	Q96NA8
	H3BT82
	Q9UNK0
	P32856
	Q13190
	O14662
	P61266
	O43752
	O60499
	Q13277

	B7ZBM8
	A0AVG3
	Q12846
DC-STAMP	Q9H295
	Q5T1A1
	Q5T197
	E9PJX3
	Q9BR26
ENV	Q9UQF0
	Q9N2K0
	P60507
	P60608
	B6SEH9
	P60508
	B6SEH8
	P61550
	P60509
	Q9N2J8
Fusão muscular (Myomaker)	H0Y5B2
	H7C1S0
	Q9HCN3
	A6NDV4
	K4DI83
Fusão muscular (Myomixer)	NP_001302423.1
	ACT64390.1
	XP_018884517.1
	XP_017826615.1

	XP_020012665.1
	XP_017402927.1
	XP_019498363.1
	ELW65617.1
	ERE90100.1
	XP_017813001.1
	XP_017733785.1
	XP_017531750.1
	XP_020142594.1
	XP_019649987.1
	XP_019805280.1
	NP_001170939.1
	NP_001170941.1
	XP_019590171.1
	XP_019062106.1
	EPQ04443.1
	EPY76709.1
	XP_017652630.1
	XP_017459263.1
	OBS58441.1
	XP_017459262.1
	XP_017894180.1
	XP_020746447.1
	ELK00259.1
	XP_019312826.1
	XP_017200354.1
	BAH40091.1

HA	P03452
	Q9Q0U6
	P03460
JUNÇÃO DA LACUNA	P36382
	P17302
	P36383
	P08034
	P35212
Outros	FGFRL1
	GAPDH

[00358] Em algumas modalidades, o fusossoma compreende uma proteína de geração de curvatura, por exemplo, Epsin1, dinamina ou uma proteína que compreende um domínio BAR. Consultar, por exemplo, Kozlovet al, CurrOp StrucBio 2015, Zimmerberget al. Nat Rev 2006, Richard et al, Biochem J 2011.

PROTEÍNAS DE NÃO MAMÍFEROS

PROTEÍNAS VIRAIS

[00359] Em algumas modalidades, o fusogênio pode incluir uma proteína de não mamífero, por exemplo, uma proteína viral. Em algumas modalidades, um fusogênio viral é uma proteína de fusão de membrana viral de Classe I, uma proteína de membrana viral de Classe II, uma proteína de fusão de membrana viral Classe III, uma glicoproteína de membrana viral ou outras proteínas de fusão viral ou um homólogo das mesmas, um fragmento da mesma, uma variante da mesma, ou uma fusão de proteína que compreende um ou mais proteínas ou fragmentos das mesmas.

[00360] Em algumas modalidades, as proteínas de fusão de membrana viral de Classe I incluem, porém sem limitação, proteína F de baculovírus, por exemplo, proteínas F do gênero nucleopoliedrovírus

(NPV), por exemplo, proteína F de Spodoptera exigua MNPV (SeM-NPV) e proteína F de Lymantria dispar MNPV (LdMNPV) e proteínas F de paramixovírus.

[00361] Em algumas modalidades, as proteínas de membrana virais de Classe II incluem, porém sem limitação, encefalite E óssea do carrapato (TBEV E), Semliki Forest Virus E1/E2.

[00362] Em algumas modalidades, proteínas de fusão de membrana viral de Classe III incluem, porém sem limitação, rabdovírus G (por exemplo, proteína G fusogênica do Vírus de Estomatite Vesicular (VSV-G)), glicoproteína B do herpesvírus (por exemplo, vírus do Herpes Simples 1 (HSV -1) gB)), glicoproteína B do vírus de Epstein Barr (EBV gB), togotovírus G, baculovírus gp64 (por exemplo, Autographa California Multiple NPV (AcMNPV) gp64) e glicoproteína do vírus da doença de Borna (BDV) (BDV G).

[00363] Exemplos de outros fusogênios virais, por exemplo, glicoproteínas de membrana e proteínas de fusão virais, incluem, porém sem limitação: proteínas sincicias virais, tais como hde iemaglutinina nfluenza (HA) ou mutantes, ou proteínas de fusão das mesmas; proteína de envelope do vírus da imunodeficiência humana tipo 1 (HIV-1 ENV), gp120 de LFA-1 de ligação de HIV para formar sincício de linfócitos, gp41 de HIV, gp160 de HIV ou Transativador de Transcrição (TAT) de HIV; glicoproteína viral VSV-G, glicoproteína viral do vírus da estomatite vesicular da família Rhabdoviridae; glicoproteínas gB e gH-gL do vírus varicela-zoster (VZV); vírus da leucemia murina (MLV)-10A1; Glicoproteína do vírus da leucemia do macaco gibão (GaLV); glicoproteínas tipo G em Raiva, Mokola, vírus da estomatite vesicular e Togavírus; proteína de projeção de superfície JHM do vírus da hepatite murina; glicoproteínas de membrana e pico de coronavírus respiratório suíno; glicoproteína em pico de bronquite infecciosa aviária e seu precursor; proteína em pico de coronavírus entérico bovino; os genes F e

H, HN ou G do vírus do sarampo; vírus da cinomose canina, vírus da doença de Newcastle, vírus da parainfluenza humana 3, vírus símio 41, vírus Sendai e vírus sincicial respiratório humano; gH de herpesvírus humano 1 e vírus da varicela de símio, com a proteína chaperona gL; herpesvírus gB humano, bovino e cercopiticina; glicoproteínas de envelope do vírus da leucemia murina Friend e vírus do macaco Mason Pfizer; hemaglutinina neuraminidase do vírus da caxumba e glicoproteínas F1 e F2; glicoproteínas de membrana da encefalomielite equina venezuelana; proteína F de paramixovírus; Proteína gp160 de SIV; Proteína G do vírus Ebola; ou proteína de fusão do vírus Sendai, ou um homólogo da mesma, um fragmento da mesma, uma variante da mesma, e uma proteína de fusão que compreende uma ou mais proteínas ou fragmentos das mesmas.

[00364] Fusogênios de não mamíferos incluem fusogênios virais, homólogos dos mesmos, fragmentos dos mesmos e proteínas de fusão que compreendem uma ou mais proteínas ou fragmentos das mesmas. Os fusogênios virais incluem fusogênios de classe I, fusogênios de classe II, fusogênios de classe III e fusogênios de classe IV. Nas modalidades, os fusogênios de classe I, como gp41 do vírus da imunodeficiência humana (HIV), têm uma conformação pós-fusão característica com um trímero de assinatura de grampos α-helicoidais com uma estrutura de bobina espiral central. As proteínas de fusão viral de classe I incluem proteínas com um feixe central de seis hélices pós-fusão. Proteínas de fusão virais de classe I incluem influenza HA, parainfluenza F, HIV Env, Ebola GP, hemaglutininas de ortomixovírus, proteínas F de paramixovírus (por exemplo, sarampo, (Katoh et al. BMC Biotechnology 2010, 10:37)), Proteínas ENV de retrovírus e fusogênios de filovírus e coronavírus. Nas modalidades, os fusogênios virais de classe II, tais como a glicoproteína E da dengue, têm uma assinatura estrutural de folhas β formando um ectodomínio alongado

que se renova para resultar em um trímero de grampos. Nas modalidades, o fusogênio viral de classe II não possui a bobina espiral central. O fusogênio viral de classe II pode ser encontrado em alfavírus (por exemplo, proteína E1) e flavivírus (por exemplo, glicoproteínas E). Os fusogênios virais de Classe II incluem fusogênios do vírus Semliki Forest, Sinbis, vírus da rubéola e vírus da dengue. Nas modalidades, os fusogênios virais de classe III, tais como a glicoproteína G do vírus da estomatite vesicular, combinam assinaturas estruturais encontradas nas classes I e II. Nas modalidades, um fusogênio viral de classe III compreende hélices α (por exemplo, formando um feixe de seis hélices para dobrar a proteína como com os fusogênios virais de classe I) e folhas β com um peptídeo de fusão anfifílico em sua extremidade, uma reminiscência do fusogênios virais de classe II. Fusogênios virais de classe III podem ser encontrados em rabdovírus e herpesvírus. Nas modalidades, os fusogênios virais de classe IV são proteínas transmembranares pequenas (FAST) associadas à fusão (doi:10.1038/sj.emboj.7600767, Nesbitt, Rae L., “Targeted Intracellular Therapeutic Delivery Using Liposomes Formulated with Multifunctional FAST proteins” (2012). Electronic Thesis and Dissertation Repository. Artigo 388), que são codificados por reovírus não envelopados. Nas modalidades, os fusogênios virais de classe IV são suficientemente pequenos para não formarem grampos (doi: 10.1146/annurev-cellbio-101512-122422, doi:10.1016/j.devcel.2007.12.008).

[00365] Proteína fusogênica ou proteínas do envelope viral podem ser realvejadas por mutação de resíduos de aminoácidos em uma proteína de fusão ou uma proteína de alvejamento (por exemplo, a proteína hemaglutinina). Em algumas modalidades, o fusogênio é sofre mutação aleatória. Em algumas modalidades, o fusogênio sofre mutação racional. Em algumas modalidades, o fusogênio é submetido à evolução direcionada. Em algumas modalidades, o fusogênio é truncado e

apenas um subconjunto do peptídeo é usado no vetor retroviral ou VLP. Por exemplo, resíduos de aminoácidos na proteína hemaglutinina do sarampo podem sofrer mutação para alterar as propriedades de ligação da proteína, redirecionando a fusão (doi: 10.1038/nbt942, Molecular Therapy vol. 16 no. 8, 1.427 a 1.436, agosto de 2008, doi: 10.1038/nbt1060, DOI: 10.1128/JVI.76.7.3558–3563.2002, DOI: 10.1128/JVI.75.17.8016–8020.2001, doi: 10.1073pnas.0604993103).

[00366] Em algumas modalidades, a proteína fusogênica ou proteína do envelope viral é realvejada por meio de i) mutar o aminoácido que reside na sequência da proteína fusogênica natural ou na sequência da proteína do envelope viral e/ou ii) submeter à engenharia genética a proteína fusogênica ou proteína do envelope viral para conter sequências polipeptídicas que permitem que o fusogênio ou a proteína do envelope viral alvejem e fusionem ou infectem as células hospedeiras fora de seu alcance normal.

[00367] Em algumas modalidades, os fusossomas compreendem um ou mais fusogênios em sua superfície externa (por exemplo, integrados na membrana celular) para direcionar uma célula ou tipo de tecido específico. Os fusogênios incluem, sem limitação, fusogênios baseados em proteínas, lipídios e produtos químicos. O fusogênio pode se ligar a um parceiro em uma superfície de células-alvo. Em algumas modalidades, o fusossoma que compreende o fusogênio integrará a membrana em uma bicamada de lipídio de uma célula-alvo.

[00368] Em algumas modalidades, o fusogênio é um fusogênio de paramixovírus. Em algumas modalidades, o fusogênio é uma proteína F do vírus Nipah, uma proteína F do vírus do sarampo, uma proteína F do paramixovírus de tupaia, uma proteína F do paramixovírus, uma proteína F do vírus Hendra, uma proteína F do Henipavírus, uma proteína F do Morbilivírus, uma proteína F do respirovírus, uma proteína F do vírus Sendai, uma proteína F do rubulavírus ou uma proteína F do

avulavírus.

[00369] Em algumas modalidades, o fusogênio é um fusogênio de poxviridae.

[00370] Os fusogênios exemplificativos adicionais são divulgados no documento n° U.S. 9.695.446. U.S. 2004/0028687, U.S. 6.416.997, U.S. 7.329.807, U.S. 2017/0112773, U.S. 2009/0202622. WO 2006/027202 e U.S. 2004/0009604 cujos conteúdos são incorporados a título de referência em suas totalidades.

[00371] Em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos da Tabela 1D ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência à mesma ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% identidade de sequência a uma porção da sequência, por exemplo, uma porção de 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento. Por exemplo, em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos que tem pelo menos 80% de identidade a quaisquer sequências de aminoácidos da Tabela 1D. Em algumas modalidades, uma sequência de ácidos nucleicos descrita no presente documento codifica uma sequência de aminoácidos da Tabela 1D ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98%, ou 99% identidade de sequência à mesma, ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% identidade de sequência em relação a uma porção da sequência, por exemplo, uma porção de 40, 50, 60, 80, 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento.

[00372] Em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos estabe-

lecida em qualquer uma das SEQ ID NOS: 1 a 56, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com a mesma, ou uma sequência de aminoácidos tendo pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com uma porção da sequência, por exemplo, uma porção de 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento. Por exemplo, em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos com pelo menos 80% de identidade com relação a uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 1 a 56. Em algumas modalidades, uma sequência de ácidos nucleicos descrita no presente documento codifica uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 1 a 56, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com a mesma, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com uma porção da sequência, por exemplo, uma porção de 40, 50, 60, 80, 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento.

TABELA 1D. Agrupamentos de sequência f de paramixovírus. Coluna 1, o id do genbank inclui o id de genbank de toda a sequência do genoma do vírus que é a sequência centroide do agrupamento. Coluna 2, os nucleotídeos de cds fornecem os nucleotídeos correspondentes ao cds do gene em todo o genoma. Coluna 3, nome completo do gene, fornece o nome completo do gene, incluindo id de genbank, espécie de vírus, cepa e nome da proteína. Coluna 4, sequência, fornece a sequência de aminoácidos do gene. A coluna 5, nº de sequências/agrupamento, fornece o número de sequências que se agrupam com essa sequência centroide.

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
KP317927	5630-7399	gb:KP317927:5630-7399|Organismo:vírus sincitial respiratório humano|Nome de cepa:Kilifi_9465_7_RSVB_2011|Nome da Proteína:glicoproteína de fusão|Símbolo do Gene:F	MIPQARTELNLGQITMELLIHRSSAIFLTLAINALYLTSSQNITEEF YQSTCSAVSRGYLSALRTGWYTSVITIELSNIKETKCNGTDTKVK LIKQELDKYKNAVTELQLLMQNTPAANNRARREAPQYMNYTINT TGSLNVSISKKRKRRFLGFLLGVGSAIASGIAVSKVLHLEGEVNK IKNALLSTNKAVVSLSNGVSVLTSKVLDLKNYINNQLLPIVNQQS CRISNIETVIEFQQKNSRLLEITREFSVNAGVTTPLSTYMLTNSEL LSLINDMPITNDQKKLMSSNVQIVRQQSYSIMSIIKEEVLAYVVQL PIYGVIDTPCWKLHTSPLCTTNIKEGSNICLTRTDRGWYCDNAG SVSFFPQADTCKVQSNRVFCDTMNSLTLPSEVSLCNTDIFNSKY DCKIMTSKTDISSSVITSLGAIVSCYGKTKCTASNKNRGIIKTFSN GCDYVSNKGVDTVSVGNTLYYVNKLEGKNLYVKGEPIINYYDPL VFPSDEFDASISQVNEKINQSLAFIRRSDELLHNVNTGKSTTNIMI TAIIIVIIVVLLSLIAIGLLLYCKAKNTPVTLSKDQLSGINNIAFSK	993	1
AB524405	4556-6217	gb:AB524405:4556-6217|Organismo:círus da doença de Newcastle|Nome da Cepa:Goose/Alaska/415/91|Nome da Proteína:proteína de fusão|Símbolo do Gene:F	MDPKPSTSYLHAFPLIFVAISLVFMAGRASALDGRPLAAAGIVVT GDKAVNIYTSSQTGTIIIKLLPNMPKDKEQCAKSPLDAYNRTLTT LLAPLGDSIRRIQESVTTSGGERQERLVGAIIGGVALGVATAAQI TAASALIQANQNAANILKLKESIAATNEAVHEVTSGLSQLAVAVG KMQQFVNDQFNKTAQEIDCIKITQQVGVELNLYLTELTTVFGPQI TSPALTQLTIQALYNLAGGNMDYMLTKLGVGNNQLSSLISSGLIS GNPILYDSQTQLLGIQVTLPSVGNLNNMRATYLETLSVSTNKGF ASALVPKVVTQVGSVIEELDTSYCIETDLDLYCTRIVTFPMSPGIF SCLGGNTSACMYSKTEGALTTPYMTLKGSVIANCKMTTCRCAD PPGIISQNYGEAVSLIDKKVCNILTLDGITLRLSGEFDATYQKNISI QDSQVVITGNLDISTELGNVNNSISNALDKLEESNSKLDKVNVRL TSTSALITYIVLTTIALICGIVSLVLACYIMYKQKAQQKTLLWLGNN TLDQMRATTKM	418	2

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AF266286	4875-7247	gb:AF266286:4875-7247|Organismo:cepa do vírus do Sarampo AIK-C|Nome da Cepa:Cepa do vírus do Sarampo (vacina AIK-C)|Nome da Proteína:proteína de fusão|Símbolo do Gene:F	MSIMGLKVNVSAIFMAVLLTLQTPTGQIHWGNLSKIGVVGIGSAS YKVMTRSSHQSLVIKLMPNITLLNNCTRVEIAEYRRLLRTVLEPIR DALNAMTQNIRPVQSVASSRRHKRFAGVVLAGAALGVATAAQI TAGIALHQSMLNSQAIDNLRASLETTNQAIEAIRQAGQEMILAVQ GVQDYINNELIPSMNQLSCDLIGQKLGLKLLRYYTEILSLFGPSL RDPISAEISIQALSYALGGDINKVLEKLGYSGGDLLGILESRGIKA RITHVDTESYFIVLSIAYPTLSEIKGVIVHRLEGVSYNIGSQEWYT TVPKYVATQGYLISNFDESSCTFMPEGTVCSQNALYPMSPLLQ ECLRGYTKSCARTLVSGSFGNRFILSQGNLIANCASILCKCYTTG TIINQDPDKILTYIAADNCPVVEVNGVTIQVGSRRYPDAVYLHRID LGPPILLERLDVGTNLGNAIAKLEDAKELLESSDQILRSMKGLSS TCIVYILIAVCLGGLIGIPALICCCRGRCNKKGEQVGMSRPGLKP DLTGTSKSYVRSL	128	3
AB503857	3068-4687	gb:AB503857: 3068-4687|Organismo: Metapneumovírus Humano|Nome da Cepa: Jpn03-1|Nome da proteína: Precursor da Glicoproteína de fusão|Símbolo do Gene: F	MSWKVVIIFSLLITPQHGLKESYLEESCSTITEGYLSVLRTGWYT NVFTLEVGDVENLTCSDGPSLIKTELDLTKSALRELKTVSADQLA REEQIEKPRQSRFVLGAIALGVATAAAVTAGVAIAKTIRLESEVT AIKNALKTTNEAVSTLGNGVRVLATAVRELKDFVSKNLTRAINKN KCDIDDLKMAVSFSQFNRRFLNVVRQFSDNAGITPAISLDLMTD AELARAVSNMPTSAGQIKLMLENRAMVRRKGFGILIGVYGSSVI YMVQLPIFGVIDTPCWIVKAAPSCSEKKGNYACLLREDQGWYC QNAGSTVYYPNEKDCETRGDHVFCDTAAGINVAEQSKECNINIS TTNYPCKVSTGRHPISMVALSPLGALVACYKGVSCSIGSNRVGII KQLNKGCSYITNQDADTVTIDNTVYQLSKVEGEQHVIKGRPVSS SFDPIKFPEDQFNVALDQVFENIENSQALVDQSNRILSSAEKGN TGFIIVIILIAVLGSSMILVSIFIIIKKTKKPTGAPPELSGVTNNGFIPH S	125	4

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
EU277658	5078-6700	gb: EU277658: 5078-6700|Organismo:Vírus da parainfluenza bovina 3|Nome da cepa:Q5592|Nome da proteína:proteína de fusão|Símbolo do gene:F	MIIIVITMILSLTPSSLCQIDITKLQSVGVLVNSPKGIKISQNFETRY LILSLIPKIEDSHSCGNQQIDQYKKLLDRLIIPLYDGLKLQKDVIVV NHESHNNTNLRTKRFFGEIIGTIAIGIATSAQITAAVALVEAKQAR SDIDKLKEAIKDTNKAVQSIQSSVGNLIVAVKSVQDYVNNEIVPSI TRLGCEAAGLQLGIALTQHYSELTNIFGDNIGTLGEKGVKLQGIA SLYRTNITEVFTTSTVDQYDIYDLLFTESIKMRVIDVDLSDYSITLQ VRLPLLTKVSNTQIYKVDSISYNIQGKEWYIPLPHHIMTKGAFLG GADIKECIESFSNYICPSDPGFILNHEMENCLSGNITQCPKTIVTS DIVPRYAFVDGGVIANCIPTTCTCNGIDNRINQSPDQGIKIITYKE CQIVGINGMLFKTNQEGTLAKYTFDNIKLNNSVALNPIDISLELNK AKSDLEESKRWIEKSNQKLDSIGSWHQSSVTIIIIIVMIVVLLIINAIII MIMIRYLRDRNRHLNNKDSEPYVLTNRQ	93	5
AB040874	4546-6162	gb:AB040874:4546-6162|Organismo:vírus da caxumba|Nome da Cepa:Miyahara|Nome da Proteína:proteína de fusão|Símbolo do Gene:F	MKVFLVTCLGFAVFSSSVCVNINILQQIGYIKQQVRQLSYYSQSS SSYIVVKLLPNIQPTDNSCEFKSVTQYNKTLSNLLLPIAENINNIAS PSSGSRRHKRFAGIAIGIAALGVATAAQVTAAVSLVQAQTNARAI AAMKNSIQATNRAVFEVKEGTQRLAIAVQAIQDHINTIMNTQLNN MSCQILDNQLATSLGLYLTELTTVFQPQLINPALSPISIQALRSLL GSMTPAVVQATLSTSISAAEILSAGLMEGQIVSVLLDEMQMIVKI NIPTIVTQSNALVIDFYSISSFINNQESIIQLPDRILEIGNEQWSYPA KNCKLTRHHIFCQYNEAERLSLESKLCLAGNISACVFSPIAGSYM RRFVALDGTIVANCRSLTCLCKSPSYPIYQPDHHAVTTIDLTACQ TLSLDGLDFSIVSLSNITYAENLTISLSQTINTQPIDISTELSKVNAS LQNAVKYIKESNHQLQSVNVNSKIGAIIVAALVLSILSIIISLLFCCW AYVATKEIRRINFKTNHINTISSSVDDLIRY	89	6

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AB475097	4908-6923	gb:AB475097:4908-6923|Organismo:Canine vírus da cinomose|Nome da Cepa:M25CR|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MNPHEQTIPMHEKIPKRSKTQTHTQQDLPQQHSTKSAESKTSRARHSITSAQRSTHYDPRTADWPDYYIMKRTRSCKQASYRSDNIPAHGDHDGIIHHTPESVSQGAKSRLKMGQSNAVKSGSQCTWLVLWCIGVASLFLCSKAQIHWNNLSTIGIIGTDSVHYKIMTRPSHQYLVIKLMPNVSLIDNCTKAELDEYEKLLSSILEPINQALTLMTKNVKPLQSVGSGRRQRRFAGVVLAGAALGVATAAQITAGIALHQSNLNAQAIQSLRTSLEQSNKAIEEIREATQETVIAVQGVQDYVNNELVPAMQHMSCELVGQRLGLKLLRYYTELLSIFGPSLRDPISAEISIQALSYALGGEIHKILEKLGYSGNDMIAILESRGIKTKITHVDLPGKFIILSVSYPTLSEVKGVIVHRLEAVSYNIGSQEWYTTVPRYVATNGYLISNFDESSCVFVSESAICSQNSLYPMSPLLQQCIRGDTSSCARTLVSGTMGNKFILSKGNIVANCASILCKCYSTSTIINQSPDKLLTFIASDTCPLVEIDGVTIQVGSRQYPDMVYESKVALGPAISLERLDVGTNLGNALKKLDDAKVLIDSSNQILETVRRSSFNFGSLLSVPILSCTALALLLLICCCKRRYQQTHKQNTKVDPTFKPDLTGTSRSYVRSL	46	7
AJ849636	5526-7166	gb:AJ849636:5526-7166|Organismo:vírus da Peste des petits ruminants|Nome da Cepa:Turkey 2000|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MTRVAILTFLFLFPNAVACQIHWGNLSKIGIVGTGSASYKVMTRPSHQTLVIKLMPNITAIDNCTKSEIAEYKRLLITVLKPVEDALSVITKNVRPIQTLTPGRRTRRFAGAVLAGVALGVATAAQITAGVALHQSLMNSQAIESLKTSLEKSNQAIEEIRLANKETILAVQGVQDYINNELVPSVHRMSCELVGHKLGLKLLRYYTEILSIFGPSLRDPIAAEISIQALSYALGGDINRILDKLGYSGGDFLAILESKGIKARVTYVDTRDYFIILSIAYPTLSEIKGVIVHKIEAITYNIGAQEWYTTIPKYVATQGYLISNFDETSCVFTPDGTVCSQNALYPMSPLLQECFQGSTKSCARTLVSGTISNRFILSKGNLIANCASVLCKCYTTETVISQDPDKLLTVVASDKCPVVEVDGVTIQVGSREYPDSVYLHKIDLGPAISLEKLDVGTNLGNAVTRLENAKELLDASDQILKTVKGVPFGGNMYIALAACIGVSLGLVTLICCCKGRCKNKEVPISKINPGLKPDLTGTSKSYVRSL	34	8

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AF017149	6618-8258	gb:AF017149|Organismo:vírus Hendra|Nome da Cepa:AF017149 DESCONHECIDA|Nome da Proteína:fusão|Símbolo do Gene:F	MATQEVRLKCLLCGIIVLVLSLEGLGILHYEKLSKIGLVKGITRKY KIKSNPLTKDIVIKMIPNVSNVSKCTGTVMENYKSRLTGILSPIKG AIELYNNNTHDLVGDVKLAGVVMAGIAIGIATAAQITAGVALYEA MKNADNINKLKSSIESTNEAVVKLQETAEKTVYVLTALQDYINTN LVPTIDQISCKQTELALDLALSKYLSDLLFVFGPNLQDPVSNSMTI QAISQAFGGNYETLLRTLGYATEDFDDLLESDSIAGQIVYVDLSS YYIIVRVYFPILTEIQQAYVQELLPVSFNNDNSEWISIVPNFVLIRN TLISNIEVKYCLITKKSVICNQDYATPMTASVRECLTGSTDKCPR ELVVSSHVPRFALSGGVLFANCISVTCQCQTTGRAISQSGEQTL LMIDNTTCTTVVLGNIIISLGKYLGSINYNSESIAVGPPVYTDKVDI SSQISSMNQSLQQSKDYIKEAQKILDTVNPSLISMLSMIILYVLSIA ALCIGLITFISFVIVEKKRGNYSRLDDRQVRPVSNGDLYYIGT	29	9
AB005795	4866-6563	gb:AB005795:4866-6563|Organismo:Vírus Sendai|Nome da Cepa:Ohita|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MATYIQRVQCISALLSVVLTTLVSCQIPRDRLSNIGVIVDEGKSLK IAGSHESRYIVLSLVPGIDLENGCGTAQVIQYKSLLNRLLIPLRDA LDLQEALITVTNDTMTGADVPQSRFFGAVIGTIALGVATSAQITA GIALAEAREAKRDIALIKESMTKTHKSIELLQNAVGEQILALKTLQ DFVNDEIKPAISELGCETAALRLGIKLTQHYSELLTAFGSNFGTIG EKSLTLQALSSLYSANITEIMTTIRTGQSNIYDVIYTEQIKGTVIDV DLERYMVTLSVKIPILSEVPGVLIHKASSISYNIDGEEWYVTVPSH ILSRASFLGGANIADCVESRLTYICPRDPAQLIPDSQQKCILGDTT RCPVTKVVDNIIPKFAFVNGGVVANCIASTCTCGTGRRPISQDR SKGVVFLTHDNCGLIGVNGIELYANRKGHDATWGVQNLTVGPAI AIRPVDISLNLAAATDFLQDSRAELEKARKILSEVGRWYNSGATL ITIIVVMIVVLVVIIVIVIVLYRLRRSMLMSNPAGRISRDTYTLEPKIR HMYTNGGFDAMTEKR	23	10

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AF457102	5088-6755	gb:AF457102|Organismo:cepa do vírus 1 de parainfluenza humana Washington/1964|Nome da Cepa:Washington 1964|Nome da Proteína:F glicoproteína|Símbolo do Gene:F	MQKSEILFLVYSSLLLSSSLCQIPVEKLSNVGVIINEGKLLKIAGSY ESRYIVLSLVPSIDLQDGCGTTQIIQYKNLLNRLLIPLKDALDLQE SLITITNDTTVTNDNPQTRFFGAVIGTIALGVATAAQITAGIALAEA REARKDIALIKDSIVKTHNSVELIQRGIGEQIIALKTLQDFVNDEIR PAIGELRCETTALKLGIKLTQHYSELATAFSSNLGTIGEKSLTLQA LSSLYSANITEILSTTKKDKSDIYDIIYTEQVKGTVIDVDLEKYMVT LLVKIPILSEIPGVLIYRASSISYNIEGEEWHVAIPNYIINKASSLGG ADVTNCIESKLAYICPRDPTQLIPDNQQKCILGDVSKCPVTKVIN NLVPKFAFINGGVVANCIASTCTCGTNRIPVNQDRSRGVTFLTY TNCGLIGINGIELYANKRGRDTTWGNQIIKVGPAVSIRPVDISLNL ASATNFLEESKTELMKARAIISAVGGWHNTESTQIIMIIIVCILIIIC GILYYLYRVRRLLVMINSTHNSPVNAYTLESRMRNPYMGNNSN	21	11
AB910309	4951-6582	gb:AB910309:4951-6582|Organismo:morbilivírus Felino|Nome da Cepa:SS1|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MGKIRVIIISSLLLSNITTAQVGWDNLTSIGVISTKQYDYKITTLNT NQLMVIKMVPNISSIINCTKPELMKYRELVLGVIRPINESLELMNS YINMRAGSERFIGAVIAGVALGVATAAQITSGIALHNSIMNKRQIQ ELRKALSTTNKAIDEIRIAGERTLIAVQGVQDYINNIIIPMQDKLQC DILSSQLAIALLRYYTNILTVFGPSIRDPVTSIISIQALSQAFNGNL QALLDGLGYTGRDLRDLLESRSITGQIIHADMTDLFLVLRINYPSI TEMQGVTIYELNSITYHIGPEEWYTIMPNFIAVQGFLTSNFDERK CSITKSSILCQQNSIYPMSTEMQRCIKGEIRFCPRSKAVGTLVNR FILTKGNLMANCLGVICRCYSSGQIITQDPSKLITIISQEECKEVG VDGIRIMVGPRKLPDVIFNARLEVGVPISLSKLDVGTDLAIASAKL NNSKALLEQSDKILDSMSKLDSINSRITGLILAIMAIFIITVTIIWIIYK RCRNKDNKFSTSIEPLYIPPSYNSPHSVVKSI	12	12

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
KT071755	4310-6070	gb:KT071755:4310-6070|Organismo:Paramixovírus aviário 2|Nome da cepa:APMV-2/Procarduelis nipalensis/China/Suiling/53/2013|Nome da proteína:proteína de fusão|Símbolo do gene:F	MIAALFISLFATCGALDNSVLAPVGIASAQEWQLAAYTNTLSGTI AVRFVPVLPGNLSTCAQATLAEYNKTVTNILGPLKENLETLLSEP TKTAARFVGAIIGTVALGVATSAQITAAVALNQAQENARNIWRLK ESIRKTNEAVLELKDGLASTAIALDKVQKFINEDIIPQIKEIDCQVV ANKLGVYLSLYLTELTTIFGAQITNPALTPLSYQALYNLCGGDMG KLTELIGVKAKDINSLYEANLITGQVIGYDSESQIILIQVSYPSVSE VTGVRATELVTVSVTTPKGEGRAIAPKYVAQSRVVTEELDTSTC RFSKTTLYCRSIITRPLPPLIANCLNGLYQDCQYTTEIGALSSRFI TVNGGIIANCRATICKCVNPPKIIVQSDASSLTVIDSAICKDVVLD NVQLRLEGKLSAQYFTNITIDLSQITTSGSLDISSEIGSINNTVNKV EELIAESNAWLQAVNPHLVNNTSIIVLCVLAAIFVVWLVALTGCLA YYIKKSSATRMVGIGSSPAGNPYVAQSATKM	12	13
AY029299	4598-6265	gb:AY029299|Organismo:Pramixovírus 6 Aviário|Nome da Cepa:APMV-6/duck/Taiwan/Y1/98|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MGARLGPLAMAPGRYVIIFNLILLHRVVSLDNSRLLQQGIMSATE REIKVYTNSITGSIAVRLIPNLPQEVLKCSAGQIKSYNDTLNRIFTP IKANLERLLATPSMLEDNQNPAPEPRLIGAIIGTAALGLATAAQVT AALALNQAQDNAKAILNLKESITKTNEAVLELKDATGQIAIALDKT QRFINDNILPAINNLTCEVAGAKVGVELSLYLTELSTVFGSQITNP ALSTLSIQALMSLCGNDFNYLLNLMGAKHSDLGALYEANLINGRI IQYDQASQIMVIQVSVPSISSISGLRLTELFTLSIETPVGEGKAVV PQFVVESGQLLEEIDTQACTLTDTTAYCTIVRTKPLPELVAQCLR GDESRCQYTTGIGMLESRFGVFDGLVIANCKATICRCLAPEMIIT QNKGLPLTVISQETCKRILIDGVTLQIEAQVSGSYSRNITVGNSQI APSGPLDISSELGKVNQSLSNVEDLIDQSNQLLNRVNPNIVNNT AIIVTIVLLVLLVLWCLALTISILYVSKHAVRMIKTVPNPYVMQAKS PGSATQF	11	14

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AY141760	5028-6665	gb:AY141760|Organismo:Paramixovírus da Fer-de-Lance|Nome da Cepa:ATCC VR-895|Nome da Proteína:Proteína de fusão F|Símbolo do Gene:F	MTRITILQIILTLTLPVMCQVSFDNLEQVGVMFDKPKFLKITGPAS TATMIIKLIPTLGTMESCGTSAVNEYKKTLDTILVPLRDTINKLSTD ITVVEGTSNISNKREKRFVGIAIAVGAVALATSAQITAGIALSNTIK NAEAIESIKSSIQASNQAIQKVIDAQGRTVTVINGIQDHINSVINPA LNQLGCDVAKNTLAISLTQYFSKLSLLFGPNLRNPVEQPLSVQAI AGLMDGDINAVVSQLGYTQSDLLDLLSTESIVGTVTAIDMVNYMI QIEMSFPQYITIPDTKVLEGHKITFNDKGSEWQTQVPSTIAVRDIL IAGVDPDGCSITSTSYICKNDPTYAMSEVLTNCFRGNTQECPRA RITSTFATRFAIARSTVIANCVAAVCLCGDPGIPVVQKAEVTLTA MTLDQCSLITVDGLQIKPSKSIANVTANFGNITLGPVVSVGDLDL SAELTKVQSDLKEAQDKLDESNAILQGINNKILTAPTSIALIVVSV VVILLIIGMISWLVWLTKAVRRSNTRSERVTPSAYNNLGFIK	8	15
EU877976	4330-6410	gb:EU877976:4330-6410|Organismo:Paramixovírus aviário 4|Nome da cepa:APMV-4/KR/YJ/06|Nome da proteína:proteína de fusão|Símbolo do gene:F	MRLSRTILTLILGTLTGYLMGAHSTNVNEGPKSEGIRGDLIPGAG IFVTQVRQLQIYQQSGYHDLVIRLLPLLPAELNDCQREVVTEYNN TVSQLLQPIKTNLDTLLADGGTRDADIQPRFIGAIIATGALAVATV AEVTAAQALSQSKTNAQNILKLRDSIQATNQAVFEISQGLEATAT VLSKLQTELNENIIPSLNNLSCAAMGNRLGVSLSLYLTLMTTLFG DQITNPVLTPISYSTLSAMAGGHIGPVMSKILAGSVTSQLGAEQL IASGLIQSQVVGYDSQYQLLVIRVNLVRIQEVQNTRVVSLRTLAV NRDGGLYRAQVPPEVVERSGIAERFYADDCVLTTTDYICSSIRS SRLNPELVKCLSGALDSCTFERESALLSTPFFVYNKAVVANCKA ATCRCNKPPSIIAQYSASALVTITTDTCADLEIEGYRFNIQTESNS WVAPNFTVSTSQIVSVDPIDISSDIAKINSSIEAAREQLELSNQILS RINPRIVNDESLIAIIVTIVVLSLLVIGLIVVLGVMYKNLKKVQRAQA AMMMQQMSSSQPVTTKLGTPF	8	16

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AB176531	4793-6448	gb:AB176531:4793-6448|Organismo:Parainfluenza vírus 2 humano|Nome da Cepa:Nishio|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MHHLHPMIVCIFVMYTGIVGSDAIAGDQLLNIGVIQSKIRSLMYYTDGGASFIVVKLLPNLPPSNGTCNITSLDAYNVTLFKLLTPLIENLSKISTVTDTKTRQKRFAGVVVGLAALGVATAAQITAAVAIVKANANAAAINNLASSIQSTNKAVSDVIDASRTIATAVQAIQDRINGAIVNGITSASCRAHDALIGSILNLYLTELTTIFHNQITNPALTPLSIQALRILLGSTLPIVIESKLNTNFNTAELLSSGLLTGQIISISPMYMQMLIQINVPTFIMQPGAKVIDLIAISANHKLQEVVVQVPNRILEYANELQNYPANDCVVTPNSVFCRYNEGSPIPESQYQCLRGNLNSCTFTPIIGNFLKRFAFANGVLYANCKSLLCRCADPPHVVSQDDTQGISIIDIKRCSEMMLDTFSFRITSTFNATYVTDFSMINANIVHLSPLDLSNQINSINKSLKSAEDWIADSNFFANQARTAKTLYSLSAIALILSVITLVVVGLLIAYIIKLVSQIHQFRSLAATTMFHRENPAFFSKNNHGNIYGIS	7	17
BK005918	4677-6302	gb:BK005918|Organismo:Rubulavírus suíno|Nome da cepa:UNKNOWN-BK005918|Nome da proteína:proteína de fusão|Símbolo do gene:F	MPQQQVAHTCVMLWGIISTVSGINTEALSQYGVVVTNVRQLTYYTQAGSTYLAVRLLPSLASPDQSCALHSIINYNATLQAILSPIAENLNLISTALREQHRKKRFAGVAIGLTALGVATAAQATAAVALVRANKNAEKVEQLSQALGETNAAISDLIDATKNLGFAVQAIQNQINTAILPQIHNLSCQVIDAQLGNILSLYLTELTTVFQPQLTNPALSPLTIQALRAVLGTTLPALLSEKLKSNIPLGDLMSSGLLKGQLVGLNLQNMLMIIELYIPTLSTHSTAKVLDLVTISSHVNGREVEIQVPNRVLELGSEVLGYGGSECALTMSHILCPFNDARVLSTDMKYCLQGNITHCIFSPVVGSFLRRFALVNGVVIANCADMSCVCFDPQEIIYQNFQEPTTVIDIKKCGKVQLDTLTFTISTFANRTYGPPAYVPPDNIIQSEPLDISGNLIAVNNSLSSALNHLATSEILRNEQIWTSSLGISTIVALVIIGILIICLVVTWAALWALLKEVRGLNSAVNSQLSSYVMGDKFIRY	7	18

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
KC237063	4530-6185	gb:KC237063:4530-6185|Organismo:vírus da Parainfluenza 5|Nome da cepa:08-1990|Nome da proteína:proteína de fusão|Símbolo do gene:F|Segmento:4	MGTRIQFLMVSCLLAGTGSLDPAALMQIGVIPTNVRQLMYYTEASSAFIVVKLMPTIDSPISGCNITSISSYNATMTKLLQPIGENLETIRYQLIPTRRRRRFVGVVIGLAALGVATAAQVTAAVALVKANKNAAAILNLKNAIQKTNAAVADVVQATQSLGTAVQAVQDHINSVVSPAITAANCKAQDAIIGSILNLYLTELTTIFHNQITNPALSPITIQALRILLGSTLPTVVRKSFNTQISAAELLSSGLLTGQIVGLDLTYMQMVIKIELPTLTVQPATQIIDLVTISAFINNREVMAQLPTRIIVTGSLIQAYPASQCTITPNTVYCRYNDAQVLSDDTMACLQGNLTRCTFSPVVGSFLTRFVLFDGIVYANCRSMLCKCMQPAAVILQPSSSPVTVIDMHKCVSLQLDNLRFTITQLANITYNSTIKLETSQILPIDPLDISQNLAAVNKSLSDALQHLAQSDTYLSAITSATTTSVLSIIAICLGSLGLILIILISVVVWKLLTIVAANRNRMENFVYHNSAFHHSRSDLSEKNQPATLGTR	7	19
AY729016	5862-7523	gb:AY729016:5862-7523|Organismo:Vírus da pneumonia de Murganho|Nome da Cepa:15; ATCC VR-25|Nome da Proteína:Precursor de glicoproteína de fusão|Símbolo do Gene:F	MIPGRIFLVLLVIFNTKPIHPNTLTEKFYESTCSVETAGYKSALRTGWHMTVMSIKLSQINIESCKSSNSLLAHELAIYSSAVDELRTLSSNALKSKRKKRFLGLILGLGAAVTAGVALAKTVQLESEIALIRDAVRNTNEAVVSLTNGMSVLAKVVDDLKNFISKELLPKINRVSCDVHDITAVIRFQQLNKRLLEVSREFSSNAGLTHTVSSFMLTDRELTSIVGGMAVSAGQKEIMLSSKAIMRRNGLAILSSVNADTLVYVIQLPLFGVMDTDCWVIRSSIDCHNIADKYACLARADNGWYCHNAGSLSYFPSPTDCEIHNGYAFCDTLKSLTVPVTSRECNSNMYTTNYDCKISTSKTYVSTAVLTTMGCLVSCYGHNSCTVINNDKGIIRTLPDGCHYISNKGVDRVQVGNTVYYLSKEVGKSIVVRGEPLVLKYDPLSFPDDKFDVAIRDVEHSINQTRTFLKASDQLLDLSENRENKNLNKSYILTTLLFVVMLIIIMAVIGFILYKVLKMIRDNKLKSKSTPGLTVLS	6	20

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
AB543336	5174-6805	gb:AB543336:5174-6805|Organismo:Vírus da parainfluenza Humana 4a|Nome da Cepa:M-25|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MGVKGLSLIMIGLLISPITNLDITHLMNLGTVPTAIRSLVYYTYTKPSYLTVDLIPNLKNLDQNCNYSSLNYYNKTALSLIQPIADNINRLTKPITSSEIQSRFFGAVIGTIALGVATAAQVTAAIGLAKAQENAKLILTLKKAATETNEAVRDLANSNKIVVKMISAIQNQINTIIQPAIDQINCQIKDLQVANILNLYLTEITTVFHNQLTNPALESISIQALKSLLGPTLPEVLSKLDLNNISAASVMASGLIKGQIIAVDIPTMTLVLMVQIPSISPLRQAKIIDLTSITIHTNSQEVQAVVPARFLEIGSEILGFDGSVCQITKDTIFCPYNDAYELPIQQKRCLQGQTRDCVFTPVAGTFPRRFLTTYGTIVANCRDLVCSCLRPPQIIYQPDENPVTIIDKDLCTTLTLDSITIEIQKSINSTFRREVVLESTQVRSLTPLDLSTDLNQYNQLLKSAEDHIQRSTDYLNSINPSIVNNNAIIILIILCILLILTVTICIIWLKYLTKEVKNVARNQRLNRDADLFYKIPSQIPVPR	5	21
AF298895	4834-6450	gb:AF298895|Organismo:vírus Tioman|Nome da Cepa:AF298895 DESCONHECIDA|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MRIALTAVIVSIHFDLAFPMNKNSLLSVGLVHKSVKNLYFYSQGSPSYIVVKLVPTLGNVPGNCTLNSLVRYKSTVSSLLSPLAENLEYLQKTLTVSRGGRRRRFAGVAIGLAALGVAAAAQATAAVALVEARQNAAQIQSLSEAIQNTNLAVNELKTAIGASATAIQAIQTQINEVINPAINRLSCEILDAQLASMLNLYLIHLTTVFQNQLTNPALTPLSIQSLQSLLQGTSSVLTNITSSSKLALNDALVTGLITGQVVGLNMTSLQIVIAAYVPSVAKLSNAVVHNFIRITTSVNGTEVIIQSPTIIMEQNEVMYDLKTGHCTESDLNIYCPYVDAQLLSPGMTNCINGRLNDCTFSKVVGSFPTRFAAVEGAILANCKYLQCNCLTPPYIITPLNGEMISMIDLSKCQRLDLGTIVFDINNPVNVTFNGNYRADVGQMIVTNPLDISAELNQINTSLSNAQGFLSKSDAWLHVSQWVTNSGTIFIILIIGLIVGIVYMIINTYVVVQIIKEINRMRTSDRAHLLKGSISSIST	5	22

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
FJ215863	4499-6130	gb:FJ215863:4499-6130|Organismo:Paramixovírus aviário 8|Nome da cepa:ganso/Delaware/1053/76|Nome da Proteína:proteína de fusão|Símbolo do gene:F	MGQISVYLINSVLLLLVYPVNSIDNTLIAPIGVASANEWQLAAYTTSLSGTIAVRFLPVLPDNMTTCLRETITTYNNTVNNILGPLKSNLDALLSSETYPQTRLIGAVIGSIALGVATSAQITAAVALKQAQDNARNILALKEALSKTNEAVKELSSGLQQTAIALGKIQSFVNEEILPSINQLSCEVTANKLGVYLSLYLTELTTIFGAQLTNPALTSLSYQALYNLCGGNMAMLTQKIGIKQQDVNSLYEAGLITGQVIGYDSQYQLLVIQVNYPSISEVTGVRATELVTVSVTTDKGEGKAIVPQFVAESRVTIEELDVASCKFSSTTLYCRQVNTRALPPLVASCLRGNYDDCQYTTEIGALSSRYITLDGGVLVNCKSIVCRCLNPSKIISQNTNAAVTYVDATICKTIQLDDIQLQLEGSLSSVYARNISIEISQVTTSGSLDISSEIGNINNTVNRVEDLIHQSEEWLAKVNPHIVNNTTLIVLCVLSALAVIWLAVLTAIIIYLRTKLKTISALAVTNTIQSNPYVNQTKRESKF	5	23
JN689227	4689-6521	gb:JN689227:4689-6521|Organismo:Vírus Tailam|Nome da Cepa:TL8K|Nome da proteína:proteína de fusão|Símbolo do gene:F	MKLSVVYTTLLVSTFYSDLARSQLALSELTKIGVIPGRSYDLKISTQASYQYMVVKLIPNLTGLNNCTNGTIEAYKKMLNRLLSPIDAALRKMKDAVNDKPPESVGNVKFWGAVIGGVALGVATSAQITAGVALHNSIQNANAILALKDSIRQSNKAIQELQTAMSTTVVVLNALQDQINNQLVPAINSLGCQVVANTLGLKLNQYFSEISLVFGPNLRDPTSETLSIQALSRAFNGDFDSMLSKLKYDDSDFLDLLESDSIRGRIIDVSLSDYLITIQIEYPALLSIKDAVIQTFNLISYNTRGTEWISIFPKQLLVRGTYISNIDISQCVIAATSIICKSDTSTPISSATWSCATGNITNCARTRVVNAHVPRFALYGGVVFANCAPVVCKCQDPLYSINQEPKVTNVMVDVDACKEMYLDGLYITLGKTQISRAMYAEDVSLGGPISVDPIDLGNEINSINSAINRSEEHLNHANELLDKVNPRIVNVKTFGVMIGLLVLVVLWCVITLVWLICLTKQLARTAYAGSMGSRASTVNSLSGFVG	5	24

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
JX857409	4831-6615	gb:JX857409:4831-6615|Organismo:vírus da parainfluenza suína 1|Nome da cepa:S206N|Nome da proteína:proteína de fusão|Símbolo do gene:F	MQVTTLRPAIILSIALLVTGQVPRDKLANLGIIIKDSKALKIAGSYENRYIVLSLVPTIDNVNGCGSIQIAKYKEMLERLLIPIKDALDLQESLIVIDNETVNNNYSPQYRFVGAIIGTIALGVATAAQVTAGVALMEAREAKRDISMLKEAIEKTQNSIEKLQNSAGEQILALKMLQDYVNGEIKPAIEELGCETAALKLGIALTQHYTELTNAFGSNLGSIGEKSLTLQALSSLYKTNITNILTATNLGKTDIYDIIYAEQVKGRVIDVDLKRYMVTISVKIPILSEIPGVLIYEVSSISYNIDGAEWYAAVPDHILSKSAYIGGADISDCIESRLTYICPQDPAQIIADNQQQCFFGHLDKCPITKVIDNLVPKFAFINGGVVANCIASTCTCGEERIQVSQDRNKGVTFLTHNNCGLIGINGIEFHANKKGSDATWNVSPIGVGPAVSLRPVDISLQIVAATNFLNSSRKDLMKAKEILNQVGNLKDLTTITIINIVIIILLICVIGLGILYHQLRSALGMRDKMSVLNNSSYSLEPRTAQVQVIKPTSFMG	5	25
AY640317	2932-4571	gb:AY640317:2932-4571|Organismo:Metapneumovírus aviário|Nome da Cepa:LAH A|Nome da Proteína:F|Símbolo do Gene:F	MDVRICLLLFLISNPSSCIQETYNEESCSTVTRGYKSVLRTGWYTNVFNLEIGNVENITCNDGPSLIDTELVLTKNALRELKTVSADQVAKESRLSSPRRRRFVLGAIALGVATAAAVTAGVALAKTIRLEGEVKAIKNALRNTNEAVSTLGNGVRVLATAVNDLKEFISKKLTPAINQNKCNIADIKMAISFGQNNRRFLNVVRQFSDSAGITSAVSLDLMTDDELVRAINRMPTSSGQISLMLNNRAMVRRKGFGILIGVYDGTVVYMVQLPIFGVIETPCWRVVAAPLCRKRRGNYACILREDQGWYCTNAGSTAYYPNKDDCEVRDDYVFCDTAAGINVALEVDQCNYNISTSKYPCKVSTGRHPVSMVALTPLGGLVSCYESVSCSIGSNKVGIIKQLGKGCTHIPNNEADTITIDNTVYQLSKVVGEQRTIKGAPVVNNFNPILFPVDQFNVALDQVFESIDRSQDLIDKSNDLLGADAKSKAGIAIAIVVLVILGIFFLLAVIYYCSRVRKTKPKHDYPATTGHSSMAYVS	4	26

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
KU646513	4641-6498	gb:KU646513:4641-6498|Organismo:Paramixovírus aviário 13 ganso/Kazakhstan/5751/2013|Nome da cepa:APMV-13/ganso da testa grande/Northern Kazakhstan/5751/2013|Nome da proteína:proteína de fusão|Símbolo do gene:F	MARFSWEIFRLSTILLIAQTCQGSIDGRLTLAAGIVPVGDRPISIYTSSQTGIIVVKLIPNLPDNKKDCAKQSLQSYNETLSRILTPLATAMSAIRGNSTTQVRENRLVGAIIGSVALGVATAAQITAATALIQANQNAANIARLANSIAKTNEAVTDLTEGLGTLAIGVGKLQDYVNEQFNNTAVAIDCLTLESRLGIQLSLYLTELMGVFGNQLTSPALTPITIQALYNLAGGNLNALLSRLGASETQLGSLINSGLIKGMPIMYDDANKLLAVQVELPSIGKLNGARSTLLETLAVDTTRGPSSPIIPSAVIEIGGAMEELDLSPCITTDLDMFCTKIISYPLSQSTLSCLNGNLSDCVFSRSEGVLSTPYMTIKGKIVANCKQVICRCMDPPQILSQNYGEALLLIDENTCRSLELSGVILKLAGTYESEYTRNLTVDPSQVIITGPLDISAELSKVNQSIDSAKENIAESNKFLSQVNVKLLSSSAMITYIVATVVCLIIAITGCVIGIYTLTKLKSQQKTLLWLGNNAEMHGSRSKTSF	4	27
AF326114	4818-6482	gb:AF326114|Organismo:Vírus Menangle|Nome da Cepa:AF326114 DESCONHECIDA|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MMPRVLGMIVLYLTHSQILCINRNTLYQIGLIHRSVKKVNFYSQGSPSYIVVKLVPTLAAIPPNCSIKSLQRYKETVTSLVQPISDNLGYLQDKLVTGQSRRRRRFAGVAIGLAALGVAAAAQATAAVALVETRENAGKIQALSESIQNTNQAVHSLKTALGFSATAIQAIQNQVNEVINPAINKLSCEVLDSQLASMLNLYLIHLTTVFQTQLTNPALTPLSIQALTSVLQGTSGVLMNSTNSTLTQPIDLLATGLITGQIISVNMTSLQLIIATFMPSIAELPNAVLHSFFRITTSVNLTEVMIQSPEFIMEQNGVFYDFNTAHCQLGDNNVYCPYIDAARLSSMMTNCINGNLGECVFSRVIGSFPSRFVSLNGAILANCKFMRCNCLSPEKIITPLDGEMISLIDLRVCQKLTLGTITFEISQPVNVSFQGGFVANAGQIIVTNPFDISAELGQINNSLNDAQGFLDQSNNWLKVSGWINNSGSLFIAGIVVIGLIVLCIVIIIYINVQIIREVNRLRSFIYRDYVLDHDKAPYSPESSSPHRKSLKTVS	3	28

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
GU206351	5441-7468	gb:GU206351:5441-7468|Organismo:Paramixovírus aviário 5|Nome da cepa:periquito australiano/Kunitachi/74|Nome da proteína:proteína de fusão|Símbolo do gene:F	MLQLPLTILLSILSAHQSLCLDNSKLIHAGIMSTTEREVNVYAQSI TGSIVVRLIPNIPSNHKSCATSQIKLYNDTLTRLLTPIKANLEGLIS AVSQDQSQNSGKRKKRFVGAVIGAAALGLATAAQVTATVALNQ AQENARNILRLKNSIQKTNEAVMELKDAVGQTAVAIDKTQAFINN QILPAISNLSCEVLGNKIGVQLSLYLTELTTVFGNQLTNPALTTLS LQALYNLCGDDFNYLINLLNAKNRNLASLYEANLIQGRITQYDSM NQLLIIQVQIPSISTVSGMRVTELFTLSVDTPIGEGKALVPKYVLS SGRIMEEVDLSSCAITSTSVFCSSIISRPLPLETINCLNGNVTQCQ FTANTGTLESRYAVIGGLVIANCKAIVCRCLNPPGVIAQNLGLPIT IISSNTCQRINLEQITLSLGNSILSTYSANLSQVEMNLAPSNPLDIS VELNRVNTSLSKVESLIKESNSILDSVNPQILNVKTVIILAVIIGLIV VWCFILTCLIVRGFMLLVKQQKFKGLSVQNNPYVSNNSH	3	29
JQ001776	6129-8166	gb:JQ001776:6129-8166|Organismo:Vírus do cedro|Nome da cepa:CG1a|Nome da proteína:glicoproteína de fusão|Símbolo do gene:F	MSNKRTTVLIIISYTLFYLNNAAIVGFDFDKLNKIGVVQGRVLNYKI KGDPMTKDLVLKFIPNIVNITECVREPLSRYNETVRRLLLPIHNML GLYLNNTNAKMTGLMIAGVIMGGIAIGIATAAQITAGFALYEAKKN TENIQKLTDSIMKTQDSIDKLTDSVGTSILILNKLQTYINNQLVPNL ELLSCRQNKIEFDLMLTKYLVDLMTVIGPNINNPVNKDMTIQSLS LLFDGNYDIMMSELGYTPQDFLDLIESKSITGQIIYVDMENLYVVI RTYLPTLIEVPDAQIYEFNKITMSSNGGEYLSTIPNFILIRGNYMS NIDVATCYMTKASVICNQDYSLPMSQNLRSCYQGETEYCPVEA VIASHSPRFALTNGVIFANCINTICRCQDNGKTITQNINQFVSMID NSTCNDVMVDKFTIKVGKYMGRKDINNINIQIGPQIIIDKVDLSNEI NKMNQSLKDSIFYLREAKRILDSVNISLISPSVQLFLIIISVLSFIILLI IIVYLYCKSKHSYKYNKFIDDPDYYNDYKRERINGKASKSNNIYY VGD	3	30

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
LC168749	4869-7235	gb:LC168749:4869-7235|Organismo:morbilivírus da peste bovina|Nome da cepa:Lv|Nome da proteína:proteína F|Símbolo do gene:F	MGILFAALLAMTNPHLATGQIHWGNLSKIGVVGTGSASYKVMTQSSHQSLVIKLMPNVTAIDNCTKTEIMEYKRLLGTVLKPIREALNAITKNIKPIQSSTTSRRHKRFAGVVLAGAALGVATAAQITAGIALHQSMMNSQAIESLKASLETTNQAIEEIRQAGQEMVLAVQGVQDYINNELVPAMGQLSCEIVGQKLGLKLLRYYTEILSLFGPSLRDPVSAELSIQALSYALGGDINKILEKLGYSGSDLLAILESKGIKAKITYVDIESYFIVLSIAYPSLSEIKGVIVHRLESVSYNIGSQEWYTTVPRYVATQGYLISNFDDTPCAFTPEGTICSQNALYPMSPLLQECFRGSTRSCARTLVSGSIGNRFILSKGNLIANCASILCKCYTTGSIISQDPDKILTYIAADQCPVVEVGGVTIQVGSREYSDAVYLHEIDLGPPISLEKLDVGTNLWNAVTKLEKAKDLLDSSDLILENIKGVSVTNTGYILVGVGLIAVVGILIITCCCKKRRSDNKVSTMVLNPGLRPDLTGTSKSYVRSL	2	31
LC187310	6250-7860	gb:LC187310:6250-7860|Organismo:paramixovírus aviário 10|Nome da cepa:rAPMV-10-FI324/YmHA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MTRTRLLFLLTCYIPGAVSLDNSILAPAGIISASERQIAIYTQTLQGTIALRFIPVLPQNLSSCAKDTLESYNSTVSNLLLPIAENLNALLKDADKPSQRIIGAIIGSVALGVATTAQVTAALAMTQAQQNARNIWKLKESIKNTNQAVLELKDGLQQSAIALDKVQSFINSEILPQINQLGCEVAANKLGIFLSLYLTEITTVFKNQITNPALSTLSYQALYNLCGGNMAALTKQIGIKDTEINSLYEAELITGQVIGYDSADQILLIQVSYPSVSRVQGVRAVELLTVSVATPKGEGKAIAPSFIAQSNIIAEELDTQPCKFSKTTLYCRQVNTRTLPVRVANCLKGKYNDCQYTTEIGALASRYVTITNGVVANCRSIICRCLDPEGIVAQNSDAAITVIDRSTCKLIQLGDITLRLEGKLSSSYSKNITIDISQVTTSGSLDISSELGSINNTITKVEDLISKSNDWLSKVNPTLISNDTIIALCVIAGIVVIWLVIITILSYYILIKLKNVALLSTMPKKDLNPYVNNTKF	2	32

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_005283	5277-6935	gb:NC_005283:5277-6935|Organismo:Morbilivírus de golfinho|Nome da cepa:NC_005283 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MAASNGGVMYQSFLTIIILVIMTEGQIHWGNLSKIGIVGTGSASY KVMTRPNHQYLVIKLMPNVTMIDNCTRTEVTEYRKLLKTVLEPV KNALTVITKNIKPIQSLTTSRRSKRFAGVVLAGVALGVATAAQITA GVALHQSIMNSQSIDNLRTSLEKSNQAIEEIRQASQETVLAVQG VQDFINNELIPSMHQLSCEMLGQKLGLKLLRYYTEILSIFGPSLR DPVSAEISIQALSYALGGDINKILEKLGYSGADLLAILESRGIKAKV THVDLEGYFIVLSIAYPTLSEVKGVIVHKLEAVSYNLGSQEWYTT LPKYVATNGYLISNFDESSCAFMSEVTICSQNALYPMSPLLQQC LRGSTASCARSLVSGTIGNRFILSKGNLIANCASVLCKCYSTGTII SQDPDKLLTFVAADKCPLVEVDGITIQVGSREYPDSVYVSRIDL GPAISLEKLDVGTNLGSALTKLDNAKDLLDSSNQILENVRRSSF GGAMYIGILVCAGALVILCVLVYCCRRHCRKRVQTPPKATPGLK PDLTGTTKSYVRSL	2	33
NC_005339	5374-7602	gb:NC_005339:5374-7602|Organismo:vírus Mossman|Nome da cepa:NC_005339 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MSNYFPARVIIIVSLITAVSCQISFQNLSTIGVFKFKEYDYRVSGD YNEQFLAIKMVPNVTGVENCTASLIDEYRHVIYNLLQPINTTLTAS TSNVDPYAGNKKFFGAVIAGVALGVATAAQVTAGVALYEARQN AAAIAEIKESLHYTHKAIESLQISQKQTVVAIQGIQDQINTNIIPQIN ALTCEIANQRLRLMLLQYYTEMLSSFGPIIQDPLSGHITVQALSQ AAGGNITGLMRELGYSSKDLRYILSVNGISANIIDADPEIGSIILRI RYPSMIKIPDVAVMELSYLAYHAAGGDWLTVGPRFILKRGYSLS NLDITSCTIGEDFLLCSKDVSSPMSLATQSCLRGDTQMCSRTAV QDREAPRFLLLQGNLIVNCMSVNCKCEDPEETITQDPAYPLMVL GSDTCKIHYIDGIRIKLGKVQLPPITVLNTLSLGPIVVLNPIDVSNQ LSLVETTVKESEDHLKNAIGALRSQSRVGGVGIVAIVGLIIATVSL VVLVISGCCLVKYFSRTATLESSLTTIEHGPTLAPKSGPIIPTYINP VYRHD	2	34

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_007454	4635-6384	gb:NC_007454:4635-6384|Organismo:J-virus|Nome da cepa:NC_007454 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MKPVALIYLTILAFTVKVRSQLALSDLTKIGIIPAKSYELKISTQAA QQLMVIKLIPNVNGLTNCTIPVMDSYKKMLDRILKPIDDALNHVK NAIQDKQGDGVPGVRFWGAIIGGVALGVATSAQITAGVALHNSI QNANAILQLKESIRNSNKAIEELQAGLQSTVLVINALQDQINSQLV PAINTLGCSVIANTLGLRLNQYFSEISLVFGPNLRDPTSQTLSIQA IAKAFNGDFDSMMKKMHYTDSDFLDLLESDSIRGRIISVSLEDYL IIIQIDYPGLTTIPNSVVQTFNLITYNYKGTEWESIFPRELLIRGSYI SNIDISQCVGTSKSMICKSDTSTTISPATWACATGNLTSCARTRV VNSHSTRFALSGGVLFANCAPIACRCQDPQYSINQEPKTTNVM VTSEDCKELYIDGFYLTLGKKMLDRAMYAEDVALGGSVSVDPID IGNELNSINESINKSHEYLDKANELLEQVNPNIVNVSSFSFILVISI LLIIWFIVTLVWLIYLTKHMNFIVGKVAMGSRSSTVNSLSGFVG	2	35
NC_009489	4620-6500	gb:NC_009489:4620-6500|Organismo:Vírus Mapuera|Nome da cepa:BeAnn 370284|Nome da proteína:proteína de fusão|Símbolo do gene:F	MRSSLFLVLTLLVPFAHSIDSITLEQYGTVITSVRSLAYFLETNPT YISVRLMPAIQTDSSHCSYHSIENYNLTLTKLLLPLQENLHQITDS LSSRRRKKRFAGVAVGLAALGVATAAQVTAAIAVVKAKENSAKI AQLTSAISETNRAVQDLIEGSKQLAVAVQAIQDQINNVIQPQLTN LSCQVADAQVGTILNMYLTELTTVFHPQITNSALTPITIQALRSLL GSTLPQVVTSTIKTDVPLQDLLTSGLLKGQIVYLDLQSMIMVVSV SVPTIALHSMAKVYTLKAISAHVNNAEVQMQVPSRVMELGSEIM GYDIDQCEETSRYLFCPYNGGSILSATMKMCLNGNISQCVFTPI YGSFLQRFVLVDGVIVANCRDMTCACKSPSKIITQPDSLPVTIID STSCSNLVLDTLELPIISINNATYRPVQYVGPNQIIFSQPLDLLSQ LGKINSSLSDAIEHLAKSDEILEQIQWDSPQGYTLIALTSVLAFVV VAIVGLLISTRYLIFEIRRINTTLTQQLSSYVLSNKIIQY	2	36

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_017937	4534-6330	gb:NC_017937:4534-6330|Organismo:vírus Nariva|Nome da cepa:NC_017937 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MAEQEKTPLRYKILLIIIVINHYNITNVFGQIHLANLSSIGVFVTKTLDYRTTSDPTEQLLVINMLPNISNIQDCAQGVVNEYKHLISSLLTPINDTLDLITSNINPYSGRNKLFGEIIAGAALTVATSAQITAGVALYEARQNAKDIAAIKESLGYAYKAIDKLTTATREITVVINELQDQINNRLIPRINDLACEVWATRLQAMLLQYYAEIFSVIGPNLQDPLSGKISIQALARAAGGNIKLMVDELNYSGQDLSRLVKVGAIKGQIIDADPSLGVVIIKMRYPNIIKIPNVAISELSYVSYSSDGQDWITTGPNYIVTRGYSIANIQTSSCSVGDDFVLCDRDMTYPMSQVTQDCLRGNIALCSRMVVRDREAPRYLILQGNMVANCMSITCRCEEPESEIYQSPDQPLTLLTRDTCDTHVVDGIRIRLGVRKLPTISVINNITLGPIITTDPIDVSNQLNAVVSTIDQSAELLHQAQRVLSERARGARDHILATAAIVICVVLAVLILVLLIGLVYLYRTQNEILVKTTMLEQVPTFAPKSFPMESQIYSGKTNKGYDPAE	2	37
NC_025256	6865-8853	gb:NC_025256:6865-8853|Organismo:Paramixovírus de morcego Eid_hel/GH-M74a/GHA/2009|Nome da cepa:BatPV/Eid_hel/GH-M74a/GHA/2009|Nome da proteína:proteína de fusão|Símbolo do gene:F	MKKKTDNPTISKRGHNHSRGIKSRALLRETDNYSNGLIVENLVRNCHHPSKNNLNYTKTQKRDSTIPYRVEERKGHYPKIKHLIDKSYKHIKRGKRRNGHNGNIITIILLLILILKTQMSEGAIHYETLSKIGLIKGITREYKVKGTPSSKDIVIKLIPNVTGLNKCTNISMENYKEQLDKILIPINNIIELYANSTKSAPGNARFAGVIIAGVALGVAAAAQITAGIALHEARQNAERINLLKDSISATNNAVAELQEATGGIVNVITGMQDYINTNLVPQIDKLQCSQIKTALDISLSQYYSEILTVFGPNLQNPVTTSMSIQAISQSFGGNIDLLLNLLGYTANDLLDLLESKSITGQITYINLEHYFMVIRVYYPIMTTISNAYVQELIKISFNVDGSEWVSLVPSYILIRNSYLSNIDISECLITKNSVICRHDFAMPMSYTLKECLTGDTEKCPREAVVTSYVPRFAISGGVIYANCLSTTCQCYQTGKVIAQDGSQTLMMIDNQTCSIVRIEEILISTGKYLGSQEYNTMHVSVGNPVFTDKLDITSQISNINQSIEQSKFYLDKSKAILDKINLNLIGSVPISILFIIAILSLILSIITFVIVMIIVRRYNKYTPLINSDPSSRRSTIQDVYIIPNPGEHSIRSAARSIDRDRD	2	38

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025347	4471-6386	gb:NC_025347:4471-6386|Organismo:Paramixovírus aviário 7|Nome da cepa:APMV-7/ANDORINHA/Tennessee/4/75|Nome da proteína:proteína de fusão|Símbolo do gene:F	MRVRPLIIILVLLVLLWLNILPVIGLDNSKIAQAGIISAQEYAVNVYSQSNEAYIALRTVPYIPPHNLSCFQDLINTYNTTIQNIFSPIQDQITSITSASTLPSSRFAGLVVGAIALGVATSAQITAAVALTKAQQNAQEIIRLRDSIQNTINAVNDITVGLSSIGVALSKVQNYLNDVINPALQNLSCQVSALNLGIQLNLYLTEITTIFGPQITNPSLTPLSIQALYTLAGDNLMQFLTRYGYGETSVSSILESGLISAQIVSFDKQTGIAILYVTLPSIATLSGSRVTKLMSVSVQTGVGEGSAIVPSYVIQQGTVIEEFIPDSCIFTRSDVYCTQLYSKLLPDSILQCLQGSMADCQFTRSLGSFANRFMTVAGGVIANCQTVLCRCYNPVMIIPQNNGIAVTLIDGSLCKELELEGIRLTMADPVFASYSRDLIINGNQFAPSDALDISSELGQLNNSISSATDNLQKAQESLNKSIIPAATSSWLIILLFVLVSISLVIGCISIYFIYKHSTTNRSRNLSSDIISNPYIQKAN	2	39
NC_025348	4790-6570	gb:NC_025348:4790-6570|Organismo:vírus Tuhoko 2|Nome da cepa:NC_025348 DECONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MAPCVLFLSSLLLISTISPSHGINQPALRRIGAIVSSVKQLKFYSKTKPNYIIVKLLPTINLSKSNCNLTSINRYKESVIEIIKPLADNIDNLNQKLLPKNRRKRMAGVAIGLAALGVAAAAQATAAVALVEARKNTQMIQSLADSIQDTNAAVQAVNIGLQNSAVAIQAIQNQINNVINPALDRLNCEVLDAQIASILNLYLIKSVTIFQNQLTNPALQQLSIQMLSIVMQDTAKILGNFTIGDKFDQHDLLGSGLITGQVVGVNLTNLQLIIAAFIPSIAPLPQAYIIDLISITISVNDTEAVIQIPERIMEHGSSIYQFGGKQCVYGQFSAYCPFSDAVLMTQDLQLCMKGNIEHCIFSSVLGSFPNRFASVDGVFYANCKYMSCACSDPLQVIHQDDSVNLMVIDSSVCRSLTLGHVTFPIIAFSNVSYQMKTNISIEQMIVTSPLDLSTELKQINNSVNIANTFLDSSNRALKTSIFGTSSQIILIVLLIFTCLLILYVIFLTYIIKILIKEVKRLRDGNSRTGSKLSFINPDV	2	40

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025350	4663-6428	gb:NC_025350:4663-6428|Organismo:vírus Tuhoko 3|Nome da cepa:NC_025350 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MLWLTILIALVGNHESTCMNINFLQSLGQINSQKRFLNFYTQQPPSYMVIRLVPTLQLSANNCTLGSIVRYRNAIKELIQPMDENLRWLSSNLIPQRRGKRFAGVAVGLAALGVAVAAQATAAVALVEARANAEKIASMSQSIQETNKAVTSLSQAVSASGIAIQAIQNEINNVIHPILNQVQCDVLDARVGNILNLYLIKVTTIFQNQLTNPALQRLSTQALSMLMQSTSSYLRNLSSSESAINADLSMTNLIEAQIVGINMTNLQLVLAVFIPSIARLNGALLYDFISITISSNQTEVMLQIPHRVLEIGNSLYTFEGTQCEMTKLNAYCLYSDAIPVTESLRDCMNGLFSQCGFVRIIGSFANRFASVNGVIYANCKHLTCSCLQPDEIITQDTNVPLTIIDTKRCTKISLGHLTFTIREYANVTYSLRTEIANSQITVVSPLDLSSQLTTINNSLADATNHIMNSDRILDRLNSGLYSKWVIIFLICASIVSLIGLVFLGFLIRGLILELRSKHRSNLNKASTYSIDSSIGLT	2	41
NC_025352	5950-8712	gb:NC_025352:5950-8712|Organismo:Vírus Mojiang|Nome da cepa:Tongguan1|Nome da proteína:proteína de fusão|Símbolo do gene:F	MALNKNMFSSLFLGYLLVYATTVQSSIHYDSLSKVGVIKGLTYNYKIKGSPSTKLMVVKLIPNIDSVKNCTQKQYDEYKNLVRKALEPVKMAIDTMLNNVKSGNNKYRFAGAIMAGVALGVATAATVTAGIALHRSNENAQAIANMKSAIQNTNEAVKQLQLANKQTLAVIDTIRGEINNIIPVINQLSCDTIGLSVGIRLTQYYSEIITAFGPALQNPVNTRITIQAISSVFNGNFDELLKIMGYTSGDLYEILHSELIRGNIIDVDVDAGYIALEIEFPNLTLVPNAVVQELMPISYNIDGDEWVTLVPRFVLTRTTLLSNIDTSRCTITDSSVICDNDYALPMSHELIGCLQGDTSKCAREKVVSSYVPKFALSDGLVYANCLNTICRCMDTDTPISQSLGATVSLLDNKRCSVYQVGDVLISVGSYLGDGEYNADNVELGPPIVIDKIDIGNQLAGINQTLQEAEDYIEKSEEFLKGVNPSIITLGSMVVLYIFMILIAIVSVIALVLSIKLTVKGNVVRQQFTYTQHVPSMENINYVSH	2	42

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025363	4622-6262	gb:NC_025363:4622-6262|Organismo:Paramixovírus aviário 12|Nome da cepa:Piadeira-american/Itália/3920_1/2005|Nome da proteína:proteína de fusão|Símbolo do gene:F	MAIPVPSSTALMIFNILVSLAPASALDGRLLLGAGIVPTGDRQVN VYTSSQTGIIALKLLPNLPKDKENCAEVSIRSYNETLTRILTPLAQ SMAAIRGNSTVSTRGREPRLVGAIIGGVALGVATAAQITAATALI QANQNAENIARLAKGLAATNEAVTDLTKGVGSLAIGVGKLQDYV NEQFNRTGEAIECLTIESRVGVQLSLYLTEVIGVFGDQITSPALS DISIQALYNLAGGNLNVLLQKMGIEGTQLGSLINSGLIKGRPIMYD DGNKILGIQVTLPSVGRINGARATLLEAIAVATPKGNASPLIPRAV ISVGSLVEELDMTPCVLTPTDIFCTRILSYPLSDSLTTCLKGNLSS CVFSRTEGALSTPYVSVHGKIVANCKSVVCRCVEPQQIISQNYG EALSLIDESLCRILELNGVILKMDGQFTSEYTKNITIDPVQVIISGPI DISSELSQVNQSLDSALENIKESNSYLSKVNVKLISSSAMITYIVIT VICLILTFVALVLGIYSYTKIRSQQKTLIWMGNNIARSKEGNRF	2	43
NC_025373	4617-6582	gb:NC_025373:4617-6582|Organismo:Paramixovírus aviário 3|Nome da cepa:Turkey/Wisconsin/68|Nome da proteína:proteína de fusão|Símbolo do gene:F	MASPMVPLLIITVVPALISSQSANIDKLIQAGIIMGSGKELHIYQES GSLDLYLRLLPVIPSNLSHCQSEVITQYNSTVTRLLSPIAKNLNHL LQPRPSGRLFGAVIGSIALGVATSAQISAAIALVRAQQNANDILAL KAAIQSSNEAIKQLTYGQEKQLLAISKIQKAVNEQVIPALTALDCA VLGNKLAAQLNLYLIEMTTIFGDQINNPVLTPIPLSYLLRLTGSEL NDVLLQQTRSSLSLIHLVSKGLLSGQIIGYDPSVQGIIIRIGLIRTQ RIDRSLVFXPYVLPITISSNIATPIIPDCVVKKGVIIEGMLKSNCIEL ERDIICKTINTYQITKETRACLQGNITMCKYQQSRTQLSTPFITYN GVVIANCDLVSCRCIRPPMIITQVKGYPLTIINRNLCTELSVDNLIL NIETNHNFSLNPTIIDSQSRLIATSPLEIDALIQDAQHHAAAALLKV EESNAHLLRVTGLGSSSWHIILILTLLVCTIAWLIGLSIYVCRIKND DSTDKEPTTQSSNRGIGVGSIQYMT	2	44

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025386	5548-7206	gb:NC_025386:5548-7206|Organismo:Vírus Salem|Nome da cepa:NC_025386 DESCONHECIDO|Nome da proteína:proteína de fusão|Símbolo do gene:F	MNPLNQTLIAKVLGFLLLSSSFTVGQIGFENLTRIGVHQVKQYGYKLAHYNSHQLLLIRMIPTVNGTHNCTHQVITRYREMVREIITPIKGALDIMKKAVSPDLVGARIFGAIVAGAALGIATSAQITAGVALHRTKLNGQEISKLKEAVSLTNEAVEQLQYSQGKSILAIQGIQDFINFNVVPLLEEHTCGIAKLHLEMALMEYFQKLILVFGPNLRDPIGSTIGIQALATLFQNNMFEVSLRLGYAGDDLEDVLQSNSIRANIIEAEPDSGFIVLAIRYPTLTLVEDQVITELAHITFNDGPQEWVATIPQFVTYRGLVLANIDVSTCTFTERNVICARDQTYPMIIDLQLCMRGNIAKCGRTRVTGSTASRFLLKDGNMYANCIATMCRCMSSSSIINQEPSHLTTLIVKETCSEVMIDTIRITLGERKHPPIDYQTTITLGQPIALAPLDVGTELANAVSYLNKSKVLLEHSNEVLSSVSTAHTSLTATIVLGIVVGGLAILIVVMFLFLEAQVIKVQRAMMLCPITNHGYLPNEDLLTRGHSIPTIG	2	45
NC_025390	4805-6460	gb:NC_025390:4805-6460|Organismo:Paramixovírus aviário 9|Nome da cepa:pato/Nova York/22/1978|Nome da proteína:proteína de fusão|Símbolo do gene:F	MGYFHLLLILTAIAISAHLCYTTTLDGRKLLGAGIVITEEKQVRVYTAAQSGTIVLRSFRVVSLDRYSCMESTIESYNKTVYNILAPLGDAIRRIQASGVSVERIREGRIFGAILGGVALGVATAAQITAAIALIQANENAKNILRIKDSITKTNEAVRDVTNGVSQLTIAVGKLQDFVNKEFNKTTEAINCVQAAQQLGVELSLYLTEITTVFGPQITSPALSKLTIQALYNLAGVSLDVLLGRLGADNSQLSSLVSSGLITGQPILYDSESQILALQVSLPSISDLRGVRATYLDTLAVNTAAGLASAMIPKVVIQSNNIVEELDTTACIAAEADLYCTRITTFPIASAVSACILGDVSQCLYSKTNGVLTTPYVAVKGKIVANCKHVTCRCVDPTSIISQNYGEAATLIDDQLCKVINLDGVSIQLSGTFESTYVRNVSISANKVIVSSSIDISNELENVNSSLSSALEKLDESDAALSKVNVHLTSTSAMATYIVLTVIALILGFVGLGLGCFAMIKVKSQAKTLLWLGAHADRSYILQSKPAQSST	2	46

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025403	4826-6649	gb:NC_025403:4826-6649|Organismo:vírus Achimota 1|Nome da cepa:NC_025403 DESCONHECIDO|Nome da proteína:proteína de fusão|Símbolo do gene:F	MWIMIILSLFQIIPGVTPINSKVLTQLGVITKHTRQLKFYSHSTPSYLVVKLVPTINTESTVCNFTSLSRYKDSVRELITPLAKNIDNLNSILTIPKRRKRMAGVVIGLAALGVAAAAQATAAVALIEAKKNTEQIQALSESIQNTNKAVSSIEKGLSSAAIAVQAIQNQINNVINPALTALDCGVTDAQLGNILNLYLIKTLTVFQKQITNPALQPLSIQALNIIMQETSSVLRNFTKTDEIEHTDLLTSGLITGQVVGVNLTNLQLIIAAFIPSIAPLNQAYILDFIRITVNINNSESMIQIPERIMEHGISLYQFGGDQCTFSDWSAYCPYSDATLMAPGLQNCFRGQAADCVFSTVMGSFPNRFVSVQGVFYVNCKFIRCACTQPQRLITQDDSLSLTQIDAKTCRMLTLGFVQFSINEYANVTYSFKNNVTAGQLIMTNPIDLSTEIKQMNDSVDEAARYIEKSNAALNKLMYGGRSDIVTTVLLVGFILLVVYVIFVTYILKILMKEVARLRNSNHPDLIKPYNYPM	2	47
NC_025404	4772-6647	gb:NC_025404:4772-6647|Organismo:vírus Achimota 2|Nome da cepa:NC_025404 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MLNSFYQIICLAVCLTTYTVISIDQHNLLKAGVIVKSIKGLNFYSRGQANYIIVKLIPNVNVTDTDCDIGSIKRYNETVYSLIKPLADNIDYLRTQFAPTKRKKRFAGVAIGLTALGVATAAQVTAAVALVKAQENARKLDALADSIQATNEAVQDLSTGLQAGAIAIQAIQSEINHVINPALERLSCEIIDTRVASILNLYLIRLTTVFHRQLVNPALTPLSIQALNHLLQGETEGLVKNESKMTDSKIDLLMSGLITGQVVGVNIKHMQLMIAVFVPTTAQLPNAYVINLLTITANINNSEVLVQLPNQILERSGIIYQFRGKDCVSSPNHMYCPYSDASILSPELQLCLQGRLEMCLFTQVVGSFPTRFASDKGIVYANCRHLQCACSEPEGIIYQDDTSAITQIDASKCSTLKLDMLTFKLSTYANKTFDASFSVGKDQMLVTNLLDLSAELKTMNASVAHANKLIDKSNLLIQSNALIGHSNTIFIVVIVILAVMVLYLIIVTYIIKVIMVEVSRLKRMNIYSIDK	2	48

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_025410	4958-6751	gb:NC_025410:4958-6751|Organismo:vírus Tuhoko 1|Nome da cepa:NC_025410 DESCONHECIDA|Nome da proteína:proteína de fusão|Símbolo do gene:F	MVTIIKPLILLVTVILQISGHIDTTALTSIGAVIASSKEIMYYAQSTPNYIVIKLIPNLPNIPSQCNFSSIAYYNKTLLDLFTPISDNINMLHQRLSNTGRNRRFAGVAIGLAALGVATAAQVTAAFALVEAKSNTAKIAQIGQAIQNTNAAINSLNAGIGGAVTAIQAIQTQINGIITDQINAATCTALDAQIGTLLNMYLLQLTTTFQPQIQNPALQPLSIQALHRIMQGTSIVLSNLTDSSKYGLNDALSAGLITGQIVSVDLRLMQITIAANVPTLSRLENAIAHDIMRITTNVNNTEVIVQLPETIMEHAGRLYQFNKDHCLSSTQRFFCPYSDAKLLTSKISSCLSGIRGDCIFSPVVGNFATRFISVKGVIIANCKFIRCTCLQPEGIISQLDDHTLTVIDLKLCNKLDLGLIQFDLQVLSNISYEMTLNTSQNQLILTDPLDLSSELQTMNQSINNAANFIEKSNSLLNSSTYEFNRSVALLVALILLSLTILYVIVLTCVVKLLVHEVSKNRRHIQDLESHHK	2	49
NC_028249	4850-7055	gb:NC_028249:4850-7055|Organismo:vírus da cinomose focina|Nome da cepa:PDV/Wadden_Sea.NLD/1988|Nome da proteína:proteína de fusão|Símbolo do gene:F	MTRVKKLPVPTNPPMHHSLDSPFLNPEHATGKISITDDTSSQLTNFLYHKYHKTTINHLSRTISGTDPPSAKLNKFGSPILSTYQIRSALWWIAMVILVHCVMGQIHWTNLSTIGIIGTDSSHYKIMTRSSHQYLVLKLMPNVSIIDNCTKAELDEYEKLLNSVLEPINQALTLMTKNVKSLQSLGSGRRQRRFAGVVIAGAALGVATAAQITAGVALYQSNLNAQAIQSLRASLEQSNKAIDEVRQASQNIIIAVQGVQDYVNNEIVPALQHMSCELIGQRLGLKLLRYYTELLSVFGPSLRDPVSAEISIQALSYALGGEIHKILEKLGYSGNDMVAILETKGIRAKITHVDLSGKFIVLSISYPTLSEVKGVVVHRLEAVSYNIGSQEWYTTVPRYVATNGYLISNFDESSCVFVSESAICSQNSLYPMSPILQQCLRGETASCARTLVSGTLGNKFILSKGNIIANCASILCKCHSTSKIINQSPDKLLTFIASDTCSLVEIDGVTIQVGSRQYPDVVYASKVILGPAISLERLDVGTNLGSALKKLNDAKVLIESSDQILDTVKNSYLSLGTLIALPVSIGLGLILLLLICCCKKRYQHLFSQSTKVAPVFKPDLTGTSKSYVRSL	2	50

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
NC_028362	5217-6842	gb:NC_028362:5217-6842|Organismo:vírus da parainfluenza caprina 3|Nome da cepa:JS2013|Nome da proteína:proteína de fusão|Símbolo do gene:F	MIKKIICIFSMPILLSFCQVDIIKLQRVGILVSKPKSIKISQNFETRYLVLNLIPNIENAQSCGDQQIKQYKKLLDRLIIPLYDGLRLQQDIIVVDNNLKNNTNHRAKRFFGEIIGTIALGVATSAQITAAVALVEAKQARSDIERVKNAVRDTNKAVQSIQGSVGNLIVAVKSVQDYVNNEIVPSIKRLGCEAAGLQLGIALTQHYSELTNIFGDNIGTLKEKGIKLQGIASLYHTNITEIFTTSTVDQYDIYDLLFTESIKMRVIDVDLNDYSITLQVRLPLLTKISDAQIYNVDSVSYNIGGTEWYIPLPRNIMTKGAFLGGANLQDCIESFSDYICPSDPGFILNRDIENCLSGNITQCPKTLVISDIVPRYAFVDGGVIANCLSTTCTCNGIDNRINQAPDQGIKIITYKDCQTIGINGMLFKTNQEGTLAAYTPVDITLNNSVNLDPIDLSIELNRARSDLAESKEWIKRSEAKLDSVGSWYQSSTTEIIQIVMIIVLFIINIIVLIVLIKYSRSQNQSMNNHMNEPYILTNKVQ	2	51
AF079780	5919-7580	gb:AF079780|Organismo:Paramixovírus da Tupaia|Nome da Cepa:UNKNOWN-AF079780|Nome da Proteína:Proteína de fusão|Símbolo do Gene:F	MASLLKTICYIYLITYAKLEPTPKSQLDLDSLASIGVVDAGKYNYKLMTTGSEKLMVIKLVPNITYATNCNLTAHTAYTKMIERLLTPINQSLYEMRSVITERDGGTIFWGAIIAGAALGVATAAAITAGVALHRAEQNARNIAALKDALRNSNEAIQHLKDAQGHTVLAIQGLQEQINNNIIPKLKESHCLGVNNQLGLLLNQYYSEILTVFGPNLQNPVSASLTIQAIAKAFNGDFNSLMTNLNYDPTDLLDILESNSINGRIIDVNLNEKYIALSIEIPNFITLTDAKIQTFNRITYGYGSNEWLTLIPDNILEYGNLISNVDLTSCVKTKSSYICNQDTSYPISSELTRCLRGDTSSCPRTPVVNSRAPTFALSGGHIYANCAKAACRCEKPPMAIVQPATSTLTFLTEKECQEVVIDQINIQLAPNRLNKTIITDGIDLGPEVIINPIDVSAELGNIELEMDKTQKALDRSNKILDSMITEVTPDKLLIAMIVVFGILLLWLFGVSYYAFKIWSKLHFLDSYVYSLRNPSHHRSNGHQNHSFSTDISG	1	52

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
EU403085	4664-6585	gb:EU403085:4664-6585|Organismo:Paramixovírus aviário 3|Nome da cepa:APMV3/PKT/Netherland/449/75|Nome da proteína:proteína de fusão|Símbolo do gene:F	MQPGSALHLPHLYIIIALVSDGTLGQTAKIDRLIQAGIVLGSGKEL HISQDSGTLDLFVRLLPVLPSNLSHCQLEAITQYNKTVTRLLAPI GKNLEQVLQARPRGRLFGPIIGSIALGVATSAQITAAIALVRAQQ NANDILALKNALQSSNEAIRQLTYGQDKQLLAISKIQKAVNEQILP ALDQLDCAVLGTKLAVQLNLYLIEMTTIFGEQINNPVLATIPLSYIL RLTGAELNNVLMKQARSSLSLVQLVSKGLLSGQVIGYDPSVQG LIIRVNLMRTQKIDRALVYQPYVLPITLNSNIVTPIAPECVIQKGTII EGMSRKDCTELEQDIICRTVTTYTLARDTRLCLQGNISSCRYQQ SGTQLHTPFITYNGAVIANCDLVSCRCLRPPMIITQVKGYPLTIIT RSVCQELSVDNLVLNIETHHNFSLNPTIIDPLTRVIATTPLEIDSLI QEAQDHANAALAKVEESDKYLRAVTGGNYSNWYIVLVIVLLFGN LGWSLLLTVLLCRSRKQQRRYQQDDSVGSERGVGVGTIQYMS	1	53
KX258200	4443-6068	gb:KX258200:4443-6068|Organismo:paramixovírus aviário 14|Nome da cepa:APMV14/pato/Japão/11OG0352/2011|Nome da proteína:proteína de fusão|Símbolo do gene:F	MEKGTVLFLAALTLYNVKALDNTKLLGAGIASGKEHELKIYQSSV NGYIAVKLIPFLPSTKRECYNEQLKNYNATINRLMGPINDNIKLVL SGVKTRTREGKLIGAIIGTAALGLATAAQVTAAIALEQAQDNARAI LTLKESIRNTNNAVSELKTGLSEVSIALSKTQDYINTQIMPALSNL SCEIVGLKIGIQLSQYLTEVTAVFGNQITNPALQPLSMQALYQLC GGDFSLLLDKIGADRNELESLYEANLVTGRIVQYDTADQLVIIQV SIPSVSTLSGYRVTELQSISVDMDHGEGKAVIPRYIVTSGRVIEE MDISPCVLTATAVYCNRLLTTSLPESVLKCLDGDHSSCTYTSNS GVLETRYIAFDGMLIANCRSIVCKCLDPPYIIPQNKGKPLTIISKEV CKKVTLDGITLLIDAEFTGEYGLNITIGPDQFAPSGALDISTELGK LNNSINKAEDYIDKSNELLNRVNVDIVNDTAVIVLCVMSALVVVW CIGLTVGLIYVSKNTLRAVAIKGTSIENPYVSSGKHAKNSS	1	54

ID de Genbank	Nucleotídeos de CDS	Nome Completo do Gene	Sequência	n° de Sequências/ Agrupamento	SEQ ID NO
KY511044	4592-6247	gb:KY511044:4592-6247|Organismo:Paramixovírus aviário UPO216|Nome da cepa:APMV-15/WB/Kr/UPO216/2014|Nome da proteína:proteína de fusão|Símbolo do gene:F	MIFTMYHVTVLLLLSLLTLPLGIQLARASIDGRQLAAAGIVVTGEK AINLYTSSQTGTIVVKLLPNVPQGREACMRDPLTSYNKTLTSLLS PLGEAIRRIHESTTETAGLVQARLVGAIIGSVALGVATSAQITAAA ALIQANKNAENILKLKQSIAATNEAVHEVTDGLSQLAVAVGKMQ DFINTQFNNTAQEIDCIRISQQLGVELNLYLTELTTVFGPQITSPA LSPLSIQALYNLAGGNLDVLLSKIGVGNNQLSALISSGLISGSPIL YDSQTQLLGIQVTLPSVSSLNNMRAIFLETLSVSTDKGFAAALIP KVVTTVGTVTEELDTSYCIETDIDLFCTRIVTFPMSPGIYACLNGN TSECMYSKTQGALTTPYMSVKGSIVANCKMTTCRCADPASIISQ NYGEAVSLIDSSVCRVITLDGVTLRLSGSFDSTYQKNITIRDSQVI ITGSLDISTELGNVNNSINNALDKIEESNQILESVNVSLTSTNALIV YIICTALALICGITGLILSCYIMYKMRSQQKTLMWLGNNTLDQMR AQTKM	1	55
NC_025360	6104-8123	gb:NC_025360:6104-8123|Organismo:Paramixovírus do salmão do Atlântico|Nome da cepa:ASPV/Yrkje371/95|Nome da proteína:proteína de fusão|Símbolo do gene:F	MDGPKFRFVLLILLTAPARGQVDYDKLLKVGIFEKGTANLKISVS SQQRYMVIKMMPNLGPMNQCGIKEVNLYKESILRLITPISTTLNYI KSEIQVEREVALQPNGTIVRFFGLIVAAGALTLATSAQITAGIALH NSLENAKAIKGLTDAIKESNLAIQKIQDATAGTVIALNALQDQVNT NIIPAINTLGCTAAGNTLGIALTRYYSELIMIFGPSLGNPVEAPLTI QALAGAFNGDLHGMIREYGYTPSDIEDILRTNSVTGRVIDVDLV GMNIVLEINLPTLYTLRDTKIVNLGKITYNVDGSEWQTLVPEWLAI RNTLMGGVDLSRCVVSSRDLICKQDPVFSLDTSIISCLNGNTES CPRNRVVNSVAPRYAVIRGNILANCISTTCLCGDPGVPIIQKGDN TLTAMSINDCKLVGVDGYVFRPGPKAVNVTFNLPHLNLGPEVN VNPVDISGALGKVEQDLASSRDHLAKSEKILSGINPNIINTEMVLV AVILSLVCAMVVIGIVCWLSILTKWVRSCRADCRRPNKGPDLGPI MSSQDNLSF	1	56

[00373] Em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequênais de aminoácidos da Tabela 2 ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência à mesma ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% identidade de sequência a uma porção da sequência, por exemplo, uma porção de 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento. Por exemplo, em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos que tem pelo menos 80% de identidade a quaisquer sequências de aminoácidos da Tabela 2. Em algumas modalidades, uma sequência de ácidos nucleicos descrita no presente documento codifica uma sequência de aminoácidos da Tabela 2 ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98%, ou 99% identidade de sequência à mesma, ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% identidade de sequência em relação a uma porção da sequência, por exemplo, uma porção de 40, 50, 60, 80, 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento.

[00374] Em algumas modalidades, um fusogênio descrito no presente documento compreende uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 57 a 132, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com a mesma, ou uma sequência de aminoácidos que tem pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com uma porção da sequência, por exemplo, uma porção de 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento. Por exemplo, em algumas modalidades, um fusogênio descrito no presente documento compreende

uma sequência de aminoácidos com pelo menos 80% de identidade com relação a uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 57 a 132. Em algumas modalidades, uma sequência de ácidos nucleicos descrita no presente documento codifica uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 57 a 132, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com a mesma, ou uma sequência de aminoácidos com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com uma porção da sequência, por exemplo, uma porção de 40, 50, 60, 80, 100, 200, 300, 400, 500 ou 600 aminoácidos de comprimento.

TABELA 2. Agrupamentos de sequências de proteína g, h e hn de paramixovírus. Coluna 1, o id do genbank inclui o id de genbank de toda a sequência do genoma do vírus que é a sequência centroide do agrupamento. Coluna 2, os nucleotídeos de cds fornecem os nucleotídeos correspondentes ao cds do gene em todo o genoma. Coluna 3, nome completo do gene, fornece o nome completo do gene, incluindo id de genbank, espécie de vírus, cepa e nome da proteína. Coluna 4, sequência, fornece a sequência de aminoácidos do gene. A coluna 5, nº de sequências/agrupamento, fornece o número de sequências que se agrupam com essa sequência centroide.

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KU950686	4643 a 5638	gb:KU950686:4643-5638|Organismo:Vírus sincicial respiratório humano|Nome da cepa:RSVA/Homo sapiens/USA/TH_10506/ 2014|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MSKTKDQRTAKTLERTWDTLNHLLFISSCLYKLNLKSIAQITLSILAMIISTSLIIAAIIFIASANHKVTLTTAIIQDATNQIKNTTPTYLTQNPQLGISFSNLSGTTLQSTTILASTTPSAESTPQSTTVKIINTTTTQILPSKPTTKQRQNKPQNKPNNDFHFEVFNFVPCSICSNNPTCWAICKRIPNKKPGKKTTTKPTKKPTLKTTKKDPKPQTTKPKEALTTKPTGKPTINTTKTNIRTTLLTSNTKGNPEHTSQEETLHSTTSEGYLSPSQVYTTSGQEETLHSTTSEGYLSPSQVYTTSEYLSQSLSSSNTTK	706	57
AB524405	6424 a 8274	gb:AB524405:6424-8274|Organismo:vírus da doença de Newcastle|Nome da Cepa:Ganso/ Alasca/415/91|Nome da Proteína:proteína hemaglutinina-neuraminidase|Símbolo do Gene:HN	MERGVSQVALENDEREAKNTWRLVFRVTVLFLTIVTLAISAAALAFSMNASTPQDLEGIPVAISKVEDKITSALGASQDVMDRIYKQVALESPLALLNTESTIMNALTSLSYQINGAANASGCGAPVPDPDYIGGIGKELIVDDTSDVTSFYPSAFQEHLNFIPAPTTGSGCTRIPSFDMSATHYCYTHNVILSGCRDHSHSHQYLALGVLRTSATGRVFFSTLRSINLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYQSTDPTLMVHGRLGFDGQYHERDLDVHTLFGDWVANYPGVGGGSFINNRVWFPVYGGLKPGSPTDKRQEGQYAIYKRYNDTCPDDQEYQVRMAKSAYKPNRFGGKRVQQAILSIGVSTTLADDPVLTVTSNTITLMGAEGRVMTVGTSHYLYQRGSSYYSPAILYPLTIANKTATLQDPYKFNAFTRPGSVPCQASARCPNSCVTGVYTDPYPIVFHKNHTLRGVFGTMLDDEQARLNPVSAVFDSIARSRVTRVSSSSTKAAYTTSTCFKVVKTGKVYCLSIAEISNTLFGEFRIVPLLVEILRDEGRSEARSALTTQGHPGWNDEVVDPIFCAVTNQTDHRQKLEEYAQSWP	418	58
JQ582844	4686-5636	gb:JQ582844:4686-5636|Organismo:Vírus sincicial respiratório humano|Nome da cepa:NH1067|Nome da proteína:glicoproteína de ligação ao receptor|Símbolo do gene:G	MSKNKNQRTARTLEKTWDTLNHLIVISSCLYKLNLKSIAQIALSVLAMIISTSLIIAAIIFIISANHKVTLTTVTVQTIKNHTEKNITTYLTQVSPERVSPSKQPTTTPPIHTNSATISPNTKSETHHTTAQTKGRTTTPTQNNKPSTKPRPKNPPKKPKDDYHFEVFNFVPCSICGNNQLCKSICKTIPNNKPKKKPTTKPTNKPPTKTTNKRDPKTPAKTLKKETTINPTTKKPTPKTTERDTSTPQSTVLDTTTSKHTERDTSTPQSTVLDTTTSKHTIQQQSLHSITPENTPNSTQTPTASEPSTSNSTQKL	278	59

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB254456	7271-9136	gb:AB254456:7271-9136|Organismo:Vírus do Saram-po|Nome da Cepa:SSPE-Kobe-1|Nome da Proteína:Hemaglutinina|Símbolo do Gene:H	MSPHRDRINAFYRDNPHPKGSRIVINREHLMIDRPYVLLAVLFVMFLS LIGLLAIAGIRLHRAAIYTAEIHKSLSTNLDVTNSIEHQVKDVLTPLFKIIG DEVGLRTPQRFTDLVKFISDKIKFLNPDREYDFRDLTWCINPPERIKLD YDQYCADVAAEELMNALVNSTLLEARATNQFLAVSKGNCSGPTTIRG QFSNMSLSLLDLYLSRGYNVSSIVTMTSQGMYGGTYLVGKPNLSSKG SELSQLSMHRVFEVGVIRNPGLGAPVFHMTNYFEQPVSNDFSNCMV ALGELRFAALCHREDSVTVPYQGSGKGVSFQLVKLGVWKSPTDMQS WVPLSTDDPVIDRLYLSSHRGVIADNQAKWAVPTTRTDDKLRMETCF QQACKGKNQALCENPEWAPLKDNRIPSYGVLSVNLSLTVELKIKIASG FGPLITHGSGMDLYKTNHDNVYWLTIPPMKNLALGVINTLEWIPRFKV SPNLFTVPIKEAGEDCHAPTYLPAEVDGDVKLSSNLVILPGQDLQYVL ATYDTSRVEHAVVYYVYSPSRSFSYFYPFRLPIKGVPIELQVECFTWD QKLWCRHFCVLADSESGGHITHSGMVGMGVSCTVTREDGTNRRQG CQ	128	60
AB040874	6614-8362	gb:AB040874:6614-8362|Organismo:Vírus da Cachum-ba|Nome da Cepa:Miyahara|Nome da Proteína:hemaglutinina-neuraminidase|Símbolo do Gene:HN	MEPSKLFTMSDNATFAPGPVINAADKKTFRTCFRILVLSVQAVTLILVI VTLGELVRMINDQGLSNQLSSIADKIRESATMIASAVGVMNQVIHGVT VSLPLQIEGNQNQLLSTLATICTGKKQVSNCSTNIPLVNDLRFINGINK FIIEDYATHDFSIGHPLNMPSFIPTATSPNGCTRIPSFSLGKTHWCYTH NVINANCKDHTSSNQYISMGILVQTASGYPMFKTLKIQYLSDGLNRKS CSIATVPDGCAMYCYVSTQLETDDYAGSSPPTQKLTLLFYNDTVTER TISPTGLEGNWATLVPGVGSGIYFENKLIFPAYGGVLPNSSLGVKSAR EFFRPVNPYNPCSGPQQDLDQRALRSYFPSYFSNRRVQSAFLVCAW NQILVTNCELVVPSNNQTLMGAEGRVLLINNRLLYYQRSTSWWPYEL LYEISFTFTNSGQSSVNMSWIPIYSFTRPGSGNCSGENVCPTACVSG VYLDPWPLTPYSHQSGINRNFYFTGALLNSSTTRVNPTLYVSALNNLK VLAPYGNQGLFASYTTTTCFQDTGDASVYCVYIMELASNIVGEFQILP VLTRLTIT	87	61

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB736166	6709-8427	gb:AB736166:6709-8427|Organismo:Respirovírus 3 Humano|Nome da Cepa:ZMLS/2011|Nome da Proteína:hemaglutinina-neuraminidase|Símbolo do Gene:HN	MEYWKHTNHGKDAGNELETATATHGNRLTNKITYILWTITLVLLSIVFII VLINSIKSEKAHESLLQDINNEFMEVTEKIQVASDNTNDLIQSGVNTRL LTIQSHVQNYIPISLTQQISDLRKFISEITIRNDNQEVPPQRITHDVGIKP LNPDDFWRCTSGLPSLMRTPKIRLMPGPGLLAMPTTVDGCVRTPSLV INDLIYAYTSNLITRGCQDIGKSYQVLQIGIITVNSDLVPDLNPRISHTFN INDNRKSCSLALLNTDVYQLCSTPKVDERSDYASSGIEDIVLDIVNYDG SISTTRFKNNNISFDQPYAALYPSVGPGIYYKGKIIFLGYGGLEHPINEN AICNTTGCPGKTQRDCNQASHSPWFSDRRMVNSIIVVDKGLNSVPKL KVWTISMRQNYWGSEGRLLLLGNKIYIYTRSTSWHSKLQLGIIDITDYS DIRIKWTWHNVLSRPGNNECPWGHSCPDGCITGVYTDAYPLNPTGSI VSSVILDSQKSRVNPVITYSTATERVNELAIRNKTLSAGYTTTSCITHY NKGYCFHIVEINHKSLNTFQPMLFKTEIPKSCS	78	62
KJ627396	6166-6885	gb:KJ627396:6166-6885|Organismo:metapneumovírus humano|Nome da cepa:HMPV/Homo sapiens/PER/FLI1305/2010/A|Nome da proteína:glicoproteína G de fixação|Símbolo do gene:G	MEVKVENIRAIDMLKARVKNRVARSKCFKNASLILIGITTLSIALNIYLIIN YTIQKTTSESEHHTSSPPTESNKETSTIPIDNPDITPNSQHPTQQSTES LTLYPASSMSPSETEPASTPGITNRLSLADRSTTQPSESRTKTNSTVH KKNKKNISSTISRTQSPPRTTAKAVSRTTALRMSSTGERPTTTSVQSD SSTTAQNHEETGPANPQASVSTM	71	63

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB475097	7079-8902	gb:AB475097:7079-8902|Organismo:vírus da Cinomose caninca|Nome da Cepa:M25CR|Nome da Proteína:hemaglutinina|Símbolo do Gene:H	MLSYQDKVGAFYKDNARANSSKLSLVTEEQGGRRPPYLLFVLLILLV GILALLAIAGVRFRQVSTSNVEFGRLLKDDLEKSEAVHHQVMDVLTPL FKIIGDEIGLRLPQKLNEIKQFILQKTNFFNPNREFDFRDLHWCINPPSK IKVNFTNYCDAIGVRKSIASAANPILLSALSGGRGDIFPPYRCSGATTS VGRVFPLSVSLSMSLISKTSEIISMLTAISDGVYGKTYLLVPDYIEREFD TQKIRVFEIGFIKRWLNDMPLLQTTNYMVLPENSKAKVCTIAVGELTLA SLCVDESTVLLYHDSNGSQDSILVVTLGIFGATPMNQVEEVIPVAHPS VERIHITNHRGFIKDSVATWMVPALVSEQQEGQKNCLESACQRKSYP MCNQTSWEPFGGVQLPSYGRLTLPLDASIDLQLNISFTYGPVILNGD GMDYYENPLLDSGWLTIPPKNGTILGLINKASRGDQFTVTPHVLTFAP RESSGNCYLPIQTSQIMDKDVLTESNLVVLPTQNFRYVVATYDISREN HAIVYYVYDPIRTISYTYPFRLTTKGRPDFLRIECFVWDDDLWCHQFY RFESDITNSTTSVEDLVRIRFSCNRSKP	45	64
AJ849636	7326-9155	gb:AJ849636:7326-9155|Organismo: vírus da Peste des petits ruminants|Nome da Cepa:Turkey 2000|Nome da Proteína:haemaglutinina|Símbolo do Gene:H	MSAQRERINAFYKDNPHNKNHRVILDRERLVIERPYILLGVLLVMFLSL IGLLAIAGIRLHRATVGTSEIQSRLNTNIELTESIDHQTKDVLTPLFKIIGD EVGIRIPQKFSDLVKFISDKIKFLNPDREYDFRDLRWCMNPPERVKINF DQFCEYKAAVKSIEHIFESPLNKSKKLQSLTLGPGTGCLGRTVTRAHF SELTLTLMDLDLEMKHNVSSVFTVVEEGLFGRTYTVWRSDARDPSTD LGIGHFLRVFEIGLVRDLGLGPPVFHMTNYLTVNMSDDYRRCLLAVG ELKLTALCSSSETVTLGERGVPKREPLVVVILNLAGPTLGGELYSVLP TSDLMVEKLYLSSHRGIIKDDEANWVVPSTDVRDLQNKGECLVEACK TRPPSFCNGTGSGPWSEGRIPAYGVIRVSLDLASDPGVVITSVFGPLI PHLSGMDLYNNPFSRAVWLAVPPYEQSFLGMINTIGFPNRAEVMPHI LTTEIRGPRGRCHVPIELSRRVDDDIKIGSNMVILPTIDLRYITATYDVS RSEHAIVYYIYDTGRSSSYFYPVRLNFKGNPLSLRIECFPWRHKVWC YHDCLIYNTITDEEVHTRGLTGIEVTCNPV	34	65

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB005795	6693-8420	gb:AB005795:6693-8420|Organismo:Vírus Sendai|Nome da Cepa:Ohita|Nome da Proteína:proteína hemaglutinina-neuraminidase|Símbolo do Gene:HN	MDGDRSKRDSYWSTSPGGSTTKLVSDSERSGKVDTWLLILAFTQWALSIATVIICIVIAARQGYSMERYSMTVEALNTSNKEVKESLTSLIRQEVITRAANIQSSVQTGIPVLLNKNSRDVIRLIEKSCNRQELTQLCDSTIAVHHAEGIAPLEPHSFWRCPAGEPYLSSDPEVSLLPGPSLLSGSTTISGCVRLPSLSIGEAIYAYSSNLITQGCADIGKSYQVLQLGYISLNSDMFPDLNPVVSHTYDINDNRKSCSVVATGTRGYQLCSMPIVDERTDYSSDGIEDLVLDILDLKGRTKSHRYSNSEIDLDHPFSALYPSVGSGIATEGSLIFLGYGGLTTPLQGDTKCRIQGCQQVSQDTCNEALKITWLGGKQVVSVLIQVNDYLSERPRIRVTTIPITQNYLGAEGRLLKLGDQVYIYTRSSGWHSQLQIGVLDVSHPLTISWTPHEALSRPGNEDCNWYNTCPKECISGVYTDAYPLSPDAANVATVTLYANTSRVNPTIMYSNTTNIINMLRIKDVQLEAAYTTTSCITHFGKGYCFHIIEINQKSLNTLQPMLFKTSIPKLCKAES	23	66
AF457102	6903-8630	gb:AF457102|Organismo:Cepa do vírus 1 de parainfluenza humana Washington/1964|Nome da Cepa:Washington 1964|Nome da Proteína:Glicoproteína HN|Símbolo do Gene:HN	MAEKGKTNSSYWSTTRNDNSTVNTHINTPAGRTHIWLLIATTMHTVLSFIIMILCIDLIIKQDTCMKTNIMTVSSMNESAKIIKETITELIRQEVISRTINIQSSVQSGIPILLNKQSRDLTQLIEKSCNRQELAQICENTIAIHHADGISPLDPHDFWRCPVGEPLLSNNPNISLLPGPSLLSGSTTISGCVRLPSLSIGDAIYAYSSNLITQGCADIGKSYQVLQLGYISLNSDMYPDLNPVISHTYDINDNRKSCSVIAAGTRGYQLCSLPTVNETTDYSSEGIEDLVFDILDLKGKTKSHRYKNEDITFDHPFSAMYPSVGSGIKIENTLIFLGYGGLTTPLQGDTKCVINRCTNVNQSVCNDALKITWLKKRQVVNVLIRINNYLSDRPKIVVETIPITQNYLGAEGRLLKLGKKIYIYTRSSGWHSNLQIGSLDINNPMTIKWAPHEVLSRPGNQDCNWYNRCPRECISGVYTDAYPLSPDAVNVATTTLYANTSRVNPTIMYSNTSEIINMLRLKNVQLEAAYTTTSCITHFGKGYCFHIVEINQASLNTLQPMLFKTSIPKICKITS	21	67

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KJ627397	6146-6888	gb:KJ627397:6146-6888|Organismo:Metapneumovírus humano|Nome da cepa:HMPV/Homo sapiens/PER/FPP00098/2010/B|Nome da proteína:glicoproteína G de fixação|Símbolo do gene:G	MEVRVENIRAIDMFKAKMKNRIRSSKCYRNATLILIGLTALSMALNIFLII DYATLKNMTKVEHCVNMPPVEPSKKSPMTSAADLNTKLNPQQATQL TTEDSTSLAATSENHLHTETTPTSDATISQQATDEHTTLLRPINRQTT QTTTEKKPTGATTKKDKEKETTTRTTSTAATQTLNTTNQTSNGREAT TTSARSRNGATTQNSDQTIQAADPSSKPYHTQTNTTTAHNTDTSSLS S	21	68
AF017149	8913-10727	gb:AF017149|Organismo:Hendra vírus|Nome da Cepa:UNKNOWN-AF017149|Nome da Proteína:glicoproteína|Símbolo do Gene:G	MMADSKLVSLNNNLSGKIKDQGKVIKNYYGTMDIKKINDGLLDSKILG AFNTVIALLGSIIIIVMNIMIIQNYTRTTDNQALIKESLQSVQQQIKALTDK IGTEIGPKVSLIDTSSTITIPANIGLLGSKISQSTSSINENVNDKCKFTLP PLKIHECNISCPNPLPFREYRPISQGVSDLVGLPNQICLQKTTSTILKP RLISYTLPINTREGVCITDPLLAVDNGFFAYSHLEKIGSCTRGIAKQRII GVGEVLDRGDKVPSMFMTNVWTPPNPSTIHHCSSTYHEDFYYTLCA VSHVGDPILNSTSWTESLSLIRLAVRPKSDSGDYNQKYIAITKVERGK YDKVMPYGPSGIKQGDTLYFPAVGFLPRTEFQYNDSNCPIIHCKYSK AENCRLSMGVNSKSHYILRSGLLKYNLSLGGDIILQFIEIADNRLTIGSP SKIYNSLGQPVFYQASYSWDTMIKLGDVDTVDPLRVQWRNNSVISRP GQSQCPRFNVCPEVCWEGTYNDAFLIDRLNWVSAGVYLNSNQTAE NPVFAVFKDNEILYQVPLAEDDTNAQKTITDCFLLENVIWCISLVEIYDT GDSVIRPKLFAVKIPAQCSES	14	69

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AF212302	8943-10751	gb:AF212302|Organismo:Vírus Nipah|Nome da Cepa:AF212302 DESCONHECIDA|Nome da Proteína:glicoproteína de fixação|Símbolo do Gene:G	MPAENKKVRFENTTSDKGKIPSKVIKSYYGTMDIKKINEGLLDSKILSA FNTVIALLGSIVIIVMNIMIIQNYTRSTDNQAVIKDALQGIQQQIKGLADKI GTEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPP LKIHECNISCPNPLPFREYRPQTEGVSNLVGLPNNICLQKTSNQILKPK LISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGVSKQRII GVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYYVLCA VSTVGDPILNSTYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEK GRYDKVMPYGPSGIKQGDTLYFPAVGFLVRTEFKYNDSNCPITKCQY SKPENCRLSMGIRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSI GSPSKIYDSLGQPVFYQASFSWDTMIKFGDVLTVNPLVVNWRNNTVI SRPGQSQCPRFNTCPEICWEGVYNDAFLIDRINWISAGVFLDSNQTA ENPVFTVFKDNEILYRAQLASEDTNAQKTITNCFLLKNKIWCISLVEIYD TGDNVIRPKLFAVKIPEQCT	14	70
EU439428	6751-8638	gb:EU439428:6751-8638|Organismo:vírus da parainfluenza suíno 3|Nome da cepa:92-7783_ISU-92|Nome da proteína:hemaglutinina-neuraminidase HN|Símbolo do gene:HN	MEYWKHTNSTKDTNNELGTTRDRHSSKATNIIMYIFWTTTSTILSVIFI MILINLIQENNHNKLMLQEIKKEFAVIDTKIQKTSDDISTSIQSGINTRLLT IQSHVQNYIPLSLTQQMSDLRKFINDLTTKREHQEVPIQRMTHDSGIE PLNPDKFWRCTSGNPSLTSSPKIRLIPGPGLLATSTTVNGCIRIPSLAI NNLIYAYTSNLITQGCQDIGKSYQVLQIGIITINSDLVPDLNPRVTHTFNI DDNRKSCSLALLNTDVYQLCSTPKVDERSDYASTGIEDIVLDIVTSNG LIITTRFTNNNITFDKPYAALYPSVGPGIYYKDKVIFLGYGGLEHEENG DVICNTTGCPGKTQRDCNQASYSPWFSNRRMVNSIIVVDKSIDTTFSL RVWTIPMRQNYWGSEGRLLLLGDRIYIYTRSTSWHSKLQLGVIDISDY NNIRINWTWHNVLSRPGNDECPWGHSCPDGCITGVYTDAYPLNPSG SVVSSVILDSQKSRENPIITYSTATNRVNELAIYNRTLPAAYTTTNCITH YDKGYCFHIVEINHRSLNTFQPMLFKTEVPKNCS	14	71

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KF530164	6157-6906	gb:KF530164:6157-6906|Organismo:metapneumovírus humano|Nome da cepa:HMPV/AUS/172832788/2004/B|Nome da proteína:glicoproteína G de fixação|Símbolo do gene:G	MEVRVENIRAIDMFKAKIKNRIRSSRCYRNATLILIGLTALSMALNIFLIIDHATLRNMIKTENCANMPSAEPSKKTPMTSTAGPSTKPNPQQATQWTTENSTSPAATLEGHPYTGTTQTPDTTAPQQTTDKHTALPKSTNEQITQTTTEKKTTRATTQKREKRKENTNQTTSTAATQTTNTTNQTRNASETITTSDGPRIDTTTQSSEQTARATEPGSSPYHARRGAGPR	14	72
AB910309	6960-8747	gb:AB910309:6960-8747|Organismo:morbilivírus felino|Nome da Cepa:SS1|Nome da Proteína:proteína hemaglutinina|Símbolo do Gene:H	MKNINIKYYKDSNRYLGKILDEHKIVNSQLYSLSIKVITIIAIIVSLIATIMTIINATSGRTTLNSNTDILLNQRDEIHSIHEMIFDRVYPLITAMSTELGLHIPTLLDELTKAIDQKIKIMNPPVDTVTSDLSWCIKPPNGIIIDPKGYCESMELSKTYKLLLDQLDVSRKKSLTINRKNINQCQLVDDSEIIFATVNIQSTPRFLNFGHTVSNQRITFGQGTYSSTYILTIQEDGITDVQYRVFEIGYISDQFGVFPSLIVSRVLPIRMVLGMESCTLTSDRQGGYFLCMNTLTRSIYDYVNIRDLKSLYITLPHYGKVNYTYFNFGKIRSPHEIDKLWLTSDRGQIISGYFAAFVTITIRNYNNYPYKCLNNPCFDNSENYCRGWYKNITGTDDVPILAYLLVEMYDEEGPLITLVAIPPYNYTAPSHNSLYYDDKINKLIMTTSHIGYIQINEVHEVIVGDNLKAILLNRLSDEHPNLTACRLNQGIKEQYKSDGMIISNSALIDIQERMYITVKAIPPVGNYNFTVELHSRSNTSYILLPKQFNAKYDKLHLECFNWDKSWWCALIPQFSLSWNESLSVDTAIFNLINCK	12	73

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB759118	7116-8957	gb:AB759118:7116-8957|Organismo:Paramixovírus 6 Aviário|Nome da Cepa:Pilrito-de-pescoço-ruivo/Japan/8KS0813/2008|Nome da Proteína:hemaglutinina-neuraminidase|Símbolo do Gene:HN	MASPSELNRSQATLYEGDPNSKRTWRTVYRASTLILDLAILCVSIVAIV RMSTLTPSDVTDSISSSITSLSDTYQSVWSDTHQKVNSIFKEVGISIPV TLDKMQVEMGTAVNIITDAVRQLQGVNGSAGFSITNSPEYSGGIDALI YPQKSLNGKSLAISDLLEHPSFIPAPTTSHGCTRIPTFHLGYRHWCYS HNTIESGCHDAGESIMYLSMGAVGVGHQGKPVFTTSAAVILDDGKNR KSCSVVANPNGCDVLCSLVKQTEDQDYADPTPTPMIHGRLHFNGTY TESMLDQSLFTGHWVAQYPAVGSGSVSHGRLFFPLYGGISKSSSLFP KLRAHAYFTHNEELECKNLTSKQREDLFNAYMPGKIAGSLWAQGIVIC NLTTLADCKIAVANTSTMMMAAEGRLQLVQDKVVLYQRSSSWWPVLI YYDILVSELVNARHLDIVNWVPYPQSKFPRPTWTKGLCEKPSICPAVC VTGVYQDVWVVSVGDFSNETVVIGGYLEAASERKDPWIAAANQYNW LTRRQLFTAQTEAAYSSTTCFRNTHQDKVFCLTIMEVTDNLLGDWRIA PLLYEVTVVDRQQSSRKAVAMSEAHRTRFKYYSPENKFTPQH	11	74
AY141760	6791-8485	gb:AY141760|Organismo:Pramixovírus Fer-de-Lance |Nome da Cepa:ATCC VR-895|Nome da Proteína:proteína hemaglutinina-neuraminidase HN|Símbolo do Gene:HN	MDPKSYYCNEDLRSDGGEKSPGGDLYKGIILVSTVISLIIAIISLAFIIDN KINIQSLDPLRGLEDSYLVPIKDKSESISQDIQEGIFPRLNLITAATTTTIP RSIAIQTKDLSDLIMNRCYPSVVNNDTSCDVLAGAIHSNLFSQLDPSTY WTCSSGTPTMNQTVKLLPDNSQIPGSTYSTGCVRIPTFSLGSMIYSY SHNVIYEGCNDHSKSSQYWQLGYISTSKTGEPLQQVSRTLTLNNGLN RKSCSTVAQGRGAYLLCTNVVEDERTDYSTEGIQDLTLDYIDIFGAER SYRYTNNEVDLDRPYAALYPSVGSGTVYNDRILFLGYGGLMTPYGD QAMCQAPECTSATQEGCNSNQLIGYFSGRQIVNCIIEIITVGTEKPIIRV RTIPNSQVWLGAEGRIQTLGGVLYLYIRSSGWHALAQTGIILTLDPIRIS WIVNTGYSRPGNGPCSASSRCPAQCITGVYTDIFPLSQNYGYLATVT LLSGVDRVNPVISYGTSTGRVADSQLTSSSQVAAYTTTTCFTFNQKG YCYHIIELSPATLGIFQPVLVVTEIPKICS	8	75

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
EU877976	6248-8161	gb:EU877976:6248-8161|Organismo:Paramixovírus aviário 4|Nome da cepa:APMV-4/KR/YJ/06|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MQGNMEGSRDNLTVDDELKTTWRLAYRVVSLLLMVSALIISIVILTRD NSQSIITAINQSSDADSKWQTGIEGKITSIMTDTLDTRNAALLHIPLQLN TLEANLLSALGGNTGIGPGDLEHCRYPVHDTAYLHGVNRLLINQTADY TAEGPLDHVNFIPAPVTTTGCTRIPSFSVSSSIWCYTHNVIETGCNDH SGSNQYISMGVIKRAGNGLPYFSTVVSKYLTDGLNRKSCSVAAGSGH CYLLCSLVSEPEPDDYVSPDPTPMRLGVLTWDGSYTEQAVPERIFKN IWSANYPGVGSGAIVGNKVLFPFYGGVRNGSTPEVMNRGRYYYIQD PNDYCPDPLQDQILRAEQSYYPTRFGRRMVMQGVLACPVSNNSTIA SQCQSYYFNNSLGFIGAESRIYYLNGNIYLYQRSSSWWPHPQIYLLDS RIASPGTQNIDSGVNLKMLNVTVITRPSSGFCNSQSRCPNDCLFGVY SDIWPLSLTSDSIFAFTMYLQGKTTRIDPAWALFSNHAIGHEARLFNK EVSAAYSTTTCFSDTIQNQVYCLSILEVRSELLGAFKIVPFLYRVL	8	76
AB176531	6821-8536	gb:AB176531:6821-8536|Organismo:vírus 2 da parainfluenza humana|Nome da Cepa:Nishio|Nome da Proteína:proteína hemaglutinina-neuraminidase|Símbolo do Gene:HN	MEDYSNLSLKSIPKRTCRIIFRTATILGICTLIVLCSSILHEIIHLDVSSGL MDSDDSQQGIIQPIIESLKSLIALANQILYNVAIIIPLKIDSIETVIFSALKD MHTGSMSNTNCTPGNLLLHDAAYINGINKFLVLKSYNGTPKYGPLLNI PSFIPSATSPNGCTRIPSFSLIKTHWCYTHNVMLGDCLDFTTSNQYLA MGIIQQSAAAFPIFRTMKTIYLSDGINRKSCSVTAIPGGCVLYCYVATR SEKEDYATTDLAELRLAFYYYNDTFIERVISLPNTTGQWATINPAVGS GIYHLGFILFPVYGGLISGTPSYNKQSSRYFIPKHPNITCAGNSSEQAA AARSSYVIRYHSNRLIQSAVLICPLSDMHTARCNLVMFNNSQVMMGA EGRLYVIDNNLYYYQRSSSWWSASLFYRINTDFSKGIPPIIEAQWVPS YQVPRPGVMPCNATSFCPANCITGVYADVWPLNDPEPTSQNALNPN YRFAGAFLRNESNRTNPTFYTASASALLNTTGFNNTNHKAAYTSSTC FKNTGTQKIYCLIIIEMGSSLLGEFQIIPFLRELIP	7	77

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AF052755	6584-8281	gb:AF052755|Organismo:Vírus Parainfluenza 5|Nome da Cepa:W3A|Nome da Proteína:proteína hemaglutinina-neuraminidase|Símbolo do Gene:HN	MVAEDAPVRATCRVLFRTTTLIFLCTLLALSISILYESLITQKQIMSQAG STGSNSGLGSITDLLNNILSVANQIIYNSAVALPLQLDTLESTLLTAIKSL QTSDKLEQNCSWSAALINDNRYINGINQFYFSIAEGRNLTLGPLLNMP SFIPTATTPEGCTRIPSFSLTKTHWCYTHNVILNGCQDHVSSNQFVSM GIIEPTSAGFPFFRTLKTLYLSDGVNRKSCSISTVPGGCMMYCFVSTQ PERDDYFSAAPPEQRIIIMYYNDTIVERIINPPGVLDVWATLNPGTGSG VYYLGWVLFPIYGGVIKGTSLWNNQANKYFIPQMVAALCSQNQATQV QNAKSSYYSSWFGNRMIQSGILACPLRQDLTNECLVLPFSNDQVLM GAEGRLYMYGDSVYYYQRSNSWWPMTMLYKVTITFTNGQPSAISAQ NVPTQQVPRPGTGDCSATNRCPGFCLTGVYADAWLLTNPSSTSTFG SEATFTGSYLNTATQRINPTMYIANNTQIISSQQFGSSGQEAAYGHTT CFRDTGSVMVYCIYIIELSSSLLGQFQIVPFIRQVTLS	7	78
BK005918	6560-8290	gb:BK005918|Organismo:Rubulavírus suíno|Nome da cepa:UNKNOWN-BK005918|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MSQLGTDQIMHLAQPAIARRTWRLCFRIFALFILIAIVITQIFMLTFDHTL LTTTQFLTSIGNLQSTITSWTPDVQAMLSISNQLIYTTSITLPLKISTTEM SILTAIRDHCHCPDCSSACPTRQMLLNDPRYMSGVNQFIGAPTESINI TFGPLFGIPSFIPTSTTTQGCTRIPSFALGPSHWCYTHNFITAGCADG GHSNQYLAMGTIQSASDGSPLLITARSYYLSDGVNRKSCSIAVVPGG CAMYCYVATRSETDYYAGNSPPQQLLTLVFSNDTIIERTIHPTGLANG WVMLVPGVGSGTLYNEYLLFPAYGGMQQILANQSGEINQFFTPYNAT VRCAMAQPQFSQRAAASYYPRYFSNRWIRSAIVACPYRAIYQTQCTL IPLPNRMVMMGSEGRIFTLGDRLFYYQRSSSWWPYPLLYQVGLNFLT TPPSVSSMTQVPLEHLARPGKGGCPGNSHCPATCVTGVYADVWPLT DPRSGVGGTSLVAAGGLDSTSERMAPVNYLAIGESLLSKTYLLSKTQ PAAYTTTTCFRDTDTGKIYCITIAELGKVLLGEFQIVPFLREIKIQSRY	7	79

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
EU338414	6015-7913	gb:EU338414:6015-7913|Organismo:Paramixovírus aviário 2|Nome da cepa:APMV-2/Chicken/California/Yucaipa/56|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MDFPSRENLAAGDISGRKTWRLLFRILTLSIGVVCLAINIATIAKLDHLD NMASNTWTTTEADRVISSITTPLKVPVNQINDMFRIVALDLPLQMTSL QKEITSQVGFLAESINNVLSKNGSAGLVLVNDPEYAGGIAVSLYQGDA SAGLNFQPISLIEHPSFVPGPTTAKGCIRIPTFHMGPSHWCYSHNIIAS GCQDASHSSMYISLGVLKASQTGSPIFLTTASHLVDDNINRKSCSIVA SKYGCDILCSIVIETENEDYRSDPATSMIIGRLFFNGSYTESKINTGSIF SLFSANYPAVGSGIVVGDEAAFPIYGGVKQNTWLFNQLKDFGYFTHN DVYKCNRTDIQQTILDAYRPPKISGRLWVQGILLCPVSLRPDPGCRLK VFNTSNVMMGAEARLIQVGSTVYLYQRSSSWWVVGLTYKLDVSEITS QTGNTLNHVDPIAHTKFPRPSFRRDACARPNICPAVCVSGVYQDIWPI STATNNSNIVWVGQYLEAFYSRKDPRIGIATQYEWKVTNQLFNSNTE GGYSTTTCFRNTKRDKAYCVVISEYADGVFGSYRIVPQLIEIRTTTGK SE	7	80
KC403973	6234-6964	gb:KC403973:6234-6964|Organismo:metapneumovírus humano|Nome da cepa:HMPV/USA/TN-82-518/1982/A|Nome da proteína:glicoproteína G de fixação|Símbolo do gene:G|Segmento:8	MEVKVENIRTIDMLKARVKNRVARSKCFKNASLILIGITTLSIALNIYLIIN YTMQENTSESEHHTSSSPMESSRETPTVPIDNSDTNPSSQYPTQQS TEGSTLYFAASASSPETEPTSTPDTTSRPPFVDTHTTPPSASRTKTSP AVHTKNNPRISSRTHSPPWAMTRTVRRTTTLRTSSIRKRSSTASVQP DSSATTHKHEEASPVSPQTSASTTRPQRKSMEASTSTTYNQTS	6	81

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KF015281	4511-5844	gb:KF015281:4511-5844|Organismo:Pneumovírus canino|Nome da cepa:cachorro/Bari/100-12/ITA/2012|Nome da proteína:proteína de fixação|Símbolo do gene:G	MRPAEQLIQENYKLTSLSMGRNFEVSGSTTNLNFERTQYPDTFRAVV KVNQMCKLIAGVLTSAAVAVCVGVIMYSVFTSNHKANSMQNATIRNS TSAPPQPTAGPPTTEQGTTPKFTKPPTKTTTHHEITEPAKMVTPSEDP YQCSSNGYLDRPDLPEDFKLVLDVICKPPGPEHHSTNCYEKREINLG SVCPDLVTMKANMGLNNGGGEEAAPYIEVITLSTYSNKRAMCVHNG CDQGFCFFLSGLSTDQKRAVLELGGQQAIMELHYDSYWKHYWSNS NCVVPRTNCNLTDQTVILFPSFNNKNQSQCTTCADSAGLDNKFYLTC DGLSRNLPLVGLPSLSPQAHKAALKQSTGTTTAPTPETRNPTPAPRR SKPLSRKKRALCGVDSSREPKPTMPYWCPMLQLFPRRSNS	6	82
KF973339	4624-5310	gb:KF973339:4624-5310|Organismo:Vírus sincicial respiratório tipo A|Nome da cepa:RSV-A/US/BID-V7358/2002|Nome da proteína:glicoproteína de fixação truncada|Símbolo do gene:G	MSKTKDQRAAKTLEKTWDTLNHLLFISSCLYKSNLKSIAQITLSILAMTI PTSLIIVATTFIASANNKVTPTTAIIQDATSQIKNTTPTHLTQNPQPGISF FNLSGTISQTTAILAPTTPSVEPILQSTTVKTKNTTTTQIQPSKLTTKQR QNKPPNKPNDDFHFEVFNFVPCSICSNNPTCWAICKRIPSKKPGKKT TTKPTKKQTIKTTKKDLKPQTTKPKEAPTT	6	83
FJ215864	6383-8116	gb:FJ215864:6383-8116|Organismo:Paramixovírus aviário 8|Nome da cepa:pintail/Wakuya/20/78|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MSNIASSLENIVEQDSRKTTWRAIFRWSVLLITTGCLALSIVSIVQIGNL KIPSVGDLADEVVTPLKTTLSDTLRNPINQINDIFRIVALDIPLQVTSIQK DLASQFSMLIDSLNAIKLGNGTNLIIPTSDKEYAGGIGNPVFTVDAGGS IGFKQFSLIEHPSFIAGPTTTRGCTRIPTFHMSESHWCYSHNIIAAGCQ DASASSMYISMGVLHVSSSGTPIFLTTASELIDDGVNRKSCSIVATQF GCDILCSIVIEKEGDDYWSDTPTPMRHGRFSFNGSFVETELPVSSMF SSFSANYPAVGSGEIVKDRILFPIYGGIKQTSPEFTELVKYGLFVSTPT TVCQSSWTYDQVKAAYRPDYISGRFWAQVILSCALDAVDLSSCIVKIM NSSTVMMAAEGRIIKIGIDYFYYQRSSSWWPLAFVTKLDPQELADTNS IWLTNSIPIPQSKFPRPSYSENYCTKPAVCPATCVTGVYSDIWPLTSS SSLPSIIWIGQYLDAPVGRTYPRFGIANQSHWYLQEDILPTSTASAYST TTCFKNTARNRVFCVTIAEFADGLFGEYRITPQLYELVRNN	5	84

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
JX857409	6619-8542	gb:JX857409:6619-8542|Organismo:vírus da parainfluenza suína 1|Nome da cepa:S206N|Nome da proteína:proteína hemaglutinina|Símbolo do gene:H	MEETKVKTSEYWARSPQIHATNHPNVQNREKIKEILTILISFISSLSLVL VIAVLIMQSLHNGTILRCKDVGLESINKSTYSISNAILDVIKQELITRIINT QSSVQVALPILINKKIQDLSLIIEKSSKVHQNSPTCSGVAALTHVEGIKP LDPDDYWRCPSGEPYLEDELTLSLIPGPSMLAGTSTIDGCVRLPSLAI GKSLYAYSSNLITKGCQDIGKSYQVLQLGIITLNSDLHPDLNPIISHTYD INDNRKSCSVAVSETKGYQLCSMPRVNEKTDYTSDGIEDIVFDVLDLK GSSRSFKFSNNDINFDHPFSALYPSVGSGIIWKNELYFLGYGALTTAL QGNTKCNLMGCPGATQDNCNKFISSSWLYSKQMVNVLIQVKGYLSS KPSIIVRTIPITENYVGAEGKLVGTRERIYIYTRSTGWHTNLQIGVLNIN HPITITWTDHRVLSRPGRSPCAWNNKCPRNCTTGVYTDAYPISPDAN YVATVTLLSNSTRNNPTIMYSSSDRVYNMLRLRNTELEAAYTTTSCIV HFDRGYCFHIIEINQKELNTLQPMLFKTAIPKACRISNL	5	85
KF908238	7510-9249	gb:KF908238:7510-9249|Organismo:vírus da parainfluenza humana 4b|Nome da cepa:QLD-01|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MQDSRGNTQIFSQANSMVKRTWRLLFRIVTLILLISIFVLSLIIVLQSTP GNLQSDVDIIRKELDELMENFETTSKSLLSVANQITYDVSVLTPIRQEA TETNIIAKIKDHCKDRVVKGESTCTLGHKPLHDVSFLNGFNKFYFTYR DNVQIRLNPLLDYPNFIPTATTPHGCIRIPSFSLSQTHWCYTHNTILRG CEDTASSKQYVSLGTLQTLENGDPYFKVEYSHYLNDRKNRKSCSVV AVLDGCLLYCVIMTKNETENFKDPQLATQLLTYISYNGTIKERIINPPGS SRDWVHISPGVGSGILYSNYIIFPLYGGLMENSMIYNNQSGKYFFPNS TKLPCSNKTSEKITGAKDSYTITYFSKRLIQSAFLICDLRQFLSEDCEILI PSNDHMLVGAEGRLYNIENNIFYYQRGSSWWPYPSLYRIKLNSNKKY PRIIEIKFTKIEIAPRPGNKDCPGNKACPKECITGVYQDIWPLSYPNTAF PHKKRAYYTGFYLNNSLARRNPTFYTADNLDYHQQERLGKFNLTAG YSTTTCFKQTTTARLYCLYILEVGDSVIGDFQIFPFLRSIDQAIT	5	86

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KT071757	6066-7962	gb:KT071757:6066-7962|Organismo:paramixovírus aviário 2|Nome da cepa:APMV-2/Emberiza spodocepha-la/China/Daxing'anling/974/2013|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MDALSRENLTEISQGGRRTWRMLFRILTLVLTLVCLAINIATIAKLDSID TSKVQTWTTTESDRVIGSLTDTLKIPINQVNDMFRIVALDLPLQMTTLQ KEIASQVGFLAESINNFLSKNGSAGSVLVNDPEYAGGIGTSLFHGDSA SGLDFEAPSLIEHPSFIPGPTTAKGCIRIPTFHMSASHWCYSHNIIASG CQDAGHSSMYISMGVLKATQAGSPSFLTTASQLVDDKLNRKSCSIIST TYGCDILCSLVVENEDADYRSDPPTDMILGRLFFNGTYSESKLNTSAI FQLFSANYPAVGSGIVLGDEIAFPVYGGVKQNTWLFNQLKDYGYFAH NNVYKCNNSNIHQTVLNAYRPPKISGRLWSQVVLICPMRLFINTDCRI KVFNTSTVMMGAEARLIQVGSDIYLYQRSSSWWVVGLTYKLDFQELS SKTGNILNNVSPIAHAKFPRPSYSRDACARPNICPAVCVSGVYQDIWP ISTAHNLSQVVWVGQYLEAFYARKDPWIGIATQYDWKKNVRLFNANT EGGYSTTTCFRNTKRDKAFCVIISEYADGVFGSYRIVPQLIEIRTTSKK GLPS	5	87
LC041132	6605-8437	gb:LC041132:6605-8437|Organismo:paramixovírus aviário ganso/Shimane/67/2000|Nome da cepa:ganso/Shimane/67/2000|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MQPGISEVSFVNDERSERGTWRLLFRILTIVLCLTSIGIGIPALIYSKEA ATSGDIDKSLEAVKTGMSTLSSKIDESINTEQKIYRQVILEAPVSQLNM ESNILSAITSLSYQIDGTSNSSGCGSPMHDQDFVGGINKEIWTTDNVN LGEITLTPFLEHLNFIPAPTTGNGCTRIPSFDLGLTHWCYTHNVILSGC QDYSSSFQYIALGVLKISATGHVFLSTMRSINLDDERNRKSCSISATSI GCDIICSLVTEREVDDYNSPAATPMIHGRLDFSGKYNEVDLNVGQLF GDWSANYPGVGGGSFLNGRVWFPIYGGVKEGTPTFKENDGRYAIYT RYNDTCPDSESEQVSRAKSSYRPSYFGGKLVQQAVLSIKIDDTLGLD PVLTISNNSITLMGAESRVLQIEEKLYFYQRGTSWFPSLIMYPLTVDDK MVRFEPPTIFDQFTRPGNHPCSADSRCPNACVTGVYTDGYPIVFHNN HSIAAVYGMQLNDVTNRLNPRSAVWYGVSMSNVIRVSSSTTKAAYTT STCFKVKKTQRVYCLSIGEIGNTLFGEFRIVPLLLEVYSEKGKSLKSSF DGWEDISINNPLRPLDNHRVDPILISNYTSSWP	4	88

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AF092942	4705-5478	gb:AF092942|Organismo:Vírus Sincicial Respiratório Bovino|Nome da Cepa:ATue51908|Nome da Proteína:glicoproteína de fixação|Símbolo do Gene:G	MSNHTHHLKFKTLKRAWKASKYFIVGLSCLYKFNLKSLVQTALTTLAMITLTSLVITAIIYISVGNAKAKPTSKPTIQQTQQPQNHTSPFFTEHNYKSTHTSIQSTTLSQLPNTDTTRETTYSHSINETQNRKIKSQSTLPATRKPPINPSGSNPPENHQDHNNSQTLPYVPCSTCEGNLACLSLCQIGPERAPSRAPTITLKKTPKPKTTKKPTKTTIHHRTSPEAKLQPKNNTAAPQQGILSSPEHHTNQSTTQI	3	89
AF326114	6691-847	gb:AF326114|Organismo:Vírus Menangle|Nome da Cepa:AF326114 DESCONHECIDA|Nome da Proteína:proteína de fixação|Símbolo do Gene:HN	MWNSIPQLVSDHEEAKGKFTDIPLQDDTDSQHPSGSKSTCRTLFRTVSIILSLVILVLGVTSTMFSAKYSGGCATNSQLLGVSNLINQIQKSIDSLISEVNQVSITTAVTLPIKIMDFGKSVTDQVTQMIRQCNTVCKGPGQKPGSQNVRIMPSNNLSTFQNINMSARGIAYQDVPLTFVRPIKNPQSCSRFPSYSVSFGVHCFANAVTDQTCELNQNTFYRVVLSVSKGNISDPSSLETKAETRTPKGTPVRTCSIISSVYGCYLLCSKATVPESEEMKTIGFSQMFILYLSMDSKRIIYDNIVSSTSAIWSGLYPGEGAGIWHMGQLFFPLWGGIPFLTPLGQKILNSTLDIPEVGSKCKSDLTSNPAKTKDMLFSPYYGENVMVFGFLTCYLLSNVPTNCHADYLNSTVLGFGSKAQFYDYRGIVYMYIQSAGWYPFTQIFRITLQLKQNRLQAKSIKRIEVTSTTRPGNRECSVLRNCPYICATGLFQVPWIVNSDAITSKEVDNMVFVQAWAADFTEFRKGILSLCSQVSCPINDLLSKDNSYMRDTTTYCFPQTVPNILSCTSFVEWGGDSGNPINILEIHYEVIFVAS	3	90

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
GU206351	7500-9714	gb:GU206351:7500-9714|Organismo:Paramixovírus aviário 5|Nome da cepa:Periquito-australiano/Kunitachi/74|Nome da proteína:proteína hemaglutinina neuraminidase|Símbolo do gene:HN	MDKSYYTEPEDQRGNSRTWRLLFRLIVLTLLCLIACTSVSQLFYPWLP QVLSTLISLNSSIITSSNGLKKEILNQNIKEDLIYREVAINIPLTLDRVTVE VGTAVNQITDALRQLQSVNGSAAFALSNSPDYSGGIEHLVFQRNTLIN RSVSVSDLIEHPSFIPTPTTQHGCTRIPTFHLGTRHWCYSHNIIGQGC ADSGASMMYISMGALGVSSLGTPTFTTSATSILSDSLNRKSCSIVATT EGCDVLCSIVTQTEDQDYADHTPTPMIHGRLWFNGTYTERSLSQSLF LGTWAAQYPAVGSGIMTPGRVIFPFYGGVIPNSPLFLDLERFALFTHN GDLECRNLTQYQKEAIYSAYKPPKIRGSLWAQGFIVCSVGDMGNCSL KVINTSTVMMGAEGRLQLVGDSVMYYQRSSSWWPVGILYRLSLVDII ARDIQVVINSEPLPLSKFPRPTWTPGVCQKPNVCPAVCVTGVYQDLW AISAGETLSEMTFFGGYLEASTQRKDPWIGVANQYSWFMRRRLFKT STEAAYSSSTCFRNTRLDRNFCLLIFELTDNLLGDWRIVPLLFELTIV	3	91
JQ001776	8170-10275	gb:JQ001776:8170-10275|Organismo:Vírus do cedro|Nome da cepa:CG1a|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MLSQLQKNYLDNSNQQGDKMNNPDKKLSVNFNPLELDKGQKDLNK SYYVKNKNYNVSNLLNESLHDIKFCIYCIFSLLIIITIINIITISIVITRLKVHE ENNGMESPNLQSIQDSLSSLTNMINTEITPRIGILVTATSVTLSSSINYV GTKTNQLVNELKDYITKSCGFKVPELKLHECNISCADPKISKSAMYST NAYAELAGPPKIFCKSVSKDPDFRLKQIDYVIPVQQDRSICMNNPLLDI SDGFFTYIHYEGINSCKKSDSFKVLLSHGEIVDRGDYRPSLYLLSSHY HPYSMQVINCVPVTCNQSSFVFCHISNNTKTLDNSDYSSDEYYITYFN GIDRPKTKKIPINNMTADNRYIHFTFSGGGGVCLGEEFIIPVTTVINTDV FTHDYCESFNCSVQTGKSLKEICSESLRSPTNSSRYNLNGIMIISQNN MTDFKIQLNGITYNKLSFGSPGRLSKTLGQVLYYQSSMSWDTYLKAG FVEKWKPFTPNWMNNTVISRPNQGNCPRYHKCPEICYGGTYNDIAP LDLGKDMYVSVILDSDQLAENPEITVFNSTTILYKERVSKDELNTRSTT TSCFLFLDEPWCISVLETNRFNGKSIRPEIYSYKIPKYC	3	92

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KP271123	6644-8431	gb:KP271123:6644-8431|Organismo:Vírus Teviot|Nome da cepa:Geelong|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MWSTQASKHPAMVNSATNLVDIPLDHPSSAQFPINRKRTGRLIYRLF SILCNLILISILISLVVIWSRSSRDCAKSDGLSSVDNQLSSLSRSINSLITE VNQISVTTAINLPIKLSEFGKSVVDQVTQMIRQCNAACKGPGEKPGIQ NVRINIPNNFSTYSELNRTANSLNFQSRTALFARPNPYPKTCSRFPSY SVYFGIHCFSHAVTDSSCELSDSTYYRLVIGVADKNLSDPADVKYIGE TTTPVRVQTRGCSVVSSIYGCYLLCSKSNQDYQDDFREQGFHQMFIL FLSRELKTTFFDDMVSSTTVTWNGLYPGEGSGIWHMGHLVFPLWGG IRFGTHASEGILNSTLELPPVGPSCKRSLADNGLINKDVLFSPYFGDS VMVFAYLSCYMLSNVPTHCQVETMNSSVLGFGSRAQFYDLKGIVYLY IQSAGWFSYTQLFRLSLQSKGYKLSVKQIKRIPISSTSRPGTEPCDIIH NCPYTCATGLFQAPWIVNGDSIRDRDVRNMAFVQAWSGAINTFQRP FMSICSQYSCPLSELLDSESSIMRSTTTYCFPSLTESILQCVSFIEWGG PVGNPISINEVYSSISFRPD	3	93
AY286409	7644-9542	gb:AY286409|Organismo:Vírus Mossman|Nome da Cepa:UNKNOWN-AY286409|Nome da Proteína:glicoproteína de fixação|Símbolo do Gene:G	MVDPPAVSYYTGTGRNDRVKVVTTQSTNPYWAHNPNQGLRRLIDMV VNVIMVTGVIFALINIILGIVIISQSAGSRQDTSKSLDIIQHVDSSVAITKQI VMENLEPKIRSILDSVSFQIPKLLSSLLGPGKTDPPIALPTKASTPVIPT EYPSLNTTTCLRIEESVTQNAAALFNISFDLKTVMYELVTRTGGCVTLP SYSELYTRVRTFSTAIRNPKTCQRAGQETDLNLIPAFIGTDTGILINSCV RQPVIATGDGIYALTYLTMRGTCQDHRHAVRHFEIGLVRRDAWWDP VLTPIHHFTEPGTPVFDGCSLTVQNQTALALCTLTTDGPETDIHNGAS LGLALVHFNIRGEFSKHKVDPRNIDTQNQGLHLVTTAGKSAVKKGILY SFGYMVTRSPEPGDSKCVTEECNQNNQEKCNAYSKTTLDPDKPRS MIIFQIDVGAEYFTVDKVVVVPRTQYYQLTSGDLFYTGEENDLLYQLH NKGWYNKPIRGRVTFDGQVTLHEHSRTYDSLSNQRACNPRLGCPST CELTSMASYFPLDKDFKAAVGVIALRNGMTPIITYSTDDWRNHWKYIK NADLEFSESSLSCYSPNPPLDDYVLCTAVITAKVMSNTNPQLLATSW YQYDKCHT	2	94

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AY900001	7809-9938	gb:AY900001|Organismo:Vírus J|Nome da Cepa:UNKNOWN-AY900001|Nome da Proteína:glicoproteína de fixação|Símbolo do Gene:G	MNPVAMSNFYGINQADHLREKGDQPEKGPSVLTYVSLITGLLSLFTIIAL NVTNIIYLTGSGGTMATIKDNQQSMSGSMRDISGMLVEDLKPKTDLINS MVSYTIPSQISAMSAMIKNEVLRQCTPSFMFNNTICPIAEHPVHTSYFEE VGIEAISMCTGTNRKLVVNQGINFVEYPSFIPGSTKPGGCVRLPSFSLGL EVFAYAHAITQDDCTSSSTPDYYFSVGRIADHGTDVPVFETLAEWFLDD KMNRRSCSVTAAGKGGWLGCSILVGSFTDELTSPEVNRISLSYMDTFG KKKDWLYTGSEVRADQSWSALFFSVGSGVVIGDTVYFLVWGGLNHPIN VDAMCRAPGCQSPTQSLCNYAIKPQEWGGNQIVNGILHFKHDTNEKPT LHVRTLSPDNNWMGAEGRLFHFHNSGKTFIYTRSSTWHTLPQVGILTLG WPLSVQWVDITSISRPGQSPCEYDNRCPHQCVTGVYTDLFPLGVSYEY SVTAYLDQVQSRMNPKIALVGAQEKIYEKTITTNTQHADYTTTSCFAYKL RVWCVSIVEMSPGVITTRQPVPFLYHLNLGCQDTSTGSLTPLDAHGGTY LNTDPVGNKVDCYFVLHEGQIYFGMSVGPINYTYSIVGRSREIGANMNV SLNQLCHSVYTEFLKEKEHPGTRNNIDVEGWLLKRIETLNGTKIFGLDDL EGSGPGHQSGPEDPSIAPIGHN	2	95
EF199772	6150-6944	gb:EF199772:6150-6944|Organismo:metapneumovírus aviário|Nome da cepa:PL-2|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MEVKVENVGKSQELKVKVKNFIKRSDCKKKLFALILGLVSFELTMNIM LSVMYVESNEALSLCRIQGTPAPRDNKTNTENATKETTLHTTTTTRDP EVRETKTTKPQANEGATNPSRNLTTKGDKHQTTRATTEAELEKQSK QTTEPGTSTQKHTPARPSSKSPTTTQATAQPTTPTAPKASTAPKNRQ ATTKKTETDTTTASRARNTNNPTETATTTPKATTETGKGKEGPTQHT TKEQPETTARETTTPQPRRTAGASPRAS	2	96

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
JF424833	5981-7156	gb:JF424833:5981-7156|Organismo: metapneumovírus aviário|Nome da cepa:IT/Ty/A/259-01/03|Nome da proteína:proteína de fixação|Símbolo do gene:G	MGSKLYMVQGTSAYQTAVGFWLDIGRRYILAIVLSAFGLTCTVTIALTV SVIVEQSVLEECRNYNGGDRDWWSTTQEQPTTAPSATPAGNYGGL QTARTRKSESCLHVQISYGDMYSRSDTVLGGFDCMGLLVLCKSGPIC QRDNQVDPTALCHCRVDLSSVDCCKVNKISTNSSTTSEPQKTNPAW PSQDNTDSDPNPQGITTSTATLLSTSLGLMLTSKTGTHKSGPPQALP GSNTNGKTTTDRELGSTNQPNSTTNGQHNKHTQRMTLPPSYDNTRT ILQHTTPWEKTFSTYKPTHSPTNESDQSLPTTQNSINCEHFDPQGKE KICYRVGSYNSNITKQCRIDVPLCSTYNTVCMKTYYTEPFNCWRRIW RCLCDDGVGLVEWCCTS	2	97

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
JN689227	7918-12444	gb:JN689227:7918-12444|Organismo:Vírus Tailam|Nome da cepa:TL8K|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MSQLAAHNLAMSNFYGIHQGGQSTSQKEEEQPVQGVIRYASMIVGLLS LFTIIALNVTNIIYMTESGGTMQSIKNAQGSIDGSMKDLSGTIMEDIKPKTD LINSMVSYNIPAQLSMIHQIIKNDVLKQCTPSFMFNNTICPLAENPTHSRY FEEVNLDSISECSGNEMSLELGTEPEFIEYPSFAPGSTKPGSCVRLPSFS LSSTVFAYTHTIMGHGCSELDVGDHYLAIGRIADAGHEIPQFETISSWFIN DKINRRSCTVAAGVMETWMGCVIMTETFYDDLDSLDTGKITISYLDVFG RKKEWIYTRSEILYDYTYTSVYFSIGSGVVVGDTVYFLLWGSLSSPIEET AYCYAPGCSNYNQRMCNEAQRPAKFGHRQMANAILRFKTNSMGKPSI SVRTLSPTVIPFGTEGRLIYSDFTKIIYLYLRSTSWYVLPLTGLLILGPPVSI SWVTQEAVSRPGEYPCGASNRCPKDCITGVYTDLFPLGARYEYAVTVY LNAETYRVNPTLALIDRSKIIARKKITTESQKAGYTTTTCFVFKLRIWCMS VVELAPATMTAFEPVPFLYQLDLTCKRNNGTTAMQFSGQDGMYKSGR YKSPRNECFFEKVSNKYYFVVSTPEGIQPYEVRDLTPERVSHVIMYISDV CAPALSAFKKLIPAMRPITTLTIGNWQFRPVDISGGLRVNIYRNLTRYGDL SMSAPEDPGTDTFPGTHAPSKGHEEVGHYTLPNEKLSEVTTAAVKTKE SLNLIPDTKDTRGEEENGSGLNEIITGHTTPGHIKTHPAETKVTKHTVIIPQ IEEDGSGATTSTELQDETGYHTEDYNTTNTNGSLTAPNERNNYTSGDH TVSGEDITHTITVSDRTKTTQTLPTDNTFNQTPTKIQEGSPKSESTPKDY TAIESEDSHFTDPTLIRSTPEGTIVQVIGDQFHSAVTQLGESNAIGNSEPI DQGNNLIPTTDRGTMDNTSSQSHSSTTSTQGSHSAGHGSQSNMNLTA LADTDSVTDQSTSTQEIDHEHENVSSILNPLSRHTRVMRDTVQEALTGA WGFIRGMIP	2	98
KC562242	6178-6926	gb:KC562242:6178-6926|Organismo:metapneumovírus humano|Nome da cepa:HMPV/USA/C1-334/2004/B|Nome da proteína:glicoproteína G de fixação|Símbolo do gene:G	MEVRVENIRAIDMFKAKIKNRIRNSRCYRNATLILIGLTALSMALNIFLII DHATLRNMIKTENCANMPSAEPSKKTPMTSIAGPSTKPNPQQATQW TTENSTSPAATLEGHPYTGTTQTPDTTAPQQTTDKHTALPKSTNEQIT QTTTEKKTTRATTQKRKKEKKTQTKPQVQLQPKQPTPPTKSEMQVR QSQHPTDPELTPLPKAVNRQPGQQNQAPHHIMHGEVQDPGERNTQ VSHPSS	2	99

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KC915036	6154-7911	gb:KC915036:6154-7911|Organismo:metapneumovírus aviário tipo C|Nome da cepa:GDY|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MEVKIENVGKSQELRVKVKNFIKRSDCKKKLFALILGLISFDITMNIMLS VMYVESNEALSSCRVQGTPAPRDNRTNTENTAKETTLHTMTTTRNT EAGGTKTTKPQADERATSPSKNPTIGADKHKTTRATTEAEQEKQSKQ TTEPGTSTPKHIPARPSSKSPATTKTTTQPTTPTVAKGGTAPKNRQTT TKKTEADTPTTSRAKQTNKPTGTETTPPRATTETDKDKEGPTQHTTK EQPETTAGGTTTPQPRRTTSRPAPTTNTKEGAETTGTRTTKSTQTSA SPPRPTRSTPSKTATGTNKRATTTKGPNTASTDRRQQTRTTPKQDQ QTQTKAKTTTNKAHAKAATTPEHNTDTTDSMKENSKEDKTTRDPSSK ATTKQENTSKGTTATNLGNNTEAGARTPPTTTPTRHTTEPATSTAGG HTKARTTRWKSTAARQPTRNNTTADTKTAQSKQTTPAQLGNNTTPE NTTPPDNKSNSQTNVAPTEEIEIGSSLWRRRYVYGPCRENALEHPMN PCLKDNTTWIYLDNGRNLPAGYYDSKTDKIICYGIYRGNSYCYGRIEC TCKNGTGLLSYCCNSYNWS	2	100
LC168749	7239-9196	gb:LC168749:7239-9196|Organismo:morbilivírus da peste bovina|Nome da cepa:Lv|Nome da proteína:proteína H|Símbolo do gene:H	MSSPRDRVNAFYKDNLQFKNTRVVLNKEQLLIERPYMLLAVLFVMFL SLVGLLAIAGIRLHRAAVNTAEINSGLTTSIDITKSIEYQVKDVLTPLFKII GDEVGLRTPQRFTDLTKFISDKIKFLNPDKEYDFRDINWCISPPERIKI NYDQYCAHTAAEELITMLVNSSLAGTAVLRTSLVNLGRSCTGSTTTK GQFSNMSLALSGIYSGRGYNISSMITITEKGMYGSTYLVGKHNQGAR RPSTAWQRDYRVFEVGIIRELGVGTPVFHMTNYLELPRQPELEICML ALGEFKLAALCLADNSVALHYGGLRDDHKIRFVKLGVWPSPADSDTL ATLSAVDPTLDGLYITTHRGIIAAGKAVWAVPVTRTDDQRKMGQCRR EACREKPPPFCNSTDWEPLEAGRIPAYGILTIRLGLADKPEIDIISEFGP LITHDSGMDLYTPLDGNEYWLTIPPLQNSALGTVNTLVLEPSLKISPNI LTLPIRSGGGDCYTPTYLSDLADDDVKLSSNLVILPSRNLQYVSATYD TSRVEHAIVYYIYSTGRLSSYYYPVKLPIKGDPVSLQIGCFPWGLKLW CHHFCSVIDSGTGKQVTHTGAVGIEITCNSR	2	101

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
LC187310	8144-9871	gb:LC187310:8144-9871|Organismo:Paramixovírus aviário 10|Nome da cepa:rAPMV-10-FI324/YmHA|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MDSSQMNILDAMDRESSKRTWRGVFRVTTIIMVVTCVVLSAITLSKVAHPQGFDTNELGNGIVDRVSDKITEALTVPNNQIGEIFKIVALDLHVLVSSSQQAIAGQIGMLAESINSILSQNGSASTILSSSPEYAGGIGVPLFSNKLTNGTVIKPITLIEHPSFIPGPTTIGGCTRIPTFHMASSHWCYSHNIIEKGCKDSGISSMYISLGVLQVLKKGTPVFLVTASAVLSDDRNRKSCSIISSRFGCEILCSLVTEAESDDYKSDTPTGMVHGRLYFNGTYREGLVDTETIFRDFSANYPGVGSGEIVEGHIHFPIYGGVKQNTGLYNSLTPYWLDAKNKYDYCKLPYTNQTIQNSYKPPFIHGRFWAQGILSCELDLFNLGNCNLKIIRSDKVMMGAESRLMLVGSKLLMYQRASSWWPLGITQEIDIAELHSSNTTILREVKPILSSKFPRPSYQPNYCTKPSVCPAVCVTGVYTDMWPISITGNISDYAWISHYLDAPTSRQQPRIGIANQYFWIHQTTIFPTNTQSSYSTTTCFRNQVRSRMFCLSIAEFADGVFGEFRIVPLLYELRV	2	102
NC_004074	6590-8563	gb:NC_004074:6590-8563|Organismo:Vírus Tioman|Nome da cepa:NC_004074 DESCONHECIDA|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MWATSESKAPIPANSTLNLVDVPLDEPQTITKHRKQKRTGRLVFRLLSLVLSLMTVILVLVILASWSQKINACATKEGFNSLDLQISGLVKSINSLITEVNQISITTAINLPIKLSDFGKSIVDQVTQMIRQCNAVCKGPGEKPGIQNIRINIPNNFSTYLELNNTVKSIELQRRPALLARPNPIPKSCSRFPSYSVNFGIHCFAHAITDQSCELSDKTYYRLAIGISDKNLSDPSDVKYIGEAFTPMGLQARGCSVISSIYGCYLLCSKSNQGYEADFQTQGFHQMYILFLSRDLKTTLFNDMISSTTVVWNGLYPGEGAGIWHMGYLIFPLWGGIKIGTPASTSILNSTLDLPLVGPSCKSTLEENNLINKDVLFSPYFGESVMVFGFLSCYMLSNVPTHCQVEVLNSSVLGFGSRSQLMDLKGIVYLYIQSAGWYSYTQLFRLSLQSRGYKLTVKQIRRIPISSTTRPGTAPCDVVHNCPYTCATGLFQAPWIVNGDSILDRDVRNLVFVQAWSGNFNTFQKGLISICNQYTCPLTTLLDNDNSIMRSTTTYCYPSLSEYNLQCQSFIEWGGPVGNPIGILEVHYIIKFK	2	103

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_005283	7091-8905	gb:NC_005283:7091-8905|Organismo:Morbilivírus de golfinho|Nome da cepa:NC_005283 DESCONHECIDA|Nome da proteína:proteína hemaglutinina|Símbolo do gene:H	MSSPRDKVDAFYKDIPRPRNNRVLLDNERVIIERPLILVGVLAVMFLSLVGLLAIAGVRLQKATTNSIEVNRKLSTNLETTVSIEHHVKDVLTPLFKIIGDEVGLRMPQKLTEIMQFISNKIKFLNPDREYDFNDLHWCVNPPDQVKIDYAQYCNHIAAEELIVTKFKELMNHSLDMSKGRIFPPKNCSGSVITRGQTIKPGLTLVNIYTTRNFEVSFMVTVISGGMYGKTYFLKPPEPDDPFEFQAFRIFEVGLVRDVGSREPVLQMTNFMVIDEDEGLNFCLLSVGELRLAAVCVRGRPVVTKDIGGYKDEPFKVVTLGIIGGGLSNQKTEIYPTIDSSIEKLYITSHRGIIRNSKARWSVPAIRSDDKDKMEKCTQALCKSRPPPSCNSSDWEPLTSNRIPAYAYIALEIKEDSGLELDITSNYGPLIIHGAGMDIYEGPSSNQDWLAIPPLSQSVLGVINKVDFTAGFDIKPHTLTTAVDYESGKCYVPVELSGAKDQDLKLESNLVVLPTKDFGYVTATYDTSRSEHAIVYYVYDTARSSSYFFPFRIKARGEPIYLRIECFPWSRQLWCHHYCMINSTVSNEIVVVDNLVSINMSCSR	2	104

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_007803	7978-12504	gb:NC_007803:7978-12504|Organismo:Vírus Beilong|Nome da cepa:Li|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MSQLAAHNLAMSNFYGTHQGDLSGSQKGEEQQVQGVIRYVSMIVSL LSLFTIIALNVTNIIYMTESGGTMQSIKTAQGSIDGSMREISGVIMEDVK PKTDLINSMVSYNIPAQLSMIHQIIKNDVPKQCTPSFMFNNTICPLAEN PTHSRYFEEVNLDSISECSGPDMHLGLGVNPEFIEFPSFAPGSTKPG SCVRLPSFSLSTTVFAYTHTIMGHGCSELDVGDHYFSVGRIADAGHEI PQFETISSWFINDKINRRSCTVAAGAMEAWMGCVIMTETFYDDRNSL DTGKLTISYLDVFGRKKEWIYTRSEILYDYTYTSVYFSVGSGVVVGDT VYFLIWGSLSSPIEETAYCFAPDCSNYNQRMCNEAQRPSKFGHRQM VNGILKFKTTSTGKPLLSVGTLSPSVVPFGSEGRLMYSEITKIIYLYLRS TSWHALPLTGLFVLGPPTSISWIVQRAVSRPGEFPCGASNRCPKDCV TGVYTDLFPLGSRYEYAATVYLNSETYRVNPTLALINQTNIIASKKVTT ESQRAGYTTTTCFVFKLRVWCISVVELAPSTMTAYEPIPFLYQLDLTC KGKNGSLAMRFAGKEGTYKSGRYKSPRNECFFEKVSNKYYFIVSTPE GIQPYEIRDLTPDRMPHIIMYISDVCAPALSAFKKLLPAMRPITTLTIGN WQFRPVEVSGGLRVNIGRNLTKEGDLTMSAPEDPGSNTFPGNHIPG NGILDAGYYTVEYPKE	2	105

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_009489	6559-8512	gb:NC_009489:6559-8512|Organismo:Vírus Mapuera|Nome da cepa:BeAnn 370284|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MASLQSEPGSQKPHYQSDDQLVKRTWRSFFRFSVLVVTITSLALSIITLIGVNRISTAKQISNAFAAIQANILSSIPDIRPINSLLNQLVYTSSVTLPLRISSLESNVLAAIQEACTYRDSQSSCSATMSVMNDQRYIEGIQVYSGSFLDLQKHTLSPPIAFPSFIPTSTTTVGCTRIPSFSLTKTHWCYTHNYIKTGCRDATQSNQYIALGTIYTDPDGTPGFSTSRSQYLNDGVNRKSCSISAVPMGCALYCFISVKEEVDYYKGTVPPAQTLILFFFNGTVHEHRIVPSSMNSEWVMLSPGVGSGVFYNNYIIFPLYGGMTKDKAEKRGELTRFFTPKNSRSLCKMNDSVFSNAAQSAYYPPYFSSRWIRSGLLACNWNQIITTNCEILTFSNQVMMMGAEGRLILINDDLFYYQRSTSWWPRPLVYKLDIELNYPDSHIQRVDQVEVTFPTRPGWGGCVGNNFCPMICVSGVYQDVWPVTNPVNTTDSRTLWVGGTLLSNTTRENPASVVTSGGSISQTVSWFNQTVPGAYSTTTCFNDQVQGRIFCLIIFEVGGGLLGEYQIVPFLKELKYQGAVHA	2	106
NC_017937	6334-8544	gb:NC_017937:6334-8544|Organismo:vírus Nariva|Nome da cepa:NC_017937 DESCONHECIDA|Nome da proteína:proteína de fixação|Símbolo do gene:H	MAPINYPASYYTNNAERPVVITTKSTESKGQRPLPLGNARFWEYFGHVCGTLTFCMSLIGIIVGIIALANYSSDKDWKGRIGGDIQVTRMATEKTVKLILEDTTPKLRNILDSVLFQLPKMLASIASKINTQTPPPPTTSGHSTALATQCSSNCENRPEIGYDYLRQVEQSLQRITNISIQLLEASEIHSMAGAYPNALYKIRTQDSWSVTAKECPLQAFQPNLNLIPAMIGTATGALIRNCVRQPVIVVDDGVYMLTYLAMRGSCQDHQKSVRHFEMGVITSDPFGDPVPTPLRHWTKRALPAYDGCALAVKGHAGFALCTETSVGPLRDRTAKRKPNIVLFKASLVGELSERVIPPQSWLSGFSFFSVYTVAGKGYAYHSKFHAFGNVVRVGQSEYQAKCRGTGCPTANQDDCNTAQRVSQEDNTYLHQAILSVDIDSVIDPEDVVYVIERDQYYQASAGDLYRVPETGEILYNLHNGGWSNEVQVGRIQPSDRFYMREIQLTSTRVPAPNGCNRVKGCPGGCVAVISPAFTPMHPEFNVGVGIFPMNQPHNPSIMHVQQQTELFWKPIVGGNITLHESSIACYSTVPPNPSYDLCIGVMTLLLHQGQLPQFQALSWYQPTMCNGNAPQNRRALIPVIVEDSKAMSVSSDAPRTP	2	107

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025256	9117-11015	gb:NC_025256:9117-11015|Organismo:Paramixovírus de morcego Eid_hel/GH-M74a/GHA/2009|Nome da cepa:BatPV/Eid_hel/GH-M74a/GHA/2009|Nome da proteína:glicoproteína|Símbolo do gene:G	MPQKTVEFINMNSPLERGVSTLSDKKTLNQSKITKQGYFGLGSHSERNWKKQKNQNDHYMTVSTMILEILVVLGIMFNLIVLTMVYYQNDNINQRMAELTSNITVLNLNLNQLTNKIQREIIPRITLIDTATTITIPSAITYILATLTTRISELLPSINQKCEFKTPTLVLNDCRINCTPPLNPSDGVKMSSLATNLVAHGPSPCRNFSSVPTIYYYRIPGLYNRTALDERCILNPRLTISSTKFAYVHSEYDKNCTRGFKYYELMTFGEILEGPEKEPRMFSRSFYSPTNAVNYHSCTPIVTVNEGYFLCLECTSSDPLYKANLSNSTFHLVILRHNKDEKIVSMPSFNLSTDQEYVQIIPAEGGGTAESGNLYFPCIGRLLHKRVTHPLCKKSNCSRTDDESCLKSYYNQGSPQHQVVNCLIRIRNAQRDNPTWDVITVDLTNTYPGSRSRIFGSFSKPMLYQSSVSWHTLLQVAEITDLDKYQLDWLDTPYISRPGGSECPFGNYCPTVCWEGTYNDVYSLTPNNDLFVTVYLKSEQVAENPYFAIFSRDQILKEFPLDAWISSARTTTISCFMFNNEIWCIAALEITRLNDDIIRPIYYSFWLPTDCRTPYPHTGKMTRVPLRSTYNY	2	108
NC_025347	6398-8418	gb:NC_025347:6398-8418|Organismo:Paramixovírus aviário 7|Nome da cepa:APMV-7/dove/Tennessee/4/75|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MESIGKGTWRTVYRVLTILLDVVIIILSVIALISLGLKPGERIINEVNGSIHNQLVPLSGITSDIQAKVSSIYRSNLLSIPLQLDQINQAISSSARQIADTINSFLALNGSGTFIYTNSPEFANGFNRAMFPTLNQSLNMLTPGNLIEFTNFIPTPTTKSGCIRIPSFSMSSSHWCYTHNIIASGCQDHSTSSEYISMGVVEVTDQAYPNFRTTLSITLADNLNRKSCSIAATGFGCDILCSVVTETENDDYQSPEPTQMIYGRLFFNGTYSEMSLNVNQMFADWVANYPAVGSGVELADFVIFPLYGGVKITSTLGASLSQYYYIPKVPTVNCSETDAQQIEKAKASYSPPKVAPNIWAQAVVRCNKSVNLANSCEILTFNTSTMMMGAEGRLLMIGKNVYFYQRSSSYWPVGIIYKLDLQELTTFSSNQLLSTIPIPFEKFPRPASTAGVCSKPNVCPAVCQTGVYQDLWVLYDLGKLENTTAVGLYLNSAVGRMNPFIGIANTLSWYNTTRLFAQGTPASYSTTTCFKNTKIDTAYCLSILELSDSLLGSWRITPLLYNITLSIMS	2	109

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025348	6590-8548	gb:NC_025348:6590-8548|Organismo:vírus Tuhoko 2|Nome da cepa:NC_025348 DESCONHECIDA|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MPPVPTVSQSIDEGSFTDIPLSPDDIKHPLSKKTCRKLFRIVTLIGVGLISILTIISLAQQTGILRKVDSSDFQSYVQESFKQVLNLMKQFSSNLNSLIEITSVTLPFRIDQFGTDIKTQVAQLVRQCNAVCRGPIKGPTTQNIVYPALYETSLNKTLETKNVRIQEVRQEVDPVPGPGLSNGCTRNPSFSVYHGVWCYTHATSIGNCNGSLGTSQLFRIGNVLEGDGGAPYHKSLATHLLTTRNVSRQCSATASYYGCYFICSEPVLTERDDYETPGIEPITIFRLDPDGNWVVFPNINRFTEYSLKALYPGIGSGVLFQGKLIFPMYGGIDKERLSALGLGNIGLIERRMADTCNHTEKELGRSFPGAFSSPYYHDAVMLNFLLICEMIENLPGDCDLQILNPTNMSMGSESQLSVLDNELFLYQRSASWWPYTLIYRLNMRYTGKYLKPKSIIPMVIKSNTRPGYEGCNHERVCPKVCVTGVFQAPWILSIGRDHKERVSNVTYMVAWSMDKSDRTYPAVSVCGSDTCKLTVPLGDSKVHSAYSVTRCYLSRDHMSAYCLVIFELDARPWAEMRIQSFLYKLILT	2	110
NC_025350	6451-8341	gb:NC_025350:6451-8341|Organismo:vírus Tuhoko 3|Nome da cepa:NC_025350 DESCONHECIDA|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MHNRTQSVSSIDTSSDVYLPRRKKAVTKFTFKKIFRVLILTLLLSIIIIIAVIFPKIDHIRETCDNSQILETITNQNSEIKNLINSAITNLNVLLTSTTVDLPIKLNNFGKSIVDQVTMMVRQCNAVCRGPGDRPTQNIELFKGLYHTSPPSNTSTKLSMITEASNPDDIVPRPGKLLGCTRFPSFSVHYGLWCYGHMASTGNCSGSSPSVQIIRIGSIGTNKDGTPKYVIIASASLPETTRLYHCSVTMTSIGCYILCTTPSVSETDDYSTMGIEKMSISFLSLDGYLTQLGQPTGLDNQNLYALYPGPGSGVIFRDFLIFPMMGGIRLMDAQKMLNRNITYRGFPPSETCTESELKLKQEVANMLTSPYYGEVLVLNFLYVCSLLDNIPGDCSVQLIPPDNMTLGAESRLYVLNGSLIMYKRGSSWWPYTELYQINYRVNNRAFRVRESVRINTTSTSRPGVQGCNLEKVCPKVCVSGIYQSPGIISAPVNPTRQEEGLLYFLVWTSSMSSRTGPLSSLCDHSTCRITYPIGDDTIFIGYTDSSCFMSSIKEGIYCIAFLELDNQPYSMMAIRSLSYIIN	2	111

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025352	8716-11257	gb:NC_025352:8716-11257|Organismo:Vírus Mojiang|Nome da cepa:Tongguan1|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSISGNKVFILMNTL LILTGAIITITLNITNLTAAKSQQNMLKIIQDDVNAKLEMFVNLDQLVKGE IKPKVSLINTAVSVSIPGQISNLQTKFLQKYVYLEESITKQCTCNPLSGI FPTSGPTYPPTDKPDDDTTDDDKVDTTIKPIEYPKPDGCNRTGDHFT MEPGANFYTVPNLGPASSNSDECYTNPSFSIGSSIYMFSQEIRKTDCT AGEILSIQIVLGRIVDKGQQGPQASPLLVWAVPNPKIINSCAVAAGDE MGWVLCSVTLTAASGEPIPHMFDGFWLYKLEPDTEVVSYRITGYAYL LDKQYDSVFIGKGGGIQKGNDLYFQMYGLSRNRQSFKALCEHGSCL GTGGGGYQVLCDRAVMSFGSEESLITNAYLKVNDLASGKPVIIGQTF PPSDSYKGSNGRMYTIGDKYGLYLAPSSWNRYLRFGITPDISVRSTT WLKSQDPIMKILSTCTNTDRDMCPEICNTRGYQDIFPLSEDSEYYTYI GITPNNGGTKNFVAVRDSDGHIASIDILQNYYSITSATISCFMYKDEIW CIAITEGKKQKDNPQRIYAHSYKIRQMCYNMKSATVTVGNAKNITIRR Y	2	112
NC_025363	6503-8347	gb:NC_025363:6503-8347|Organismo:Paramixovírus aviário 12|Nome da cepa:Wigeon/Italy/3920_1/2005|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MESATSQVSFENDKTSDRRTWRAVFRVLMIILALSSLCVTVAALIYSA KAAIPGNIDASEQRILSSVEAVQVPVSRLEDTSQKIYRQVILEAPVTQL NMETNILNAITSLSYQIDASANSSGCGAPVHDSDFTGGVGRELLQEA EVNLTIIRPSKFLEHLNFIPAPTTGNGCTRIPSFDLGQTHWCYTHNVVL NGCRDRGHSFQYVALGILRTSATGSVFLSTLRSVNLDDDRNRKSCSV SATPIGCEMLCSLVTETEEGDYDSIDPTPMVHGRLGFDGKYREVDLS EKEIFADWRANYPAVGGGAFFGNRVWFPVYGGLKEGTQSERDAEK GYAIYKRFNNTCPDDNTTQIANAKASYRPSRFGGRFIQQGILSFKVEG NLGSDPILSLTDNSITLMGAEARVMNIENKLYLYQRGTSWFPSALVYP LDVANTAVKVRAPYIFDKFTRPGGHPCSASSRCPNVCVTGVYTDAYP LVFSRSHDIVAVYGMQLAAGTARLDPQAAIWYGNEMSTPTKVSSSTT KAAYTTSTCFKVTKTKRIYCISIAEIGNTLFGEFRIVPLLIEVQKTPLTRR SELRQQMPQPPIDLVIDNPFCAPSGNLSRKNAIDEYANSWP	2	113

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025373	6619-8605	gb:NC_025373:6619-8605|Organismo:Paramixovírus aviário 3|Nome da cepa:Turkey/Wisconsin/68|Nome da proteína:hemaglutinina|Símbolo do gene:HN	MEPTGSKVDIVPSQGTKRTCRTFYRLLILILNLIIIILTIISIYVSISTDQHKL CNNEADSLLHSIVEPITVPLGTDSDVEDELREIRRDTGINIPIQIDNTENI ILTTLASINSNIARLHNATDESPTCLSPVNDPRFIAGINKITKGSMIYRNF SNLIEHVNFIPSPTTLSGCTRIPSFSLSKTHWCYSHNVISTGCQDHAA SSQYISIGIVDTGLNNEPYLRTMSSRLLNDGLNRKSCSVTAGAGVCW LLCSVVTESESADYRSRAPTAMILGRFNFYGDYTESPVPASLFSGRF TANYPGVGSGTQLNGTLYFPIYGGVVNDSDIELSNRGKSFRPRNPTN PCPDPEVTQSQRAQASYYPTRFGRLLIQQAILACRISDTTCTDYYLLY FDNNQVMMGAEARIYYLNNQMYLYQRSSSWWPHPLFYRFSLPHCE PMSVCMITDTHLILTYATSRPGTSICTGASRCPNNCVDGVYTDVWPL TEGTTQDPDSYYTVFLNSPNRRISPTISIYSYNQKISSRLAVGSEIGAA YTTSTCFSRTDTGALYCITIIEAVNTIFGQYRIVPILVQLISD	2	114
NC_025386	7541-9403	gb:NC_025386:7541-9403|Organismo:Vírus de Salem|Nome da cepa:NC_025386 DESCONHECIDA|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MKAMHYYKNDFADPGTNDNSSDLTTNPFISNQIKSNLSPPVLAEGHL SPSPIPKFRKILLTISFVSTIVVLTVILLVLTIRILTIIEASAGDEKDIHTILSS LLNTFMNEYIPVFKNLVSIISLQIPQMLIDLKTSSTQMMQSLKTFPRDLE TLSTVTQSVAVLLEKAKSTIPDINKFYKNVGKVTFNDPNIKVLTLEVPA WLPIVRQCLKQDFRQVISNSTGFALIGALPSQLFNEFEGYPSLAIVSEV YAITYLKGVMFENQENFLYQYFEIGTISPDGYNKPYFLRHTSVMLSTF KLSGKCTAAVDYRGGIFLCTPSPKIPKILQNPPDLPTLTVVSIPFDGRY TIRNISLMLTDEADIIYDLDTLQGRGVLQAMRFYALVRVISSSSPRHFP FCKNSWCPTADDKICDQSRRLGADGNYPVMYGLISIPAHSSYQGNV SLKLIDPKYYAYTRDASLFYNSMTDTYHYSFGTRGWVSRPIIGELLLG DDIVLTRYTVRSVSRATAGDCTTVSMCPQACSGGMNSIFYPLNFDKP QVTGVAIRQYERQQEGIIVVTMNDHYYYSVPIIKNGTLLISSVTDCFWL MGDLWCMSLMEKNNLPLGVRSLAHLTWNIHWSCS	2	115

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025390	6647-8386	gb:NC_025390:6647-8386|Organismo:paramixovírus aviário 9|Nome da cepa:pato/Nova Iorque/22/1978|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MESGISQASLVNDNIELRNTWRTAFRVVSLLLGFTSLVLTACALHFAL NAATPADLSSIPVAVDQSHHEILQTLSLMSDIGNKIYKQVALDSPVALL NTESTLMSAITSLSYQINNAANNSGCGAPVHDKDFINGVAKELFVGSQ YNASNYRPSRFLEHLNFIPAPTTGKGCTRIPSFDLAATHWCYTHNVIL NGCNDHAQSYQYISLGILKVSATGNVFLSTLRSINLDDDENRKSCSIS ATPLGCDLLCAKVTEREEADYNSDAATRLVHGRLGFDGVYHEQALP VESLFSDWVANYPSVGGGSYFDNRVWFGVYGGIRPGSQTDLLQSE KYAIYRRYNNTCPDNNPTQIERAKSSYRPQRFGQRLVQQAILSIRVEP SLGNDPKLSVLDNTVVLMGAEARIMTFGHVALMYQRGSSYFPSALLY PLSLTNGSAAASKPFIFEQYTRPGSPPCQATARCPNSCVTGVYTDAY PLFWSEDHKVNGVYGMMLDDITSRLNPVAAIFDRYGRSRVTRVSSSS TKAAYTTNTCFKVVKTKRVYCLSIAEIENTLFGEFRITPLLSEIIFDPNLE PSDTSRN	2	116
NC_025403	6692-8645	gb:NC_025403:6692-8645|Organismo:vírus Achimota 1|Nome da cepa:NC_025403 DESCONHECIDA|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MATNLSTITNGKFSQNSDEGSLTELPFFEHNRKVATTKRTCRFVFRS VITLCNLTILIVTVVVLFQQAGFIKRTESNQVCETLQNDMHGVVTMSK GVITTLNNLIEITSVNLPFQMKQFGQGIVTQVTQMVRQCNAVCKGPTI GPDIQNIVYPASYESMIKHPVNNSNILLSEIRQPLNFVPNTGKLNGCTR TPSFSVYNGFWCYTHAESDWNCNGSSPYMQVFRVGVVTSDYDYNV IHKTLHTKTSRLANVTYQCSTISTGYECYFLCSTPNVDEITDYKTPGIE SLQIYKIDNRGTFAKFPITDQLNKELLTALYPGPGNGVLYQGRLLFPM HGGMQSSELNKVNLNNTVLSQFNDNKGCNATEIKLESEFPGTFTSPY YSNQVMLNYILICEMIENLPGNCDLQIVAPKNMSMGSESQLYSINNKL YLYQRSSSRWPYPLIYEVGTRLTNRQFRLRAINRFLIKSTTRPGSEGC NIYRVCPKVCVTGVYQAPWILHVSKAGSQSIAKVLYAVAWSKDHMSR KGPLFSICDNDTCFLTKSLASEHVHSGYSITRCYLENSERHIICVVIME LDASPWAEMRIQSVIYNITLPS	2	117

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_025404	6655-8586	gb:NC_025404:6655-8586|Organismo:vírus Achimota 2|Nome da cepa:NC_025404 DESCONHECIDA|Nome da proteína:proteína de fixação|Símbolo do gene:HN	MDNSMSISTISLDAQPRIWSRHESRRTWRNIFRITSLVLLGVTVIICIWLCCEVARESELELLASPLGALIMAINTIKSSVVKMTTELNQVTFTTSIILPNKVDQFGQNVVSQVAQLVKQCNAVCRGHQDTPELEQFINQKNPTWILQPNYTTKLTNLHEIDSIIPLVDYPGFSKSCTRFPSFSEGSKFWCFTYAVVKEPCSDISSSIQVVKYGAIKANHSDGNPYLVLGTKVLDDGKFRRGCSITSSLYGCYLLCSTANVSEVNDYAHTPAYPLTLELISKDGITTDLSPTYTVQLDKWSALYPGIGSGVIFKGYLMFPVYGGLPFKSPLISASWVGPGNKWPVDFSCSEDQYSTFNFSNPYSALYSPHFSNNIVVSALFVCPLNENLPYSCEVQVLPQGNLTIGAEGRLYVIDQDLYYYQRSTSWWPYLQLYKLNIRITNRVFRVRSLSLLPIKSTTRPGYGNCTYFKLCPHICVTGVYQSPWLISIRDKRPHEEKNILYFIGWSPDEQIRQNPLVSLCHETACFINRSLATNKTHAGYSESHCVQSFERNKLTCTVFYELTAKPWAEMRVQSLLFQVDFL	2	118
NC_025410	6799-8869	gb:NC_025410:6799-8869|Organismo:vírus Tuhoko 1|Nome da cepa:NC_025410 DESCONHECIDA|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MDSRSDSFTDIPLDNRIERTVTSKKTWRSIFRVTAIILLIICVVVSSISLNQHNDAPLNGAGNQATSGFMDAIKSLEKLMSQTINELNQVVMTTSVQLPNRITKFGQDILDQVTQMVRQCNAVCRGPGVGPSIQNYVIQGHAPTVSFDPISAEYQKFVFGITEKTLITAYHNPWECLRFPSQHLFDTTWCVSYQILTQNCSDHGPRITVIQLGEIMIANNLSTVFRDPVIKYIRHHIWLRSCSVVAYYSQCTIFCTSTNKSEPSDYADTGYEQLFLATLQSDGTFTEHSMHGVNIVHQWNAIYGGVGNGVIIGRNMLIPLYGGINYYDHNTTIVQTVDLRPYPIPDSCSQTDNYQTNYLPSMFTNSYYGTNLVVSGYLSCRLMAGTPTSCSIRVIPIENMTMGSEGQFYLINNQLYYYKRSSNWIRDTQVYLLSYSDKGNIIEITSAERYIFKSVTSPDEGDCVTNHGCPSNCIGGLFQAPWILNDFKLCGSNITCPKIVTVWADQPDKRSNPMLSIAETDKLLLHKSYINYHTAVGYSTVLCFDSPKLNLKTCVVLQELMSDDKLLIRISYSIVSIMVE	2	119

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
NC_028249	7059-9010	gb:NC_028249:7059-9010|Organismo:Vírus da cinomose focina|Nome da cepa:PDV/Wadden_Sea.NLD/1988|Nome da proteína:proteína hemaglutinina|Símbolo do gene:H	MFSHQDKVGAFYKNNARANSSKLSLVTDEVEERRSPWFLSILLILLVGILILLAITGIRFHQVVKSNLEFNKLLIEDMEKTKAVHHQVKDVLTPLFKIIGDEVGLRLPQKLNEIKQFIVQKTNFFNPNREFDFRELHWCINPPSKVKVNFTQYCEITEFKEATRSVANSILLLTLYRGRDDIFPPYKCRGATTSMGNVFPLAVSLSMSLISKPSEVINMLTAISEGIYGKTYLLVTDDTEENFETPEIRVFEIGFINRWLGDMPLFQTTNYRIISNNSNTKICTIAVGELALASLCTKESTILLNLGDEESQNSVLVVILGLFGATHMDQLEEVIPVAHPSIEKIHITNHRGFIKDSVATWMVPALALSEQGEQINCLRSACKRRTYPMCNQTSWEPFGDKRLPSYGRLTLSLDVSTDLSINVSVAQGPIIFNGDGMDYYEGTLLNSGWLTIPPKNGTILGLINQASKGDQFIVTPHILTFAPRESSTDCHLPIQTYQIQDDDVLLESNLVVLPTQSFEYVVATYDVSRSDHAIVYYVYDPARTVSYTYPFRLRTKGRPDILRIECFVWDGHLWCHQFYRFQLDATNSTSVVENLIRIRFSCDRLDP	2	120
NC_028362	6951-8675	gb:NC_028362:6951-8675|Organismo:Vírus da parainfluenza caprina 3|Nome da cepa:JS2013|Nome da proteína:hemaglutinina-neuraminidase|Símbolo do gene:HN	MEYWGHTNNPDKINRKVGVDQVRDRSKTLKIITFIISMMTSIMSTVALILILIMFIQNNNNNRIILQELRDETDAIEARIQKASNDIGVSIQSGINTRLLTIQNHVQNYIPLALTQQVSSLRESINDVITKREETQSKMPIQRMTHDDGIEPLIPDNFWKCPSGIPTISASPKIRLIPGPGLLATSTTINGCIRLPSLVINNLIYAYTSNLITQGCQDIGKSYQVLQIGIITINSDLVPDLNPRITHTFDIDDNRKSCSLALRNADVYQLCSTPKVDERSDYSSIGIEDIVLDIVTSEGTVSTTRFTNNNITFDKPYAALYPSVGPGIYYDNKIIFLGYGGLEHEENGDVICNITGCPGKTQHDCNQASYSPWFSNRRMVNAIILVNKGLNKVPSLQVWTIPMRQNYWGSEGRLLLLGNKIYIYTRSTSWHSKLQLGTLDISNYNDIRIRWTHHDVLSRPGSEECPWGNTCPRGCITGVYNDAYPLNPSGSVVSSVILDSRTSRENPIITYSTDTSRVNELAIRNNTLSAAYTTTNCVTHYGKGYCFHIIEINHKSLNTLQPMLFKTEIPKSCN	2	121

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AB548428	5999-7261	gb:AB548428:5999-7261|Organismo:Metapneumovírus aviário|Nome da Cepa:VCO3/60616|Nome da Proteína:Glicoproteína de fixação|Símbolo do Gene:G	MGSELYIIEGVSSSEIVLKQVLRRSKKILLGLVLSALGLTLTSTIVISICIS VEQVKLRQCVDTYWAENGSLHPGQSTENTSTRGKTTTKDPRRLQAT GAGKFESCGYVQVVDGDMHDRSYAVLGGVDCLGLLALCESGPICQG DTWSEDGNFCRCTFSSHGVSCCKKPKSKATTAQRNSKPANSKSTPP VHSDRASKEHNPSQGEQPRRGPTSSKTTIASTPSTEDTAKPTISKPKL TIRPSQRGPSGSTKAASSTPSHKTNTRGTSKTTDQRPRTGPTPERPR QTHSTATPPPTTPIHKGRAPTPKPTTDLKVNPREGSTSPTAIQKNPTT QSNLVDCTLSDPDEPQRICYQVGTYNPSQSGTCNIEVPKCSTYGHAC MATLYDTPFNCWRRTRRCICDSGGELIEWCCTSQ	1	122
AF079780	8118-10115	gb:AF079780|Organismo:Paramixovírus da Tupaia|Nome da Cepa:AF079780 DESCONHECIDA|Nome da Proteína:hemaglutinina|Símbolo do Gene:H	MDYHSHTTQTGSNETLYQDPLQSQSGSRDTLDGPPSTLQHYSNPPP YSEEDQGIDGPQRSQPLSTPHQYDRYYGVNIQHTRVYNHLGTIYKGL KLAFQILGWVSVIITMIITVTTLKKMSDGNSQDSAMLKSLDENFDAIQE VANLLDNEVRPKLGVTMTQTTFQLPKELSEIKRYLLRLERNCPVCGTE ATPQGSKGNASGDTAFCPPCLTRQCSEDSTHDQGPGVEGTSRNHK GKINFPHILQSDDCGRSDNLIVYSINLVPGLSFIQLPSGTKHCIIDVSYT FSDTLAGYLIVGGVDGCQLHNKAIIYLSLGYYKTKMIYPPDYIAIATYTY DLVPNLRDCSIAVNQTSLAAICTSKKTKENQDFSTSGVHPFYIFTLNTD GIFTVTVIEQSQLKLDYQYAALYPATGPGIFIGDHLVFLMWGGLMTKA EGDAYCQASGCNDAHRTSCNIAQMPSAYGHRQLVNGLLMLPIKELG SHLIQPSLETISPKINWAGGHGRLYYNWEINTTYIYIEGKTWRSRPNL GIISWSKPLSIRWIDHSVARRPGARPCDSANDCPEDCLVGGYYDMFP MSSDYKTAITIIPTHHQWPSSPALKLFNTNREVRVVMILRPPNNVKKT TISCIRIMQTNWCLGFIIFKEGNNAWGQIYSYIYQVESTCPNTK	1	123

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
AY590688	6138-7935	gb:AY590688:6138-7935|Organismo:Metapneumovírus aviário|Nome da Cepa:Colorado|Nome da Proteína:glicoproteína de fixação|Símbolo do Gene:G	MEVKVENVGKSQELKVKVKNFIKRSDCKKKLFALILGLVSFELTMNIM LSVMYVESNEALSLCRIQGTPAPRDNKTNTENATKETTLHTTTTTRDP EVRETKTTKPQANEGATNPSRNLTTKGDKHQTTRATTEAELEKQSK QTTEPGTSTQKHTPTRPSSKSPTTTQAIAQLTTPTTPKASTAPKNRQA TTKKTETDTTTASRARNTNNPTETATTTPKATTETGKSKEGPTQHTT KEQPETTAGETTTPQPRRTASRPAPTTKIEEEAETTKTRTTKSTQTST GPPRPTGGAPSGAATEGSGRAAAAGGPSAASAGGRRRTEAAAERD RRTRAGAGPTAGGARARTAAASERGADTAGSAGGGPGGDGATGGL SGGAPAEREDASGGTAAAGPGDGTEADGRAPPAAALAGRTTESAA GAAGDSGRAGTAGWGSAADGRSTGGNAAAEAGAAQSGRAAPRQP SGGTAPESTAPPNSGGSGRADAAPTEEVGVGSGLWRGRYVCGPCG ESVPEHPMNPCFGDGTAWICSDDGGSLPAGCYDGGTDGVVCCGVC GGNSCCCGRVECTCGGGAGLLSCCCGSYSWS	1	124
EU403085	6620-8593	gb:EU403085:6620-8593|Organismo:Paramixovírus aviário 3|Nome da cepa:APMV3/PKT/Netherland/449/75|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MESPPSGKDAPAFREPKRTCRLCYRATTLSLNLTIVVLSIISIYVSTQT GANNSCVNPTIVTPDYLTGSTTGSVEDLADLESQLREIRRDTGINLPV QIDNTENLILTTLASINSNLRFLQNATTESQTCLSPVNDPRFVAGINRIP AGSMAYNDFSNLIEHVNFIPSPTTLSGCTRIPSFSLSKTHWCYTHNVIS NGCLDHAASSQYISIGIVDTGLNNEPYFRTMSSKSLNDGLNRKSCSVT AAANACWLLCSVVTEYEAADYRSRTPTAMVLGRFDFNGEYTEIAVPS SLFDGRFASNYPGVGSGTQVNGTLYFPLYGGVLNGSDIETANKGKSF RPQNPKNRCPDSEAIQSFRAQDSYYPTRFGKVLIQQAIIACRISNKSC TDFYLLYFDNNRVMMGAEARLYYLNNQLYLYQRSSSWWPHPLFYSI SLPSCQALAVCQITEAHLTLTYATSRPGMSICTGASRCPNNCVDGVY TDVWPLTKNDAQDPNLFYTVYLNNSTRRISPTISLYTYDRRIKSKLAV GSDIGAAYTTSTCFGRSDTGAVYCLTIMETVNTIFGQYRIVPILLRVTS R	1	125

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
FJ977568	6139-7936	gb:FJ977568:6139-7936|Organismo:metapneumovírus aviário|Nome da cepa:aMPV/MN/turkey/2a/97|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MEVKVENVGKSQELKVKVKNFIKRSDCKKKLFALILGLVSFELTMNIM LSVMYVESNEALSLCRIQGTPAPRDNKTNTENATKETTLHTTTTTRDP EVRETKTTKPQANEGATNPSRNLTTKGDKHQTTRATTEAELEKQSK QTTEPGTSTQKHTPARPSSKSPTTTQATAQPTTPTAPKASTAPKNRQ ATTKKTETDTTTASRARNTNNPTETATTTPKATTETGKGKEGPTQHT TKEQPETTARETTTPQPRRTASRPAPTTKIEEEAETTKTRTTKNTQTS TGPPRPTRSTPSKTATENNKRTTTTKRPNTASTDSRQQTRTTAEQD QQTQTRAKPTTNGAHPQTTTTPEHNTDTTNSTKGSPKEDKTTRDPS SKTPTEQEDASKGTAAANPGGSAEADRRAPPATTPTGRTTESAAGT TGDDSGAETTRRRSAADRRPTGGSTAAEAGTAQSGRATPKQPSGG TAAGNTAPPNNESSGRADAAPAEEAGVGPSIRRGRHACGPRRESAP EHPTNPCPGDGTAWTRSDGGGNLPAGRHDSGADGAARRGARGGN PRRRGRAERTRGGGAGPPSCRCGSHNRS	1	126
HG934339	5997-7166	gb:HG934339:5997-7166|Organismo:metapneumovírus aviário tipo D|Nome da cepa:Turkey/1985/Fr85.1|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MGAKLYAISGASDAQLMKKTCAKLLEKVVPIIILAVLGITGTTTIALSISIS IERAVLSDCTTQLRNGTTSGSLSNPTRSTTSTAVTTRDIRGLQTTRTR ELKSCSNVQIAYGYLHDSSNPVLDSIGCLGLLALCESGPFCQRNYNP RDRPKCRCTLRGKDISCCKEPPTAVTTSKTTPWGTEVHPTYPTQVTP QSQPATMAHQTATANQRSSTTEPVGSQGNTTSSNPEQQTEPPPSP QHPPTTTSQDQSTETADGQEHTPTRKTPTATSNRRSPTPKRQETGR ATPRNTATTQSGSSPPHSSPPGVDANMEGQCKELQAPKPNSVCKGL DIYREALPRGCDKVLPLCKTSTIMCVDAYYSKPPICFGYNQRCFCMET FGPIEFCCKS	1	127
JN032116	4659-5252	gb:JN032116:4659-5252|Organismo:Vírus sincicial respiratório|Nome da cepa:B/WI/629-12/06-07|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MSKNKNQRTARTLEKTWDTLNHLIVISSCLYKLNLKSIAQIALSVLAMII STSLIIAAIIFIISANHKVTLTTVTVQTIKNHTEKNITTYLTQVSPERVSPS KQPTTTPPIHTNSATISPNTKSEIHHTTAQTKGRTSTPTQNNKPNTKP RPKNPPKKDDYHFEVFNFVPCSICGNNQLCKSICKTIPSNKPRKNQP	1	128

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KX258200	6254-7996	gb:KX258200:6254-7996|Organismo:Paramixovírus aviário 14|Nome da cepa:APMV14/duck/Japan/11OG0352/2011|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MEGSRTVIYQGDPNEKNTWRLVFRTLTLILNLAILSVTIASIIITSKITLSEVTTLKTEGVEEVITPLMATLSDSVQQEKMIYKEVAISIPLVLDKIQTDVGTSVAQITDALRQIQGVNGTQAFALSNAPEYSGGIEVPLFQIDSFVNKSMSISGLLEHASFIPSPTTLHGCTRIPSFHLGPRHWCYTHNIIGSRCRDEGFSSMYISIGAITVNRDGNPLFITTASTILADDNNRKSCSIIASSYGCDLLCSIVTESENDDYANPNPTKMVHGRFLYNGSYVEQALPNSLFQDKWVAQYPGVGSGITTHGKVLFPIYGGIKKNTQLFYELSKYGFFAHNKELECKNMTEEQIRDIKAAYLPSKTSGNLFAQGIIYCNISKLGDCNVAVLNTSTTMMGAEGRLQMMGEYVYYYQRSSSWWPVGIVYKKSLAELMNGINMEVLSFEPIPLSKFPRPTWTAGLCQKPSICPDVCVTGVYTDLFSVTIGSTTDKDTYFGVYLDSATERKDPWVAAADQYEWRNRVRLFESTTEAAYTTSTCFKNTVNNRVFCVSIVELRENLLGDWKIVPLLFQIGVSQGPPPK	1	129

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KX940961	7978-12504	gb:KX940961:7978-12504|Organismo:Vírus Beilong|Nome da cepa:ERN081008_1S|Nome da proteína:glicoproteína de fixação|Símbolo do gene:G	MSQLAAHNLAMSNFYGTHQGDLSGSQKGEEQQVQGVIRYVSMIVGL LSLFTIIALNVTNIIYMTESGGTMQSIKTAQGSIDGSMREISGVIMEDVK PKTDLINSMVSYNIPAQLSMIHQIIKNDVLKQCTPSFMFNNTICPLAEN PTHSRYFEEVNLDSISECSGPDMHLGLGVNPEFIEFPSFAPGSTKPG SCVRLPSFSLSTTVFAYTHTIMGHGCSELDVGDHYFSVGRIADAGHEI PQFETISSWFINDKINRRSCTVAAGAMEAWMGCVIMTETFYDDLNSL DTGKLTISYLDVFGRKKEWIYTRSEILYDYTYTSVYFSVGSGVVVGDT VYFLIWGSLSSPIEETAYCFAPDCSNYNQRMCNEAQRPSKFGHRQM VNGILKFKTTSTGKPLLSVGTLSPSVVPFGSEGRLMYSEITKIIYLYLRS TSWHALPLTGLFVLGPPTSISWIVQRAVSRPGEFPCGASNRCPKDCV TGVYTDLFPLGSRYEYAATVYLNSETYRVNPTLALINQTNIIASKKVTT ESQRAGYTTTTCFVFKLRVWCISVVELAPSTMTAYEPIPFLYQLDLTC KGKNGSLAMRFTGKEGTYKSGRYKSPRNECFFEKVSNKYYFIVSTPE GIQPYEIRDLTPDRMPHIIMYISDVCAPALSAFKKLLPAMRPITTLTIGN WQFRPVEVSGGLRVSIGRNLTKEGDLTMSAPEDPGSNTFPGGHIPG NGLFDAGYYTVEYPKEWKQTTPKPSEGGNIIDKNKTPVIPSRDNPTS DSSIPHRESIEPVRPTREVLKSSDYVTIVSTDSGSGSGDFATGVPWT GVSPKAPQNGINLPGTELPHPTVLDRINTPAPSDPKVSADSDHTRDTI DPTALSKPLNHDTTGDTDTRINTGTATYGFTPGREATSSGKLANDLT NSTSVPSEAHPSASTSEASKPEKNTDNRVTQDPTSGTAERPTTNAPV DGKHSTQLTDARPNTADPERTSQHSSSTTRDEVKPSLPSTTEASTH QRTEAATPPELVNNTLNPPSTQVRSVRSLMQDAIAQAWNFVRGVTP	1	130

ID de Genbank	Nucleotídeos de CDS	ID da sequência completa	Sequência	n° de Sequências/Agrupamento	SEQ ID NO
KY511044	6454-8310	gb:KY511044:6454-8310|Organismo:paramixovírus aviário UPO216|Nome da cepa:APMV-15/WB/Kr/UPO216/2014|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MERGISEVALANDRTEEKNTWRLIFRITVLVVSVITLGLTAASLVYSMN AAQPADFDGIIPAVQQVGTSLTNSIGGMQDVLDRTYKQVALESPLTLL NMESTIMNAITSLSYKINNGGNSSGCGAPIHDPEYIGGIGKELLIDDNV DVTSFYPSAFKEHLNFIPAPTTGAGCTRIPSFDLSATHYCYTHNVILSG CQDHSHSHQYIALGVLKLSDTGNVFFSTLRSINLDDTANRKSCSISAT PLGCDILCSKVTETELEDYKSEEPTPMVHGRLSFDGTYSEKDLDVNN LFSDWTANYPSVGGGSYIGNRVWYAVYGGLKPGSNTDQSQRDKYVI YKRYNNTCPDPEDYQINKAKSSYTPSYFGSKRVQQAILSIAVSPTLGS DPVLTPLSNDVVLMGAEGRVMHIGGYTYLYQRGTSYYSPALLYPLNI QDKSATASSPYKFDAFTRPGSVPCQADARCPQSCVTGVYTDPYPLIF AKDHSIRGVYGMMLNDVTARLNPIAAVFSNISRSQITRVSSSSTKAAY TTSTCFKVIKTNRIYCMSIAEISNTLFGEFRIVPLLVEILSNGGNTARSA GGTPVKESPKGWSDAIAEPLFCTPTNVTRYNADIRRYAYSWP	1	131
NC_025360	8127-10158	gb:NC_025360:8127-10158|Organismo:Paramixovírus do salmão do Atlântico|Nome da cepa:ASPV/Yrkje371/95|Nome da proteína:proteína hemaglutinina-neuraminidase|Símbolo do gene:HN	MPPAPSPVHDPSSFYGSSLFNEDTASRKGTSEEIHLLGIRWNTVLIVL GLILAIIGIGIGASSFSASGITGNTTKEIRLIVEEMSYGLVRISDSVRQEIS PKVTLLQNAVLSSIPALVTTETNTIINAVKNHCNSPPTPPPPTEAPLKK HETGMAPLDPTTYWTCTSGTPRFYSSPNATFIPGPSPLPHTATPGGC VRIPSMHIGSEIYAYTSNLIASGCQDIGKSYQNVQIGVLDRTPEGNPE MSPMLSHTFPINDNRKSCSIVTLKRAAYIYCSQPKVTEFVDYQTPGIE PMSLDHINANGTTKTWIYSPTEVVTDVPYASMYPSVGSGVVIDGKLV FLVYGGLLNGIQVPAMCLSPECPGIDQAACNASQYNQYLSGRQVVN GIATVDLMNGQKPHISVETISPSKNWFGAEGRLVYMGGRLYIYIRSTG WHSPIQIGVIYTMNPLAITWVTNTVLSRPGSAGCDWNNRCPKACLSG VYTDAYPISPDYNHLATMILHSTSTRSNPVMVYSSPTNMVNYAQLTTT AQIAGYTTTSCFTDNEVGYCATALELTPGTLSSVQPILVMTKIPKECV	1	132

OUTRAS PROTEÍNAS

[00375] Em algumas modalidades, o fusogênio pode incluir uma proteína dependente de pH, um homólogo da mesma, um fragmento da mesma e uma fusão de proteína que compreende uma ou mais proteínas ou fragmentos das mesmas. Os fusogênios podem mediar a fusão da membrana na superfície da célula ou em um endossomo ou em outro espaço ligado à membrana da célula.

[00376] Em algumas modalidades, o fusogênio inclui um EFF-1, AFF-1, proteína de junção de lacuna, por exemplo, uma conexina (tal como Cn43, GAP43, CX43) (DOI: 10.1021/jacs.6b05191), outras proteínas de conexão de tumor, um homólogo da mesma, um fragmento da mesma, uma variante da mesma e uma fusão de proteínas que compreende uma ou mais proteínas ou fragmentos das mesmas.

FUSOGÊNIOS LIPÍDICOS

[00377] Em algumas modalidades, o fusossoma pode compreender um ou mais lipídios fusogênicos, tais como ácidos graxos saturados. Em algumas modalidades, os ácidos graxos saturados têm entre 10 a 14 carbonos. Em algumas modalidades, os ácidos graxos saturados têm ácidos carboxílicos de cadeia mais longa. Em algumas modalidades, os ácidos graxos saturados são monoésteres.

[00378] Em algumas modalidades, o fusossoma pode compreender um ou mais ácidos graxos não saturados. Em algumas modalidades, os ácidos graxos insaturados têm entre C16 e C18 ácidos graxos insaturados. Em algumas modalidades, os ácidos graxos insaturados incluem ácido oleico, mono-oleato de glicerol, glicerídeos, diacilglicerol, ácidos graxos insaturados modificados e qualquer combinação dos mesmos.

[00379] Sem se ater a nenhuma teoria, em algumas modalidades, os lipídios de curvatura negativa promovem a fusão da membrana. Em algumas modalidades, o fusossoma compreende um ou mais lipídios

de curvatura negativa, por exemplo, lipídios de curvatura negativa exogénos, na membrana. Nas modalidades, o lipídio de curvatura negativa ou um precursor do mesmo é adicionado ao meio que compreende células de origem ou fusossomas. Nas modalidades, a célula de origem é projetada para expressar ou superexpressar um ou mais genes de síntese de lipídios. O lipídio de curvatura negativa pode ser, por exemplo, diacilglicerol (DAG), colesterol, ácido fosfatídico (PA), fosfatidiletanolamina (PE) ou ácido graxo (FA).

[00380] Sem se ater a nenhuma teoria, em algumas modalidades, os lipídios de curvatura positiva inibem a fusão da membrana. Em algumas modalidades, o fusossoma compreende reduziu os níveis de um ou mais lipídios de curvatura positiva, por exemplo, lipídios de curvatura positiva exógena, na membrana. Nas modalidades, os níveis são reduzidos pela inibição da síntese do lipídeo, por exemplo, por knockout ou knockdown de um gene de síntese de lipídeos, na célula de origem. O lipídeo de curvatura positiva pode ser, por exemplo, lisofosfatidilcolina (LPC), fosfatidilinositol (PtdIns), ácido lisofosfatídico (LPA), lisofosfatidiletanolamina (LPE) ou monoacilglicerol (MAG).

FUSOGÊNIOS QUÍMICOS

[00381] Em algumas modalidades, o fusossoma pode ser tratado com produtos químicos fusogênicos. Em algumas modalidades, o produto químico fusogênico é polietilenoglicol (PEG) ou derivados do mesmo.

[00382] Em algumas modalidades, o fusogênio químico induz uma desidratação sítio entre as duas membranas que leva a um empacotamento molecular desfavorável da bicamada. Em algumas modalidades, o fusogênio químico induz a desidratação de uma área próxima à bicamada lipídica, causando o deslocamento de moléculas aquosas entre duas membranas e permitindo a interação entre as duas membranas juntas.

[00383] Em algumas modalidades, o fusogênio químico é um cátion positivo. Alguns exemplos não limitativos de cátions positivos incluem Ca2+, Mg2+, Mn2+, Zn2+, La3+, Sr3+ e H+.

[00384] Em algumas modalidades, o fusogênio químico se liga à membrana alvo, modificando a polaridade da superfície, o que altera a repulsão entre membranas dependente da hidratação.

[00385] Em algumas modalidades, o fusogênio químico é um solúvel de lipídeo solúvel. Alguns exemplos não limitantes incluem oleoilglicerol, dioleoilglicerol, trioleoilglicerol e variantes e derivados dos mesmos.

[00386] Em algumas modalidades, o fusogênio químico é um produto químico solúvel em água. Alguns exemplos não limitantes incluem polietilenoglicol, dimetilsulfóxido e variantes e derivados dos mesmos.

[00387] Em algumas modalidades, o fusogênio químico é uma pequena molécula orgânica. Um exemplo não limitante inclui brometo de n-hexila.

[00388] Em algumas modalidades, o fusogênio químico não altera a constituição, a viabilidade celular ou as propriedades de transporte de íons do fusogênio ou membrana alvo.

[00389] Em algumas modalidades, o fusogênio químico é um hormônio ou uma vitamina. Alguns exemplos não limitantes incluem ácido abscísico, retinol (vitamina A1), um tocoferol (vitamina E) e variantes e derivados dos mesmos.

[00390] Em algumas modalidades, o fusossoma compreende actina e um agente que estabilizar a actina prolimerizada. Sem se ater a qualquer teoria, a actina estabilizada em um fusossoma pode promover fusão com uma célula-alvo. Nas modalidades, o agente que estabiliza a actina polimerizada é escolhido a partir de actina, miosina, biotina-estreptavidina, ATP, proteína da síndrome de Wiskott-Aldrich neuronal (N-WASP) ou formina. Consultar, por exemplo, Langmuir. 16 de

agosto de 2011;27(16):10.061 a 10.071 e Wen et al., Nat Commun. 31 de agosto de 2016;7. Nas modalidades, o fusossoma compreende actina exógena, por exemplo, actina do tipo selvagem ou actina que compreende uma mutação que promove polimerização. Nas modalidades, o fusossoma compreende ATP ou fosfocreatina, por exemplo, ATP exógena ou fosfocreatina.

FUSOGÊNIOS DE MOLÉCULAS PEQUENAS

[00391] Em algumas modalidades, o fusossoma pode ser tratado com moléculas pequenas fusogênicas. Alguns exemplos não limitantes incluem halotano, fármacos anti-inflamatórios não esteroidais (NSAIDs), tais como meloxicam, piroxicam, tenoxicam e clorpromazina.

[00392] Em algumas modalidades, o fusogênio de molécula pequena pode estar presente em agregados semelhantes a micelas ou livre de agregados.

MODIFICAÇÕES EM FUSOGÊNIOS DE PROTEÍNA

[00393] Proteína fusogênica ou proteínas do envelope viral podem ser realvejadas por mutação de resíduos de aminoácidos em uma proteína de fusão ou uma proteína de alvejamento (por exemplo, a proteína hemaglutinina). Em algumas modalidades, o fusogênio é sofre mutação aleatória. Em algumas modalidades, o fusogênio sofre mutação racional. Em algumas modalidades, o fusogênio é submetido à evolução direcionada. Em algumas modalidades, o fusogênio é truncado e apenas um subconjunto do peptídeo é usado no vetor retroviral ou VLP. Por exemplo, resíduos de aminoácidos na proteína hemaglutinina do sarampo podem sofrer mutação para alterar as propriedades de ligação da proteína, redirecionando a fusão (doi: 10.1038/nbt942, Molecular Therapy vol. 16 no. 8, 1.427 a 1.436, agosto de 2008, doi: 10.1038/nbt1060, DOI: 10.1128/JVI.76.7.3558–3563.2002, DOI: 10.1128/JVI.75.17.8016–8020.2001, doi: 10.1073pnas.0604993103).

[00394] Proteínas fusogênicas podem ser realvejadas por covalência conjugando-se uma porção química de alvejamento à proteína de fusão ou proteína de alvejamento (por exemplo, a proteína hemaglutinina). Em algumas modalidades, o fusogênio e a porção química de alvejamento são conjugados covalentemente pela expressão de uma proteína quimérica que compreende o fusogênio ligado à porção química de alvejamento. Um alvo inclui qualquer peptídeo (por exemplo, um receptor) que é exibido em uma célula-alvo. Em alguns exemplos, o alvo é expresso em níveis mais elevados em uma célula-alvo do que células não alvo. Por exemplo, o fragmento variável de cadeia única (scFv) pode ser conjugado a fusogênios para redirecionar a atividade de fusão para células que exibem o alvo de ligação de scFv (doi:10.1038/nbt1060, DOI 10.1182/blood-2012-11-468579, doi: 10.1038/ nmeth.1514, doi:10.1006/mthe.2002.0550, HUMAN GENE THERAPY 11:817– 826, doi:10.1038/nbt942, doi:10.1371/journal. pone.0026381, DOI 10.1186/s12896-015-0142-z). Por exemplo, as proteínas de repetição de anquirina (DARPin) podem ser conjugadas a fusogênios para redirecionar a atividade de fusão em direção a células que exigem o alvo de ligação de DARPin (doi:10.1038/mt.2013.16, doi:10.1038/mt.2010.298, doi: 10.4049/jimmunol.1500956), assim como combinações de DARPins diferentes (doi:10.1038/mto.2016.3). Por exemplo, ligantes de receptor e antígenos podem ser conjugados a fusogênios para redirecionar a atividade de fusão para células que exibem o receptor alvo (DOI: 10.1089/hgtb.2012.054, DOI: 10.1128/ JVI.76.7.3558–3563.2002). Uma proteína de alvejamento também pode incluir, por exemplo, um anticorpo ou um fragmento de ligação ao antígeno do mesmo (por exemplo, Fab, Fab', F(ab')2, fragmentos Fv, fragmentos de anticorpo scFv, Fvs ligados por dissulfeto (sdFv), um fragmento Fd que consiste nos domínios VH e CH1, anticorpos lineares, anticorpos de domínio único, como sdAb (ou VL ou VH), nanocor-

pos ou domínios VHH de camelídeo), um esqueleto de fibronectina de ligação ao antígeno tipo III (Fn3), como um minicorpo de polipeptídeo de fibronectina, um ligante, uma citocina, uma quimiocina ou um receptor de células T (TCRs). Proteínas fusogênicas podem ser realvejadas por conjugação não covalente de uma porção química de alvejamento à proteína de fusão ou proteína de alvejamento (por exemplo, a proteína hemaglutinina). Por exemplo, a proteína de fusão pode ser projetada para ligar a região Fc de um anticorpo que tem como alvo um antígeno em uma célula-alvo, redirecionando a atividade de fusão para células que exibem o alvo do anticorpo (DOI: 10.1128/ JVI.75. 17.8016–8020.2001, doi: 10.1038/nm1192). Fusogênios alterados e não alterados podem ser exibidos no mesmo vetor retroviral ou VLP (doi: 10.1016/j.biomaterials.2014.01.051).

[00395] Uma porção química de alvejamento pode compreender, por exemplo, uma molécula de anticorpo humanizado, anticorpo IgA, IgG, IgE ou IgM intacto; anticorpo bi ou multiespecífico (por exemplo, Zybodies®, etc.); fragmentos de anticorpo, tais como fragmentos Fab, fragmentos Fab', fragmentos F(ab')2, fragmentos Fd', fragmentos Fd e CDRs isolados ou conjuntos dos mesmos; Fvs de cadeia única; fusões polipeptídeo-Fc; anticorpos de domínio único (por exemplo, anticorpos de domínio único de tubarão, como IgNAR ou fragmentos dos mesmos); anticorpos cameloides; anticorpos mascarados (por exemplo, Probodies®); Imunofarmacêuticos modulares pequenos (“SMIPsTM”); diacorpos de cadeia simples ou Tandem (TandAb®); VHHs; Anticalins®; Nanobodies®; minibodies; BiTE®s; proteínas de repetição de anquirina ou DARPINs®; Avimers®; DARTs; anticorpos do tipo TCR;, Adnectins®; Affilins®; Trans-bodies; Affibodies®; TrimerX®; MicroProteins; Fynomers®, Centyrins® e KALBITOR®s.

[00396] Nas modalidades, o fusogênio realvejado se liga a um marcador de superfície celular na célula-alvo, por exemplo, uma proteína,

glicoproteína, receptor, ligante de superfície celular, agonista, lipídio, açúcar, proteína transmembranar de classe I, proteína transmembranar de classe II ou proteína transmembranar de classe III.

[00397] Os fusossomas pode exibir porções químicas de alvejamento que não são conjugadas a fusogênios de proteína a fim de redirecionar a atividade de fusão em direção a uma célula que é ligada pela porção química tida como alvo ou para afetar o homing.

[00398] A porção química tida como alvo adicionada ao fusossoma pode ser modulada de modo a ter diferentes resistências de ligação. Por exemplo, scFvs e anticorpos com várias resistências de ligação podem ser usados para alterar a atividade de fusão do fusossoma em direção a células que exibe altas ou baixas quantidades do antígeno-alvo (doi:10.1128/JVI.01415-07, doi:10.1038/cgt.2014.25, DOI: 10.1002/jgm.1151). Por exemplo, DARPins com diferentes afinidades podem ser usadas para alterar a atividade de fusão do vetor retroviral ou VLP para células que exibem quantidades altas ou baixas do antígeno-alvo (doi: 10.1038/mt.2010.298). Porções de alvejamento também podem ser moduladas para direcionar diferentes regiões no ligante alvo, o que afetará a taxa de fusão com células exibindo o alvo (doi: 10.1093/protein/gzv005).

[00399] Em algumas modalidades, os fusogênios de proteína podem ser alterados para reduzir a imunorreatividade, por exemplo, conforme descrito no presente documento. Por exemplo, proteínas fusogênicas podem ser demarcadas com moléculas que reduzem interações imunológicas, tais como PEG (DOI: 10.1128/JVI.78.2.912–921.2004). Dito isso, em algumas modalidades, o fusogênio que compreende PEG, por exemplo, é um polipeptídeo PEGuilado. Os resíduos de aminoácidos no fusogênio que são alvejados pelo sistema imunológico podem ser alterados para não serem reconhecidos pelo sistema imunológico (doi: 10.1016/j.virol.2014.01.027, doi: 10.1371/journal.pone.

0046667). Em algumas modalidades, a sequência da proteína do fusogênio é alterada para se parecer com as sequências de aminoácidos encontradas em seres humanos (humanizadas). Em algumas modalidades, a sequência de proteína do fusogênio é alterada para uma sequência de proteína que se liga com menos intensidade a complexos de MHC. Em algumas modalidades, os fusogênios de proteína são derivados de vírus ou organismos que não infectam humanos (e contra os quais os seres humanos não foram vacinados), aumentando a probabilidade de que o sistema imunológico de um paciente seja naïve para os fusogênios de proteína (por exemplo, há uma resposta humoral insignificante ou imunológica adaptativa mediada por células em relação ao fusogênio) (doi: 10.1006/mthe.2002.0550, doi: 10.1371/journal.ppat.1005641, doi: 10.1038/gt.2011.209, DOI 10.1182/blood-2014-02-558163). Em algumas modalidades, a glicosilação do fusogênio pode ser alterada para alterar as interações imunológicas ou reduzir a imunorreatividade. Sem se ater a nenhuma teoria, em algumas modalidades, uma proteína fusogênica derivada de um vírus ou organismo que não infecta humanos não tem alvos de fusão naturais em pacientes e, portanto, tem alta especificidade.

ELEMENTO REGULATÓRIO ESPECÍFICO DA CÉLULA-ALVO POSITIVO

[00400] Em algumas modalidades, um ácido nucleico retroviral descrito no presente documento compreende um elemento regulatório específico de célula-alvo positivo, como um promotor específico de tecido, um intensificador específico de tecido, um sítio de splice específico de tecido, um sítio específico de tecido que estende a meia-vida de um RNA ou proteína, um sítio de promoção de exportação nuclear de mRNA específico de tecido, um sítio de aumento de tradução específico de tecido ou um sítio de modificação pós-tradução específico de tecido.

[00401] Um ácido nucleico retroviral descrito no presente documento pode compreender regiões, por exemplo, regiões não traduzidas, tais como origens de replicação, cassetes de seleção, promotores, intensificadores, sinais de iniciação de tradução (sequência Shine Dalgarno ou sequência Kozak), íntrons, uma sequência de poliadenilação, regiões 5' e 3’ não traduzidas - que interagem com proteínas celulares do hospedeiro para realizar a transcrição e tradução, e que são com capacidade para direcionar, aumentar, regular ou controlar a transcrição ou expressão de um polinucleotídeo ligado operacionalmente. Esses elementos podem ter força e especificidade variáveis. Dependendo do sistema de vetor e hospedeiro usado, qualquer número de elementos de transcrição e tradução adequados, incluindo promotores ubíquos e promotores, induzíveis pode ser usado.

[00402] Nas modalidades particulares, os elementos de controle são com capacidade para direcionar, aumentar, regular ou controlar a transcrição ou expressão de um polinucleotídeo ligado operacionalmente de uma maneira específica para células. Nas modalidades particulares, os ácidos nucleicos retrovirais compreendem uma ou mais sequências de controle de expressão que são específicas para células, tipos de células ou linhagens de células particulares, por exemplo, células-alvo; isto é, a expressão de polinucleotídeos operativamente ligados a uma sequência de controle de expressão específica para células, tipos de células ou linhagens de células particulares é expressa em células-alvo e não (ou em um nível inferior) em células não alvo.

[00403] Nas modalidades particulares, um ácido nucleico retroviral pode incluir sequências de controle exógenas, endógenas ou heterólogas, como promotores e/ou potencializadores.

[00404] Nas modalidades, o promotor compreende um sítio de reconhecimento ao qual uma RNA polimerase se liga. Uma RNA polimerase inicia e transcreve polinucleotídeos ligados operacionalmente ao

promotor. Nas modalidades particulares, os promotores operativos em células de mamíferos compreendem uma região rica em AT localizada aproximadamente 25 a 30 bases a montante do sítio no qual a transcrição é iniciada e/ou outra sequência encontrada 70 a 80 bases a montante do início da transcrição, uma região CNCAAT na qual N pode ser qualquer nucleotídeo.

[00405] Nas modalidades, um intensificador compreende um segmento de DNA que contém sequências com capacidade para fornecer transcrição aprimorada e, em alguns casos, pode funcionar independentemente da orientação em relação a outra sequência de controle. Um intensificador pode funcionar cooperativamente ou aditivamente com promotores e/ou outros elementos intensificadores. Em algumas modalidades, um segmento promotor/intensificador de DNA contém sequências com capacidade para fornecer funções de promotor e intensificador.

[00406] Sequências de controle de expressão ubíquas ilustrativas incluem, porém sem limitação, um promotor inicial imediato de citomegalovírus (CMV), um vírus símio viral 40 (SV40) (por exemplo, precoce ou tardio), um promotor LTR do vírus da leucemia murina Moloney (MoMLV), uma LTR do vírus do sarcoma de Rous (RSV), um promotor do vírus herpes simplex (HSV) (timidina quinase), promotores H5, P7.5 e P11 do vírus vaccinia, um promotor do fator de alongamento 1-alfa (EF1a), resposta de crescimento inicial 1 (EGR1), ferritina H (FerH), ferritina L (FerL), Gliceraldeído 3-fosfato desidrogenase (GAPDH), fator de iniciação da tradução eucariótica 4A1 (EIF4A1), proteína de choque térmico 70 kDa 5 (HSPA5), proteína de choque térmico 90 kDa beta, membro 1 (HSP90B1), proteína de choque térmico 70 kDa (HSP70), β-cinesina (β-KIN), o locus ROSA 26 humano Orions et al., *Nature Biotechnology* 25, 1.477 a 1.482 (2007)), um promotor Ubiquitina C (UBC), um promotor de fosfoglicerato quinase-1 (PGK), um pro-

motor de citomegalovírus/promotor de β-actina de galinha (CAG), um promotor de β-actina e um intensificador do vírus do sarcoma mieloproliferativo, região de controle negativa deletada, promotor do sítio de ligação do iniciador d1587rev substituído (MND) (Challita et al., *J Virol.* 69(2):748 a 55 (1995)).

[00407] Em algumas modalidades, um promotor pode ser emparelhado com um gene heterólogo para transmitir as funções regulatórias desse promotor no gene heterólogo. Em algumas modalidades, os elementos cis-regulatórios de um promotor do primeiro gene podem ser ligados a segmentos de um promotor de gene diferente para criar promotores quiméricos que têm propriedades de ambos os promotores.

[00408] Em algumas modalidades, o promotor é um promotor tecido-específico, por exemplo, um promotor que conduz a expressão em células do fígado, por exemplo, hepatócitos, células endoteliais sinusoidais do fígado, colangiócitos, células estreladas, células apresentadoras de antígeno residentes no fígado (por exemplo, Células de Kupffer), linfócitos imunes resistentes no fígado (por exemplo, células T, célula B ou células NK) ou fibroblastos portais. Vários promotores específicos do fígado adequados (por exemplo, promotores específicos dos hepatócitos e promotores de células endoteliais sinusoidais do fígado) são descritos na Tabela 3 abaixo. A Tabela 3 também lista vários promotores ubíquos que não são específicos a células do fígado. Em algumas modalidades, um fusossoma (por exemplo, vetor viral) descrito no presente documento compreende, em seu ácido nucleico, um promotor tendo uma sequência da Tabela 3, ou fragmento ativo de maneira transcricional da mesma, ou uma variante que tem pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com a mesma. Em algumas modalidades, um fusossoma (por exemplo, vetor viral) descrito no presente documento compreen-

de, em seu ácido nucleico, um promotor com locais de ligação a fator de transcrição da região dentro de 3 kb do sítio de início da transcrição para os genes listados na Tabela 3. Em algumas modalidades, o fusossoma (por exemplo, vetor viral) descrito no presente documento compreende, em seu ácido nucleico, uma região dentro de 2,5 kb, 2 kb, 1,5 kb, 1 kb ou 0,5 kb imediatamente a montante do sítio de início da transcrição de um gene listado na Tabela 3, ou um fragmento ativo de maneira transcricional do mesmo, ou uma variante com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com o mesmo.

[00409] Em algumas modalidades, um fusossoma (por exemplo, vetor viral) descrito no presente documento compreende, em seu ácido nucleico, um promotor que tem uma sequência estabelecida em qualquer uma das SEQ ID NOS: 133 a 142 ou um fragmento ativo de maneira transcricional da mesma, ou uma variante com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com a mesma.

[00410] Em algumas modalidades, o promotor é um promotor, por exemplo, um promotor que conduz a expressão em células do fígado, por exemplo, hepatócitos, células endoteliais sinusoidais do fígado, colangiócitos, células estreladas, células apresentadoras de antígeno residentes no fígado (por exemplo, Células de Kupffer), linfócitos imunes resistentes no fígado (por exemplo, células T, célula B ou células NK) ou fibroblastos portais. Em algumas modalidades, um fusossoma (por exemplo, vetor viral) descrito no presente documento compreende, em seu ácido nucleico, um promotor que tem uma sequência estabelecida em qualquer uma das SEQ ID NOS: 133 a 136 ou 519 a 525, ou fragmento ativo de maneira transcricional da mesma, ou uma variante com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com a mesma. Em algumas modalidades,

o promotor é um ser promotor humano hepatócito-específico (ApoE.HCR-hAAT (hApoE). Em algumas modalidades, o promotor tem uma sequência que tem pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequência à sequência apresentada na SEQ ID NO: 133. Em algumas modalidades, o promotor tem a sequência apresentada na SEQ ID NO: 133.

TABELA 3. PROMOTORES EXEMPLIFICATIVOS, POR EXEMPLO, PROMOTORES ESPECÍFICOS DE HEPATÓCITOS

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos	hAAT (serpina A1	gene da α1 antitripsina (Gene de Serpina 1)	AGATCTTGCTACCAGTGGAACAGCCACTAAGGATTCTGCAGTGAGAGCAGAGGGCCAGCTAAGTGGTAC TCTCCCAGAGACTGTCTGACTCACGCCACCCCCTCCACCTTGGACACAGGACGCTGTGGTTTCTGAGCC AGGTACAATGACTCCTTTCGGTAAGTGCAGTGGAAGCTGTACACTGCCCAGGCAAAGCGTCCGGGCAGC GTAGGCGGGCGACTCAGATCCCAGCCAGTGGACTTAGCCCCTGTTTGCTCCTCCGATAACTGGGGTGA CCTTGGTTAATATTCACCAGCAGCCTCCCCCGTTGCCCCTCTGGATCCACTGCTTAAATACGGACGAGGA CAGGGCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATGTCCCCCTGATCTGCG GCCGTGACTCTCTTAAGGTAGCCTTGCAGAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTT ACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGAT AGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTAC AGCT	519
Hepatócitos	ApoE.HCR-hAAT	Gene de apolipoproteína E/CI, gene de antitripsina α1	gttaggctcagaggcacacaggagtttctgggctcaccctgccccctccaaccccctcagttcccatcctccagcagctgtttgtgtgctgcctctgaagt ccacactgaacaaacttcagcctactcatgtccctaaaatgggcaaacattgcaagcagcaaacagcaaacacacagccctccctgcctgctgac cttggagctggggcagaggtcagagacctctctgggcccatgccacctccaacatccactcgacccctttggaatttcggtggagaggagcagaggt tgtcctggcgtggtttaggtagtgtgagaggggtacccggggatcttgctaccagtggaacagccactaaggattctgcagtgagagcagagggcc agctaagtggtactctcccagagactgtctgactcacgccacccccctccaccttggacacaggacgctgtggtttctgagccaggtacaatgactcctt tcggtaagtgcagtggaagctgtacactgcccaggcaaagcgtccgggcagcgtaggcgggcgactcagatcccagccagtggacttagcccct gtttgctcctccgataactggggtgaccttggttaatattcaccagcagcctcccccgttgcccctctggatccactgcttaaatacggacgaggacagg gccctgtctcctcagcttcaggcaccaccactgacctgggacagtgaatgatcccccctgatctgcggcctcgacggtatcgataagcttgatatcgaat tctagtcgtcgaccactttcacaatctgctagcaacctgaggaggttatcgtacgaaattcgctgtctgcgagggccagctgttggggtgagtactccct ctcaaaagcgggcatgacttctgcgctaagattgtcagtttccaaaaacgaggaggatttgatattcacctggcccgcggtgatgcctttgagggtggc cgcgtccatctggtcagaaaagacaatcttttgttgtcaagcttgaggtgtggcaggcttgagatcgatctgaccatacacttgagtgacaatgacatc cactttgcctttctctccacaggtgtccactcccaggtccaac	133
Hepatócitos	Transtiretina intensificada	Gene transtiretina	CAAATGACTTAGTTTGGCTAAAATGTAGGCTTTTAAAAATGTGAGCACTGCCAAGG GTTTTTCCTTGTTGACCCATGGATCCATCAAGTGCAAACATTTTCTAATGCACTATA TTTAAGCCTGTGCAGCTAGATGTCATTCAACATGAAATACATTATTACAACTTGCAT CTGTCTAAAATCTTGCATCTAAAATGAGAGACAAAAAATCTATAAAAATGGAAAACA TGCATAGAAATATGTGAGGGAGGAAAAAATTACCCCCAAGAATGTTAGTGCACGC AGTCACACAGGGAGAAGACTATTTTTGTTTTGTTTTGATTGTTTTGTTTTGTTTTGG TTGTTTTGTTTTGGTGACCTAACTGGTCAAATGACCTATTAAGAATATTTCATAGAA CGAATGTTCCGATGCTCTAATCTCTCTAGACAAGGTTCATATTTGTATGGGTTACTT ATTCTCTCTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGCAGTCAGATTG GCAGGGATAAGCAGCCTAGCTCAGG	520

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos	Alva	Gene da albumina	ccaccgcggtggcggccgctctagcttccttagcatgacgttccacttttttctaaggtggagcttacttctttgatttgatctttgtg aaacttttggaaattacccatcttcctaagcttctgcttctctcagttttctgcttgctcattccacttttccagctgaccctgcccccct accaacattgctccacaagcacaaattcatccagagaaaataaattctaagttttatagttgtttggatcgcataggtagctaa agaggtggcaacccacacatccttaggcatgagcttgatttttttgatttagaaccttcccctctctgttcctagactacactaca cattctgcaagcatagcacagagcaatgttctactttaattactttcattttcttgtatcctcacagcctagaaaataacctgcgtt acagcatccactcagtatcccttgagcatgaggtgacactacttaacatagggacgagatggtactttgtgtctcctgctctgt cagcagggcactgtacttgctgataccagggaatgtttgttcttaaataccatcattccggacgtgtttgccttggccagttttcc atgtacatgcagaaagaagtttggactgatcaatacagtcctctgcctttaaagcaataggaaaaggccaacttgtctacgtt tagtatgtggctgtagaaagggtatagatataaaaattaaaactaatgaaatggcagtcttacacatttttggcagcttatttaa agtcttggtgttaagtacgctggagctgtcacagctaccaatcaggcatgtctgggaatgagtacacggggaccataagtta ctgacattcgtttcccattccatttgaatacacacttttgtcatggtattgcttgctgaaattgttttgcaaaaaaaaccccttcaaat tcatatatattattttaataaatgaattttaatttatctcaatgttataaaaaagtcaattttaataattaggtacttatatacccaataa tatctaacaatcatttttaaacatttgtttattgagcttattatggatgaatctatctctatatactctatatactctaaaaaagaaga aagaccatagacaatcatctatttgatatgtgtaaagtttacatgtgagtagacatcagatgctccatttctcactgtaataccat ttatagttacttgcaaaactaactggaattctaggacttaaatattttaagttttagctgggtgactggttggaaaattttaggtaag tactgaaaccaagagattataaaacaataaattctaaagttttagaagtgatcataatcaaatattaccctctaatgaaaatat tccaaagttgagctacagaaatttcaacataagataattttagctgtaacaatgtaatttgttgtctattttcttttgagatacagtttt ttctgtctagctttggctgtcctggaccttgctctgtagaccaggttggtcttgaactcagagatctgcttgcctctgccttgcaagt gctaggattaaaagcatgtgccaccactgcctggctacaatctatgttttataagagattataaagctctggctttgtgacatta atctttcagataataagtcttttggattgtgtctggagaacatacagactgtgagcagatgttcagaggtatatttgcttaggggt gaattcaatctgcagcaataattatgagcagaattactgacacttccattttatacattctacttgctgatctatgaaacatagat aagcatgcaggcattcatcatagttttctttatctggaaaaacattaaatatgaaagaagcactttattaatacagtttagatgtg ttttgccatcttttaatttcttaagaaatactaagctgatgcagagtgaagagtgtgtgaaaagcagtggtgcagcttggcttga actcgttctccagcttgggatcgacctgcaggcatgcttccatgccaaggcccacactgaaatgctcaaatgggagacaa agagattaagctcttatgtaaaatttgctgttttacataacttttaatgaatggacaaagtcttgtgcatgggggtgggggtgggg ttagaggggaacagctccagatggcaaacatacgcaagggatttagtcaaacaacttttttggcaaagatggtatgattttgt aatggggtaggaaccaatgaaatgcgaggtaagtatggttaatgatctacagttattggttaaagaagtatattagagcgag tctttctgcacacagatcacctttcctatcaaccccgggatccccgggctgcaggaattcgatatcaagcttatcgataccgt cgacctcgaggggggggcccggtac	134

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos (por exemplo, hepatócitos de progenitores de hepatócitos)	Apoa2	Gene da apolipoproteína A-II	CCGGGCGTGGTGGCGCATGTCTGTAATCCCAGCTACTTGGGATGCTGAGGCA GGAGAATCCTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCCGAGATCATGCC ATTACGCTCCAGCCTGAGCAACAAGAGCAAAACTCCGTCTCAGGAAAACAAACA AAAAAACCTGCACATATACTTCTGAATTTAAAACAAAAGTTAAAAAACAAAGATTT CTTGGTCTCTGGTCACTACCTCCCTCATCAGCTTTGCGCCTCCACTGTCACCCT CAGGAATGTTCCACATACTCAGCGAGTATGCTTGGGGGGCAAAAGGGTGAAAG ATACAAAAGCTTCTGATATCTATTTAACTGATTTCACCCAAATGCTTTGAACCTG GGAATGTACCTCTCCCCCTCCCCCACCCCCAACAGGAGTGAGACAAGGGCCA GGGCTATTGCCCCTGCTGACTCAATATTGGCTAATCACTGCCTAGAACTGATAA GGTGATCAAATGACCAGGTGCCTTCAACCTTTACCCTGGTAGAAGCCTCTTATT CACCTCTTTTCCTGCCAGAGCCCTCCATTGGGAGGGGACGGGCGGAAGCTGTT TTCTGAATTTGTTTTACTGGGGGTAGGGTATGTTCAGTGATCAGCATCCAGGTC ATTCTGGGCTCTCCTGTTTTCTCCCCGTCTCATTACACATTAACTCAAAAACGGA CAAGATCATTTACACTTGCCCTCTTACCCGACCCTCATTCCCCTAACCCCCATA GCCCTCAACCCTGTCCCTGATTTCAATTCCTTTCTCCTTTCTTCTGCTCCCCAAT ATCTCTCTGCCAAGTTGCAGTAAAGTGGGATAAGGTTGAGAGATGAGATCTACC CATAATGGAATAAAGACACCATGAGCTTTCCATGGTATGATGGGTTGATGGTAT TCCATGGGTTGATATGTCAGAGCTTTCCAGAGAAATAACTTGGAATCCTGCTTC CTGTTGCACTCAAGTCCAAGGACCTCAGATCTCAAAAGAATGAACCTCAAATAT ACCTGAAGTGTACCCCCTTAGCCTCCACTAAGAGCTGTACCCCCTGCCTCTCAC CCCATCACCATGAGTCTTCCATGTGCTTGTCCTCTCCTCCCCCATTTCTCCAAC TTGTTTATCCTCACATAATCCCTGCCCCACTGGGCCCATCCATAGTCCCTGTCA CCTGACAGGGGGTGGGTAAACAGACAGGTATATAGCCCCTTCCTCTCCAGCCA GGGCAGGCACAGACACCAAGGACAGAGACGCTGGCTAGGTAAGATAAGGAGG CAAGATGTGTGAGCAGCATCCAAAGAGGCCTGGGCTTCAGTTGTGGAGAGGGA GAGAGCCAGGTTGGAATGGGCAGCAGGTAGGGAGATCCCTGGGGAGGAGCTG AAGCCCATTTGGCTTCAGTGTCCCCCAAACCCCCACCACCCT	521

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos (por exemplo, hepatócitos maduros)	Cyp3a4	Gene Cyp3a4	AGCTCCTGGGGCCTGCCCTCCTCCCATTAGAAAATCCTCCACTTGTCAAAAAGG AAGCCATTTGCTTTGAACTCCAATTCCACCCCCAAGAGGCTGGGACCATCTTAT TGGAGTCCTTGATGCTGTGTGACCTGCAGTGACCACTGCCCCATCATTGCTGG CTGAGGTGGTTGGGGTCCATCTGGCTATCTGGGCAGCTGTTCTCTTCTCTCCTT TCTCTCCTGTTTCCAGACATGCAGTATTTCCAGAGAGAAGGGGCCACTCTTTGG CAAAGAACCTGTCTAACTTGCTATCTATGGCAGGACCTTTGAAGGGTTCACAGG AAGCAGCACAAATTGATACTATTCCACCAAGCCATCAGCTCCATCTCATCCATG CCCTGTCTCTCCTTTAGGGGTCCCCTTGCCAACAGAATCACAGAGGACCAGCC TGAAAGTGCAGAGACAGCAGCTGAGGCACAGCCAAGAGCTCTGGCTGTATTAA TGACCTAAGAAGTCACCAGAAAGTCAGAAGGGATGACATGCAGAGGCCCAGCA ATCTCAGCTAAGTCAACTCCACCAGCCTTTCTAGTTGCCCACTGTGTGTACAGC ACCCTGGTAGGGACCAGAGCCATGACAGGGAATAAGACTAGACTATGCCCTTG AGGAGCTCACCTCTGTTCAGGGAAACAGGCGTGGAAACACAATGGTGGTAAAG AGGAAAGAGGACAATAGGATTGCATGAAGGGGATGGAAAGTGCCCAGGGGAG GAAATGGTTACATCTGTGTGAGGAGTTTGGTGAGGAAAGACTCTAAGAGAAGG CTCTGTCTGTCTGGGTTTGGAAGGATGTGTAGGAGTCTTCTAGGGGGCACAGG CACACTCCAGGCATAGGTAAAGATCTGTAGGTGTGGCTTGTTGGGATGAATTTC AAGTATTTTGGAATGAGGACAGCCATAGAGACAAGGGCAGGAGAGAGGCGATT TAATAGATTTTATGCCAATGGCTCCACTTGAGTTTCTGATAAGAACCCAGAACCC TTGGACTCCCCAGTAACATTGATTGAGTTGTTTATGATACCTCATAGAATATGAA CTCAAAGGAGGTCAGTGAGTGGTGTGTGTGTGATTCTTTGCCAACTTCCAAGGT GGAGAAGCCTCTTCCAACTGCAGGCAGAGCACAGGTGGCCCTGCTACTGGCT GCAGCTCCAGCCCTGCCTCCTTCTCTAGCATATAAACAATCCAACAGCCTCACT GAATCACTGCTGTGCAGGGCAGGAAAGCTCCATGCA	522
Hepatócitos	LP1B	Gene de apolipoproteína E/CI, gene de antitripsina α1	cggcctctagactcgagccctaaaatgggcaaacattgcaagcagcaaacagcaaacacacagccctccctgcct gctgaccttggagctggggcagaggtcagagacctctctgggcccatgccacctccaacatccactcgacccctttgg aatttcggtggagaggagcagaggttgtcctggcgtggtttaggtagtgtgagagggtggacacaggacgctgtggtt tctgagccaggggcgactcagatcccagccagtggacttagcccctgtttgctcctccgataactggggtgaccttg gttaatattcaccagcagcctcccccgttgcccctctggatccactgcttaaatacggacgaggacagggccctgtctc ctcagcttcaggcaccaccactgacctgggacagtgaatccggactctaaggtaaatataaaatttttaagtgtataat gtgttaaactactgattctaattgtttctctcttttagattccaacctttggaactgaaccggt	135

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos (por exemplo, hepatócitos de células hepáticas embrionárias em estágio inicial e endoderme)	MIR122	microRNA-122	GAATGCATGGTTAACTACGTCAGAAATGACCAGTTCAAGAGGAGAATGAGATTG GCTTCCAAATGTTGGTCAAGAGCTCTACGTAGCATGAGCCAAGGATCTATTGAA CTTAGTAGGCTCCTGTGACCGGTGACTCTTCTGTCTCTAGAAATCTGGGGAGGT GACCAGGTCATACATGGCAGTCTTCCCGTGAGGAACGTTAAACTGGTTGGAAG TTGGGGTTCTGAGGGGAAGATGTATTCACTAGGTGACCTGTCTTCTCTGCCTCG GTGGCCTCCATGGCTGCCTGCTGGCCGCACACCCCCACTCAGCAGAGGAATG GACTTTCCAATCTTGCTGAGTGTGTTTGACCAAAGGTGGTGCTGACTTAGTGGC CTAAGGTCGTGCCCTCCCTCCCCCACTGAATCGATAAATAATGCGACTTATCAG AAAGAGAAAGAATTGTTTACTTTTAAACCCTGGATCCCATAAAGGGAGAGGGGA GAGGCCTAAAGCCACAGAAGCTGTGGAAGGCGCCATCCTGCCTGCCACAGGA AGGGCCTTGGACTGAGAGGACCGGAGCTGACTGGGGGTAAGTGCGGCTCTCC CCCGGCGCCTGCCGACCCCCCTGAGTGATCAGGCCGTTCTTTGGGGTGGCCG CTGACCGAGAAATGACGGGAGG Consultar Li et al., 2011, J. Hepatol., 55:602 a 611	523

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
Hepatócitos	hemopexina	Gene da hemopexina	GCAGCTTTGGGAGTGGGCCCAGGAAGTACTGAGGATAGCAGGTGAGATCCCA GGAAGAGATGGATGTGGGGCCGAGACACTGGAGAGAGAAACAGGACTGTCAG ATAAAGGGCGTCTGTGACTCCTAGATCTCATTATGCCTACTACCATAACCTACC CCCAATTCCTAATATTCTCCTACCCTAGAGGGGGGGAAATTGTCAGAAATTTGG CTGCAACACTAGCAACACTACTCAGTACTTGAAATGCATTTTTGCATTTTTTTCA TTCAACAAATATTTCTGGAACAACTCTTATATGCCAGGCACTATTTTAGGAGTCA GGGATATATAATGGTAAACAAGACAGGCAAAACAAAGCAAAGCAACAACAACCA TCACCAGATAAGTAGACAGATGAAAGAATTTCAAGTTTTAGTAAGTAAAATAAAA CAAGCAAGGGTCTGAAATGGCTAGATAAGGTGGTCAAGAAAGGCTTCATTGAG AAGGTAGCATTTAAGCAGGAGTCAGCTAGAAATATTGTGAAATTCCAGTTACAG TTCTATTTGTTCTGGGTTGGTTAAATAAAGCTTTTTCCCCCAAGGTGGAAACTAC CAAGAAAGACTAATTACTAGTAGTGGTGGTGCTCTCTGGAAGAGAGACACCTCC TGTTTCTGCCTCATTACTGTCAACCCTTCACTTCCAGGCACTTTTTGCAAAGCCC TTTGCCAGTCAGGGAAGGCGAGAGGCTGGGCATGGGGCTTGGACATTTGACA ACAGTGAGACATTATTGTCCCCAGACTCACTAGCCCAAGGGTAAAGCTGAAGA GGCTTGGGCATGCCCCAGAAAGGCCCCTGATGAAGCTTGGAAAAAGCTGTTCT CTGAGTATTTCTAAGTAAGTTTATCTGTGTGTGTGGTTACTAAAAGTAGTAAGTA TTGCTGTCTCTAGCTGCCTTAGAGCAGGGCTTGACACAGTACACAGCAATATTA GTTCCCTCCTTTTCTCACCTCCCCCATTGTGGAGATAAACTCAATCACAAAAGG TGATCCTCAGTCTACTCACTTCCCTGACTTATGGATGCCTGGACCCATTGCCAG TGTGAGAGTCACAGCTGGACGTCAGCAGTGTAGCCCAGTTACTGCTTGAAAATT GCTGAAGGGGGTTGGGGGGCAGCTGCCGGGAAAAAGGAGTCTTGGATTCAGA TTTCTGTCCAGACCCTGACCTTATTTGCAGTGATGTAATCAGCCAATATTGGCTT AGTCCTGGGAGACAGCACATTCCCAGTAGAGTTGGAGGTGGGGGTGGTGCTG CTGCCAACT	524
Hepatócitos	HLP	Apolipoproteína, SERINA1	tgtttgctgcttgcaatgtttgcccattttagggtggacacaggacgctgtggtttctgagccaggggcgactcagatcc cagccagtggacttagcccctgtttgctcctccgataactggggtgaccttggttaatattcaccagcagcctcccccgtt gcccctctggatccactgcttaaatacggacgaggacagggccctgtctcctcagcttcaggcaccaccactgacct gggacagtgaatc	136

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
células endoteliais sinusoidais do fígado	VEC	Gene de caderina endotelial vascular	CCCCTGCCCTCCTCCTCTGCCCTCTCCTGGCATTCCTCCTTCATCATGGGACCC TCTTCTAATGGATCCCCAAATGTCAGAGGGTCCAAGTCCTCCCTCCCTCCAAGC TCATCCATGCCCATGGCCTCAGATGCCAGCCATAAGCTGTTGGGTTCCAAACCT CGACTCCAGGCTGGACTCACCCCTGTCTCCCCCACCAGCCTGACACCTCCACC TGGGTATCTAACGAGCATCTCAAACTCAACCTGCCTGAGACAGAGGAATCACTA TCCCCTCCTCCTCCAAAAATATCCTTCCATCACACTCCCCATCTTGTGCTCTGAT TTACTAAACGGCCCTGGGCCCTCTCTTTCTCAGGGTCTCTGCTTGCCCAGCTAT ATAATAAAACAAGTTTGGGACTTCCCAACCATTCACCCATGGAAAAACAGAAGC AACTCTTCAAAGGACAGATTCCCAGGATCTGCCCTGGGAGATTCCAAATCAGTT GATCTGGGGTGAGCCCAGTCCTCTGTAGTTTTTAGAAGCTCCTCCTATGTCTCT CCTGGTCAGCAGAATCTTGGCCCCTCCCTTCCCCCCAGCCTCTTGGTTCTTCTG GGCTCTGATCCAGCCTCAGCGTCACTGTCTTCCACGCCCCTCTTTGATTCTCGT TTATGTCAAAAGCCTTGTGAGGATGAGGCTGTGATTATCCCCATTTTACAGATG AGGAAACTGTGGCTCCAGGATGACACAACTGGCCAGAGGTCACATCAGAAGCA GAGCTGGGTCACTTGACTCCACCCAATATCCCTAAATGCAAACATCCCCTACAG ACCGAGGCTGGCACCTTAGAGCTGGAGTCCATGCCCGCTCTGACCAGGAGAA GCCAACCTGGTCCTCCAGAGCCAAGAGCTTCTGTCCCTTTCCCATCTCCTGAA GCCTCCCTGTCACCTTTAAAGTCCATTCCCACAAAGACATCATGGGATCACCAC AGAAAATCAAGCTCTGGGGCTAGGCTGACCCCAGCTAGATTTTTGGCTCTTTTA TACCCCAGCTGGGTGGACAAGCACCTTAAACCCGCTGAGCCTCAGCTTCCCGG GCTATAAAATGGGGGTGATGACACCTGCCTGTAGCATTCCAAGGAGGGTTAAA TGTGATGCTGCAGCCAAGGGTCCCCACAGCCAGGCTCTTTGCAGGTGCTGGGT TCAGAGTCCCAGAGCTGAGGCCGGGAGTAGGGGTTCAAGTGGGGTGCCCCAG GCAGGGTCCAGTGCCAGCCCTCTGTGGAGACAGCCATCCGGGGCCGAGGCAG CCGCCCACCGCAGGGCCTGCCTATCTGCAGCCAGCCCAGCCCTCACAAAGGA ACAATAACAGGAAACCATCCCAGGGGGAAGTGGGCCAGGGCCAGCTGGAAAA CCTGAAGGGGAGGCAGCCAGGCCTCCCTCGCCAGCGGGGTGTGGCTCCCCTC CAAAGACGGTCGGCTGACAGGCTCCACAGAGCTCCACTCACGCTCAGCCCTG GACGGACAGGCAGTCCAACGGAACAGAAACATCCCTCAGCCCACAGGCACGG TGAGTGGGGGCTCCCACACTCCCCTCCACCCCAAACCCGCCACCCTGCG	525

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
ubíquo	Promotor de núcleo EF1a	Gene EF1α	gggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggcaattgaacgggtgcctagag aaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcccgagggtgggggagaaccgt atataagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttgccgccagaacacag	137
ubíquo	EF1a	Gene EF1α	ggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggcaa ttgaaccggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcccga gggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttgccgccagaacaca ggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttcgtgccttgaattacttccacctgg ctccagtacgtgattcttgatcccgagctggagccaggggcgggccttgcgctttaggagccccttcgcctcgtgcttga gttgaggcctggcctgggcgctggggccgccgcgtgcgaatctggtggcaccttcgcgcctgtctcgctgctttcgata agtctctagccatttaaaatttttgatgacctgctgcgacgctttttttctggcaagatagtcttgtaaatgcgggccaggat ctgcacactggtatttcggtttttgggccgcggccggcgacggggcccgtgcgtcccagcgcacatgttcggcgagg cggggcctgcgagcgcggccaccgagaatcggacgggggtagtctcaagctggccggcctgctctggtgcctggc ctcgcgccgccgtgtatcgccccgccctgggcggcaaggctggcccggtcggcaccagttgcgtgagcggaaaga tggccgcttcccggccctgctccagggggctcaaaatggaggacgcggcgctcgggagagcgggcgggtgagtc acccacacaaaggaaaagggcctttccgtcctcagccgtcgcttcatgtgactccacggagtaccgggcgccgtcc aggcacctcgattagttctggagcttttggagtacgtcgtctttaggttggggggagggggttttatgcgatggagtttcccc acactgagtgggtggagactgaagttaggccagcttggcacttgatgtaattctccttggaatttggccttttgagtttgg atcttggttcattctcaagcctcagacagtggttcaaagttttttttcttccatttcaggtgtcgtga	138
ubíquo	hPGK	Gene PGK	ggggttggggttgcgccttttccaaggcagccctgggtttgcgcagggacgcggctgctctgggcgtggttccgggaa acgcagcggcgccgaccctgggtctcgcacattcttcacgtccgttcgcagcgtcacccggatcttcgccgctaccctt gtgggccccccggcgacgcttcctgctccgcccctaagtcgggaaggttccttgcggttcgcggcgtgccggacgtg acaaacggaagccgcacgtctcactagtaccctcgcagacggacagcgccagggagcaatggcagcgcgccga ccgcgatgggctgtggccaatagcggctgctcagcggggcgcgccgagagcagcggccgggaaggggcggtgc gggaggcggggtgtggggcggtagtgtgggccctgttcctgcccgcgcggtgttccgcattctgcaagcctccggag cgcacgtcggcagtcggctccctcgttgaccgaatcaccgacctctctcccca	139
ubíquo	mCMV	Citomegalovírus	ggtaggcgtgtacggtgggaggcctatataagcagagct	140

Especificidade	Nome do Promotor	Fonte de elementos cis-regulatórioes	Sequência exemplificativa	SEQ ID NO
ubíquo	Ubc	Gene da ubiquitina C	gtctaacaaaaaagccaaaaacggccagaatttagcggacaatttactagtctaacactgaaaattacatattgacc caaatgattacatttcaaaaggtgcctaaaaaacttcacaaaacacactcgccaaccccgagcgcatagttcaaaa ccggagcttcagctacttaagaagataggtacataaaaccgaccaaagaaactgacgcctcacttatccctcccctc accagaggtccggcgcctgtcgattcaggagagcctaccctaggcccgaaccctgcgtcctgcgacggagaaaag cctaccgcacacctaccggcaggtggccccaccctgcattataagccaacagaacgggtgacgtcacgacacga cgagggcgcgcgctcccaaaggtacgggtgcactgcccaacggcaccgccataactgccgcccccgcaacaga cgacaaaccgagttctccagtcagtgacaaacttcacgtcagggtccccagatggtgccccagcccatctcacccg aataagagctttcccgcattagcgaaggcctcaagaccttgggttcttgccgcccaccatgccccccaccttgtttcaa cgacctcacagcccgcctcacaagcgtcttccattcaagactcgggaacagccgccattttgctgcgctcccccaa cccccagttcagggcaaccttgctcgcggacccagactacagcccttggcggtctctccacacgcttccgtcccaccg agcggcccggcggccacgaaagccccggccagcccagcagcccgctactcaccaagtgacgatcacagcgatc cacaaacaagaaccgcgacccaaatcccggctgcgacggaactagctgtgccacacccggcgcgtccttatataa tcatcggcgttcaccgccccacggagatccctccgcagaatcgccgagaagggactacttttcctcgcctgttccgctc tctggaaagaaaaccagtgccctagagtcacccaagtcccgtcctaaaatgtccttctgctgatactggggttctaagg ccgagtcttatgagcagcgggccgctgtcctgagcgtccgggcggaaggatcaggacgctcgctgcgcccttcgtct gacgtggcagcgctcgccgtgaggaggggggcgcccgcgggaggcgccaaaacccggcgcggaggcc	141
ubíquo	SFFV	Vírus formador de foco no baço	gtaacgccattttgcaaggcatggaaaaataccaaaccaagaatagagaagttcagatcaagggcgggtacatga aaatagctaacgttgggccaaacaggatatctgcggtgagcagtttcggccccggcccggggccaagaacagatg gtcaccgcagtttcggccccggcccgaggccaagaacagatggtccccagatatggcccaaccctcagcagtttctt aagacccatcagatgtttccaggctcccccaaggacctgaaatgaccctgcgccttatttgaattaaccaatcagcctg cttctcgcttctgttcgcgcgcttctgcttcccgagctctataaaagagctcacaacccctcactcggcgcgccagtcctc cgacagactgagtcgcccggg	142

[00411] Várias sequências de consenso dentro de módulos cis-regulatórioes específicos do fígado (por exemplo, promotores) foram descritas. Em algumas modalidades, um módulo cis-regulatório específico do fígado compreende um sítio de ligação para um ou mais dentre HNF1α, C/EBP, LEF1, FOX, IRF, LEF1/TCF, Tal1β/E47 e MyoD. Em algumas modalidades, um módulo cis-regulatório específico do fígado compreende uma sequência estabelecida na Figura 1 ou Tabela 1 de Chuah et al, “Liver-Specific Transcriptional Modules Identified by Genome-Wide In Silico Analysis Enable Efficient Gene Therapy in Mice and Non-Human Primates” Mol Ther. Setembro de 2014; 22(9): 1.605 a 1.613, que é incorporado no presente documento a título de referência em sua totalidade, incluindo as sequências da Figura 1 e da Tabela 1 no mesmo. Em algumas modalidades, um módulo cis-regulatório específico do fígado compreende uma sequência humana de HS-CRM1, HS-CRM2, HS-CRM3, HS-CRM4, HS-CRM5, HS-CRM6, HS-CRM7, HS-CRM8, HS-CRM9, HS-CRM10, HS-CRM11, HS-CRM12, HS-CRM13 ou HS-CRM14 conforme descrito em Chuah et al supra. Promotores específicos de células adicionais e elementos cis-regulatórios, por exemplo, promotores específicos do fígado ou elementos cis-regulatórios, podem ser identificados usando métodos descritos em Chuah et al., supra.

[00412] Um sítio de entrada de ribossoma interno (IRES) tipicamente promove a entrada de ribossoma interno direto no códon de iniciação, como ATG, de um cístron (uma região de codificação de proteína), levando assim à tradução independente de cap do gene. Consultar, por exemplo, Jackson et al, (1990) Trends Biochem Sci 15 (12): 477 a 483) e Jackson e Kaminski. (1995) RNA 1 (10):985 a 1.000. Nas modalidades particulares, um vetor inclui um ou mais genes exógenos que codificam um ou mais agentes exógenos. Nas modalidades particulares, para alcançar a tradução eficiente de cada um dentre a plura-

lidade de agentes proteicos exógenos, as sequências polinucleotídicas podem ser separadas por uma ou mais sequências IRES ou sequências polinucleotídicas que codificam polipeptídeos autocliváveis.

[00413] Os ácidos nucleicos retrovirais no presente documento também podem compreender uma ou mais sequências Kozak, por exemplo, uma sequência de nucleotídeos curta que facilita a ligação inicial do mRNA à pequena subunidade do ribossoma e aumenta a tradução. A sequência consenso Kozak é (GCC)RCCATGG, em que R é uma purina (A ou G) (Kozak, (1986) Cell. 44(2):283 a 292, e Kozak, (1987) Nucleic Acids Res. 15(20): 8.125 a 8.148).

PROMOTORES RESPONSIVOS A UM FATOR DE TRANSCRIÇÃO HETERÓLOGO E INDUTOR

[00414] Em algumas modalidades, um ácido nucleico retroviral compreende um elemento que permite a expressão condicional do agente exógeno, por exemplo, qualquer tipo de expressão condicional incluindo, porém sem limitação, expressão induzível; expressão reprimível; expressão específica do tipo de célula ou expressão específica do tecido. Em algumas modalidades, a fim de atingir a expressão condicional do agente exógeno, a expressão é controlada submetendo-se uma célula, tecido ou organismo a um tratamento ou afecção que faz com que o agente exógeno seja expresso ou que causa um aumento ou diminuição na expressão do agente exógeno.

[00415] Exemplos ilustrativos de promotores/sistemas indutíveis incluem, porém sem limitação, promotores indutíveis por esteroides, tais como promotores para genes que codificam receptores de glucocorticoides ou de estrogênio (induzíveis por tratamento com o hormônio correspondente), promotor de metalotionina (induzível por tratamento com vários metais pesados), promotor MX-1 (induzível por interferon), o sistema regulável por mifepristona “GeneSwitch” (Sirin et al., 2003, *Gene,* 323:67), a troca de gene induzível por cumato (WO

2002/088346), sistemas reguladores dependentes de tetraciclina etc.

[00416] A expressão do transgene pode ser ativada ou reprimida pela presença ou ausência de uma molécula indutora. Em alguns casos, a molécula indutora ativa ou reprime a expressão gênica de maneira gradativa e, em alguns casos, as moléculas indutoras ativam ou reprimem a expressão gênica de uma maneira ou tudo ou nada.

[00417] Um promotor/sistema induzível comumente usado é o sistema regulado por tetraciclina (Tet). O sistema Tet ser baseia na coexpressão de dois elementos na respectiva célula-alvo: (i) no elemento de resposta da tetraciclina que contém repetições das sequências do operador Tet (TetO) fundidas a um promotor mínimo e conectado a um gene de interesse (por exemplo, um gene que codifica o agente exógeno) e (ii) o transativador transcricional (tTA), uma proteína de fusão do repressor Tet (TetR) e o domínio de transativação da proteína VP16 derivada do vírus herpes simples. Embora na versão originalmente descrita a expressão do transgene fosse ativa na ausência de tetraciclina ou seu análogo potente doxiciclina (Do), denominado de sistema Tet-OFF, a modificação de quatro aminoácidos dentro da proteína transativadora resultou em um tTA reverso (rtTA), que só se liga a TetO na presença de Dox (sistema Tet-ON). Em algumas modalidades, no transativador, o domínio VP16 foi substituído por domínios de ativação mínimos, locais potenciais de doador de splice e aceitador de splice foram removidos, e a proteína foi otimizada por códon, resultando na variante de transativador rtTA2S-M2 aprimorada com maior sensibilidade a Dox e menor atividade de linha de base. Além disso, diferentes elementos promotores responsivos a Tet foram gerados, incluindo modificação no TetO com espaçamento de 36 nucleotídeos de operadores vizinhos para aprimorar a regulação. Modificações adicionais podem ser úteis para reduzir ainda mais a atividade basal e aumentar a faixa dinâmica de expressão. Como exemplo, a variante

pTet-T11 (abreviatura: TII) exibe uma alta faixa dinâmica e baixa atividade de fundo.

[00418] A expressão condicional também pode ser alcançada com o uso de uma recombinase de DNA específica de local. De acordo com certas modalidades, o ácido nucleico retroviral compreende pelo menos um (tipicamente dois) sítio (ou locais) para recombinação mediada por uma recombinase específica de local, por exemplo, uma proteína excisiva ou integrativa, enzima, cofator ou proteína associada que está envolvida em reações de recombinação envolvendo um ou mais locais de recombinação (por exemplo, dois, três, quatro, cinco, sete, dez, doze, quinze, vinte, trinta, cinquenta, etc.), que podem ser proteínas de tipo selvagem (consultar Landy, Current Opinion in Biotechnology 3: 699 a 707 (1993)), ou mutantes, derivados (por exemplo, proteínas de fusão contendo as sequências de proteína de recombinação ou fragmentos das mesmas), fragmentos e variantes dos mesmos. Exemplos ilustrativos de recombinases incluem, porém sem limitação: Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin, ΦC31, Cin, Tn3 resolvase, TndX, XerC, XerD, TnpX, Hjc, Gin, SpCCE1 e ParA.

RIBOSWITCHES PARA REGULAR A EXPRESSÃO DO AGENTE EXÓGENO

[00419] Algumas das composições e métodos fornecidos no presente documento incluem um ou mais riboswitches ou polinucleotídeos que incluem um ou mais riboswitches. Riboswitches são uma característica comum em bactérias para regular a expressão de genes e são um meio para alcançar o controle de RNA de funções biológicas. Os riboswitches podem estar presentes na região 5' não traduzida dos mRNAs e podem permitir o controle regulatório sobre a expressão do gene por meio da ligação de um ligante de pequena molécula que induz ou suprime uma atividade do riboswitch. Em algumas modalidades, o riboswitch controla um produto gênico envolvido na geração do

ligante de molécula pequena. Riboswitches normalmente atuam de forma cis, embora tenham sido identificados riboswitches que atuam de forma trans. Os riboswitches naturais consistem em dois domínios: um domínio de aptâmero que se liga ao ligante por meio de uma estrutura de RNA dobrado tridimensional e um domínio de troca de função que induz ou suprime uma atividade no riboswitch com base na ausência ou presença do ligante. Assim, existem duas conformações sensíveis ao ligante alcançadas pelo riboswitch, representando estados ligado e desligado (Garst et al., 2011). O domínio de troca de função pode afetar a expressão de um polinucleotídeo regulando-se: um sítio de entrada de ribossoma interno, acessibilidade de doador de splice pré-mRNA na construção do gene retroviral, tradução, terminação da transcrição, degradação da transcrição, expressão de miRNA ou expressão de shRNA (Dambach e Winkler 2009). Os domínios de aptâmero e de comutação de função podem ser usados como componentes modulares, permitindo que dispositivos de RNA sintético controlem a expressão gênica como aptâmeros nativos, aptâmeros nativos qu sofreram mutação/evoluídos ou aptâmeros totalmente sintéticos que são identificados na triagem de bibliotecas de RNA aleatórias (McKeague et al 2016).

[00420] A família riboswitch purina representa uma das maiores famílias, com mais de 500 sequências encontradas (Mandal et al 2003; US20080269258; e WO2006055351). Os riboswitches de purina compartilham uma estrutura semelhante que consiste em três elementos helicoidais/estruturas de haste conservados (PI, P2, P3) com elementos de alça/junção intervenientes (JI-2, L2, J2-3, L3, J3-1). Os domínios de aptâmero da família purina de riboswitches variam naturalmente em sua afinidade/regulação por vários compostos de purina, tais como adenina, guanina, adenosina, guanosina, desoxiadenosina, desoxiguanosina, etc. devido à variação de sequência (Kim et al. 2007).

[00421] Em algumas modalidades, um ácido nucleico retroviral descrito no presente documento compreende um polinucleotídeo que codifica o agente exógeno ligado operacionalmente a um promotor e a um riboswitch. O riboswitch inclui um ou mais dentre, por exemplo, todos dentre: a.) um domínio de aptâmero, por exemplo, um domínio de aptâmero com capacidade para se ligar a um fármaco antiviral análogo de nucleosídeo e tendo ligação reduzida a guanina ou 2'-desoxiguanosina em relação ao fármaco antiviral análogo de nucleosídeo; e b.) um domínio de comutação de função, por exemplo, um domínio de comutação de função com capacidade para regular a expressão do agente exógeno, em que a ligação do análogo de nucleosídeo pelo domínio de aptâmero induz ou suprime a atividade de regulação da expressão do domínio de comutação de função, regulando assim a expressão do agente exógeno. Em algumas modalidades, o agente exógeno pode ser um polipeptídeo, um miRNA ou um shRNA. Por exemplo, em uma modalidade, o riboswitch está ligado operacionalmente a um ácido nucleico que codifica um receptor de antígeno quimérico (CAR). Em exemplos ilustrativos não limitantes fornecidos no presente documento, o gene exógeno codifica um ou mais polipeptídeos de sinalização projetados. Por exemplo, o riboswitch e o polinucleotídeo alvo que codifica um ou mais polipeptídeos de sinalização projetados podem ser encontrados no genoma de uma célula de origem, em uma partícula retroviral recombinante incompetente para replicação, em uma célula T e/ou em uma célula NK.

[00422] Os domínios de aptâmero podem ser usados, por exemplo, como componentes modulares e combinados com qualquer um dos domínios de comutação de função para afetar a transcrição de RNA. Em qualquer uma das modalidades no presente documento divulgadas, o riboswitch pode afetar a transcrição de RNA regulando-se qualquer uma das seguintes atividades: sítio de entrada ribossomal interno

(IRES), acessibilidade do doador de splice pré-mRNA, tradução, terminação da transcrição, degradação da transcrição, expressão de miRNA ou expressão de shRNA. Em algumas modalidades, o domínio de comutação de função pode controlar a ligação de um anti-IRES a um IRES (consultar, por exemplo, Ogawa, RNA (2011), 17: 478 a 488, cuja divulgação é incorporada no presente documento a título de referência em sua totalidade). Em qualquer uma das modalidades divulgadas no presente documento, a presença ou ausência do ligante de molécula pequena pode fazer com que o riboswitch afete a transcrição de RNA. Em algumas modalidades, o riboswitch pode incluir uma ribozima. Riboswitches com ribozimas podem inibir ou aumentar a degradação da transcrição de polinucleotídeos-alvo na presença do ligante de molécula pequena. Em algumas modalidades, a ribozima pode ser uma classe pistola de ribozima, uma classe cabeça de martelo de ribozima, uma classe torcida de ribozima, uma classe machadinha de ribozima ou o HDV (vírus da hepatite delta).

ELEMENTO REGULATÓRIO ESPECÍFICO DE CÉLULA NÃO ALVO

[00423] Em algumas modalidades, o elemento regulatório específico de célula não alvo ou TCSRE negativo compreende uma sequência de reconhecimento de miRNA específico de tecido, sítio de reconhecimento de protease específico de tecido, sítio de ubiquitina ligase específico de tecido, sítio de repressão transcricional específico de tecido ou sítio de repressão epigenética específico de tecido.

[00424] Em algumas modalidades, uma célula não alvo compreende um miRNA endógeno. O ácido nucleico retroviral (por exemplo, o gene que codifica o agente exógeno) pode compreender uma sequência de reconhecimento para esse miRNA. Desse modo, caso o ácido nucleico retroviral entre na célula não alvo, o miRNA pode regular negativamente a expressão do agente exógeno. Isso ajuda a atingir especificidade adicional para a célula-alvo versus células não alvo.

[00425] Em algumas modalidades, o miRNA é um pequeno RNA não codificante de 20 a 22 nucleotídeos, tipicamente excisado de estruturas precursoras de RNA de foldback de aproximadamente 70 nucleotídeos conhecidas como pré-miRNAs. De modo geral, os miRNAs regulam negativamente seus alvos de uma dentre duas maneiras, dependendo do grau de complementaridade entre o miRNA e o alvo. Primeiramente, miRNAs que se ligam com complementaridade perfeita ou quase perfeita às sequências de mRNA que codificam proteínas induzem normalmente a via de interferência mediada por RNA (RNAi). Os miRNAs que exercem seus efeitos regulatórios ligando-se a locais complementares imperfeitos dentro das regiões 3' não traduzidas (UTRs) de seus alvos de mRNA normalmente reprimem a expressão do gene alvo de maneira pós-transcricional, aparentemente no nível da tradução, por meio de um complexo RISC que é semelhante, ou possivelmente idêntico, àquele que é usado para a via de RNAi. Consistente com o controle translacional, os miRNAs que usam esse mecanismo reduzem os níveis de proteína de seus genes-alvo, porém os níveis de mRNA desses genes são minimamente afetados. Os miRNAs (por exemplo, miRNAs de ocorrência natural ou miRNAs projetados artificialmente) podem ter como alvo específico qualquer sequência de mRNA. Por exemplo, em uma modalidade, a pessoa versada na técnica pode projetar construtos de RNA de grampo curto expressos como transcrições primárias de miRNA humano (por exemplo, miR-30 ou miR-21). Esse projeto adiciona um sítio de processamento de Drosha ao construto de grampo e demonstrou aumentar muito a eficiência de knockdown (Pusch et al., 2004). A haste de grampo consiste em 22-nt de dsRNA (por exemplo, antissenso tem complementaridade perfeita com o alvo desejado) e uma alça de 15 a 19-nt de um miR humano. A adição das sequências de flanqueamento miR30 e de alça de miR em um ou ambos os lados do grampo resulta em um aumento

maior que 10 vezes no processamento de Drosha e Dicer dos grampos expressos quando comparados a projetos de shRNA convencionais sem microRNA. O aumento do processamento de Drosha e Dicer é traduzido em maior produção de siRNA/miRNA e maior potência para grampos expressos.

[00426] Centenas de genes de miRNA distintos são expressos diferencialmente durante o desenvolvimento e entre os tipos de tecido. Vários estudos sugeriram funções regulatórias importantes para miRNAs em uma ampla gama de processos biológicos, incluindo tempo de desenvolvimento, diferenciação celular, proliferação, apoptose, oncogênese, secreção de insulina e biossíntese de colesterol. (Consultar Bartel 2004 Cell 116:281 a 297; Ambros 2004 Nature 431:350 a 355; Du et al. 2005 Development 132:4.645 a 4.652; Chen 2005 N. Engl. J. Med. 353:1.768 a 1.771; Krutzfeldt et al. 2005 Nature 438:685 a 689.) Análise molecular mostrou que miRNAs têm perfis distintos de expressão em tecidos diferentes. Métodos computacionais foram usados para analisar a expressão de aproximadamente 7.000 alvos de miRNA humanos previstos. Os dados sugerem que a expressão de miRNA contribui amplamente para a especificidade da expressão de mRNA em muitos tecidos humanos. (Consultar Sood et al. 2006 PNAS EUA 103(8):2.746 a 2.751.)

[00427] Desse modo, uma abordagem com base em miRNA pode ser usada para restringir a expressão do agente exógeno a uma população de células-alvo, silenciando-se a expressão do agente exógeno em tipos de células não alvo com o uso de espécies de microRNA endógenos. O microRNA induz o silenciamento do gene pós-transcricional específico da sequência em muitos organismos, seja pela inibição da tradução do RNA mensageiro (mRNA) ou pela degradação do mRNA. Consultar, por exemplo, Brown et al. 2006 Nature Med. 12(5):585 a 591, e o documento n° WO2007/000668, dentre os quais

cada um é incorporado ao presente documento a título de referência em sua totalidade. Em algumas modalidades, o ácido nucleico retroviral compreende uma ou mais dentre (por exemplo, uma pluralidade de) sequências de reconhecimento de miRNA específicas de tecido. Em algumas modalidades, a sequência de reconhecimento de miRNA específica de tecido tem de aproximadamente 20 a 25, 21 a 24 ou 23 nucleotídeos de comprimento. Nas modalidades, a sequência de reconhecimento de miRNA específico de tecido tem complementaridade perfeita com um miRNA presente em uma célula não alvo. Em algumas modalidades, o agente exógeno não compreende GFP, por exemplo, não compreende uma proteína fluorescente, por exemplo, não compreende uma proteína repórter. Em algumas modalidades, as células fora do alvo não são células hematopoiéticas e/ou o miRNA não está presente nas células hematopoiéticas.

[00428] Em algumas modalidades, um método no presente documento compreende a expressão específica de tecido de um agente exógeno em uma célula-alvo que compreende colocar uma pluralidade de vetores retrovirais que compreendem um nucleotídeo que codifica o agente exógeno e pelo menos uma sequência alvo de microRNA específico de tecido (miRNA) em contato com uma pluralidade de células que compreendem células-alvo e células não alvo, em que o agente exógeno é preferencialmente expresso na célula-alvo, por exemplo, restrito à mesma.

[00429] Por exemplo, o ácido nucleico retroviral pode compreender pelo menos uma sequência de reconhecimento de miRNA operacionalmente ligada a uma sequência de nucleotídeos que têm um miRNA correspondente em uma célula não alvo, por exemplo, uma célula progenitora hematopoiética (HSPC), células-tronco hematopoiéticas (HSC), que impede ou reduz a expressão da sequência de nucleotídeos na célula não alvo, porém não em uma célula-alvo, por exemplo,

célula diferenciada. Em algumas modalidades, o ácido nucleico retroviral compreende pelo menos uma sequência de miRNA alvo para um miRNA que está presente em uma quantidade eficaz (por exemplo, a concentração do miRNA endógeno é suficiente para reduzir ou prevenir a expressão de um transgene) na célula não alvo, e compreende um transgene. Nas modalidades, o miRNA usado nesse sistema é fortemente expresso em células não alvo, como HSPC e HSC, porém não na progênie diferenciada de, por exemplo, linhagem mieloide e linfoide, prevenindo ou reduzindo a expressão de um transgene em populações de células-tronco sensíveis, enquanto mantém a expressão e eficácia terapêutica nas células-alvo.

[00430] Em algumas modalidades, o TSCRE ou NTSCRE negativo compreende um sítio de reconhecimento de miRNA, por exemplo, um sítio de reconhecimento de miRNA que é ligado por um miRNA endógeno a células hematopoiéticas. Os miRNAs são fornecidos na Tabela 4 a seguir. Em algumas modalidades, o ácido nucleico (por exemplo, ácido nucleico de fusossoma ou ácido nucleico retroviral) compreende uma sequência que é complementar a um miRNA da Tabela 4, ou tem pelo menos 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, ou 99% de complementaridade ao mesmo. Em algumas modalidades, o ácido nucleico (por exemplo, ácido nucleico de fusossoma ou ácido nucleico retroviral) compreende uma sequência que é perfeitamente complementar a uma sequência semente dentro de um miRNA endógeno, por exemplo, miRNA da Tabela 4. Nas modalidades, a sequência de semente tem pelo menos 6, 7, 8, 9 ou 10 nucleotídeos de comprimento.

[00431] Em algumas modalidades, o ácido nucleico (por exemplo, ácido nucleico de fusossoma ou ácido nucleico retroviral) compreende uma sequência que é complementar a um miRNA apresentado em qualquer uma das SEQ ID NOS: 143 a 160, ou uma sequência que

tem pelo menos 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de complementaridade ao mesmo. Em algumas modalidades, o ácido nucleico (por exemplo, fusossoma ácido nucleico ou ácido nucleico retroviral) compreende uma sequência que é perfeitamente complementar a uma sequência de sementes dentro de um miRNA endógeno, por exemplo, miRNA apresentado em qualquer uma das SEQ ID NOS: 143-160. Nas modalidades, a sequência de semente tem pelo menos 6, 7, 8, 9 ou 10 nucleotídeos de comprimento.

TABELA 4. SEQUÊNCIAS DE miRNA EXEMPLIFICATIVAS.

Tipo de célula silenciada	**nome de miRNA**	**miRNA maduro**	**sequência de miRNA**	**SEQ ID NO**
células hematopoiéticas	miR-142	hsa-miR-142-3p	uguaguguuuccuacuuuaugga	143
células hematopoiéticas	miR-142	hsa-miR-142-5p	cauaaaguagaaagcacuacu	144
células hematopoiéticas	mir-181a-2	hsa-miR-181a-5p	aacauucaacgcugucggugagu	145
células hematopoiéticas	mir-181a-2	hsa-miR-181a-2-3p	accacugaccguugacuguacc	146
células hematopoiéticas	mir-181b-1	hsa-miR-181b-5p	aacauucauugcugucggugggu	147
células hematopoiéticas	mir-181b-1	hsa-miR-181b-3p	cucacugaacaaugaaugcaa	148
células hematopoiéticas	mir-181c	hsa-miR-181c-5p	aacauucaaccugucggugagu	149
células hematopoiéticas	mir-181c	hsa-miR-181c-3p	aaccaucgaccguugaguggac	150
células hematopoiéticas	mir-181a-1	hsa-miR-181a	aacauucaacgcugucggugagu	151
células hematopoiéticas	mir-181a-1	hsa-miR-181a-3p	accaucgaccguugauuguacc	152
células hematopoiéticas	mir-181b-2	hsa-miR-181b-5p	aacauucauugcugucggugggu	153
células hematopoiéticas	mir-181b-2	hsa-miR-181b-2-3p	cucacugaucaaugaaugca	154

Tipo de célula silenciada	nome de miRNA	miRNA maduro	sequência de miRNA	SEQ ID NO
células hematopoiéticas	mir-181d	hsa-miR-181d-5p	aacauucauuguugucgguggguu	155
células hematopoiéticas	mir-181d	hsa-miR-181d-3p	ccaccgggggaugaaugucac	156
células hematopoiéticas	miR-223	hsa-miR-223-5p	cguguauuugacaagcugaguu	157
células hematopoiéticas	miR-223	hsa-miR-223-3p	ugucaguuugucaaauacccca	158
pDCs	miR-126	hsa-miR-126-5p	cauuauuacuuuugguacgcg	159
pDCs	miR-126	hsa-miR-126-3p	ucguaccgugaguaauaaugcg	160

[00432] Em algumas modalidades, o TSCRE ou NTSCRE negativo compreende um sítio de reconhecimento de miRNA para um miRNA descrito no presente documento. Os miRNAs exemplificativos incluem aqueles constatados em Griffiths-Jones et al. Nucleic Acids Res. 1° de janeiro de 2006, 34; Chen and Lodish, Semin Immunol. Abril de 2005; 17(2):155 a 165; Chen et al. Science. 2 de janeiro de 2004;303(5654): 83 a 86; Barad et al. Genome Res. dezembro de 2004; 14(12): 2.486 a 2.494; Krichevsky et al., RNA. Outubro de 2003;9(10):1.274 a 1.281; Kasashima et al. Biochem Biophys Res Commun. Setembro de 2004 17;322(2):403 a 410; Houbaviy et al., Dev Cell. Agosto de 2003;5(2): 351 a 358; Lagos-Quintana et al., Curr Biol. 30 de abril de 2002;12(9): 735 a 739; Calin et al., Proc Natl Acad Sci EUA. 2 de março de 2004;101(9):2.999 a 3.004; Sempere et al. Genome Biol. 2004; 5(3): R13; Metzler et al., Genes Chromosomes Cancer. Fevereiro de 2004; 39(2):167 a 169; Calin et al., Proc Natl Acad Sci EUA. 26 de novembro de 2002;99(24):15.524 a 15.529; Mansfield et al. Nat Genet. 36 de outubro de 2004;36(10):1.079 a 1.083; Michael et al. Mol Cancer Res. outubro de 2003;1(12):882 a 891; e em www.miRNA.org.

[00433] Em algumas modalidades, o TSCRE negativo ou NTSCRE compreende um sítio de reconhecimento de miRNA para um miRNA selecionado a partir de miR-1b, miR-189b, miR-93, miR-125b, miR-130, miR-32, miR-128, miR-22, miR124a, miR-296, miR-143, miR-15, miR-141, miR-143, miR-16, miR-127, miR99a, miR-183, miR-19b, miR-92, miR-9, miR-130b, miR-21, miR-30b, miR-16, miR-142-s, miR-99a, miR-212, miR-30c, miR-213, miR-20, miR-155, miR-152, miR-139, miR-30b, miR-7, miR-30c, miR-18, miR-137, miR-219, miR-1d, miR-178, miR-24, miR-122a, miR-215, miR-142-a, miR-223, miR-142, miR-124a, miR-190, miR-149, miR-193, miR-181, let-7a, miR-132, miR-27a, miR-9*, miR-200b, miR-266, miR-153, miR-135, miR-206, miR-24, miR-19a, miR-199, miR-26a, miR-194, miR-125a, miR-15a, miR-145,

miR-133, miR-96, miR-131, miR-124b, miR-151, miR-7b, miR-103 e miR-208.

[00434] Em algumas modalidades, o ácido nucleico (por exemplo, ácido nucleico de fusossoma ou ácido nucleico retroviral) compreende dois ou mais locais de reconhecimento de miRNA. Em algumas modalidades, cada um dos dois ou mais locais de reconhecimento miRNA são reconhecidos por um miRNA de qualquer um conforme descrito no presente documento, conforme estabelecido na Tabela 4. Em algumas modalidades, cada um dentre os dois ou mais locais de reconhecimento de miRNA são reconhecidos por um miRNA apresentado em qualquer uma das SEQ ID NOS: 143 a 160. Em algumas modalidades, os dois ou mais locais de reconhecimento de miRNA podem incluir 2, 3, 4, 5, 6, 7, 8, 9, 10 ou mais locais de reconhecimento de miRNA. Os dois ou mais locais de reconhecimento de miRNA podem ser posicionados em tandem no ácido nucleico para fornecer vários locais de ligação em tandem para um miRNA.

[00435] Em algumas modalidades, os dois ou mais locais de reconhecimento de miRNA podem incluir pelo menos um primeiro sítio de reconhecimento de miRNA, tal como 1, 2, 3, 4, 5, 6 ou mais primeiros locais de reconhecimento de miRNA, e pelo menos um segundo sítio de reconhecimento de miRNA, tal como 1, 2, 3, 4, 5, 6 ou mais locais de reconhecimento de segundo miRNA. Em algumas modalidades, o ácido nucleico contém dois ou mais primeiros locais de reconhecimento de miRNA e cada um dos primeiros locais de reconhecimento de miRNA estão presentes em tandem no ácido nucleico para fornecer vários locais de ligação em tandem para um primeiro miRNA e/ou o ácido nucleico contém mais dois segundos locais de reconhecimento de miRNA e cada um dos segundos locais de reconhecimento de miRNA estão presentes em tandem no ácido nucleico para fornecer vários locais de ligação em tandem para um segundo miRNA. Em al-

gumas modalidades, o primeiro sítio de reconhecimento de miRNA e o segundo sítio de reconhecimento de miRNA são reconhecidos pelo mesmo miRNA e, em algumas modalidades, o primeiro sítio de reconhecimento de miRNA e o segundo sítio de reconhecimento de miRNA são reconhecidos por diferentes miRNAs. Em algumas modalidades, o primeiro sítio de reconhecimento de miRNA e o segundo sítio de reconhecimento de miRNA são reconhecidos por miRNAs presentes na mesma célula não alvo, e em algumas modalidades, o primeiro sítio de reconhecimento de miRNA e o segundo sítio de reconhecimento de miRNA são reconhecidos por miRNAs presentes em diferentes células não alvo. Em algumas modalidades, tanto ou o primeiro sítio de reconhecimento de miRNA e segundo sítio de reconhecimento de miRNA, ou cada um dentre os mesmos, são reconhecidos por miRNAs dentre qualquer um dos descritos, tal como qualquer um estabelecido na Tabela 4. Em algumas modalidades, um ou ambos o primeiro sítio de reconhecimento de miRNA e o segundo sítio de reconhecimento de miRNA são reconhecidos por um miRNAs apresentado em qualquer uma das SEQ ID NOS: 143 a 160.

[00436] Em algumas modalidades, um ou mais dentre os locais de reconhecimento de miRNA no ácido nucleico de fuososssoma (por exemplo, ácido nucleico retroviral) são transcritos em *cis* com o agente exógeno. Em algumas modalidades, um ou mais dentre os locais de reconhecimento de miRNA no ácido nucleico do fusossoma (por exemplo, ácido nucleico retroviral) estão situados a jusante da sequência da cauda poli-A, por exemplo, entre a sequência da cauda poli-A e o WPRE. Em algumas modalidades, um ou mais dos locais de reconhecimento de miRNA no ácido nucleico do fusossoma (por exemplo, ácido nucleico retroviral) estão situados a jusante do WPRE.

MODULAÇÃO IMUNOLÓGICA

[00437] Em algumas modalidades, um vetor retroviral ou VLP des-

crito no presente documento compreende CD47 elevado. Consultar, por exemplo, a patente nº U.S. 9.050.269, que é incorporada no presente documento a título de referência em sua totalidade. Em algumas modalidades, um vetor retroviral ou VLP descrito no presente documento compreende a proteína Regulatória do Complemento elevada. Consultar, por exemplo, os documentos n° ES2627445T3 e n° US6790641, dentre os quais cada um é incorporado ao presente documento a título de referência em sua totalidade. Em algumas modalidades, um vetor retroviral ou VLP descrito no presente documento não tem uma proteína MHC, por exemplo, um MHC-1 classe 1 ou classe II, ou compreende níveis reduzidos da mesma. Consultar, por exemplo, o documento US20170165348, que é incorporado no presente documento a título de referência em sua totalidade.

[00438] Às vezes, os vetores retrovirais ou VLPs são reconhecidos pelo sistema imunológico do indivíduo. No caso de partículas de vetor viral envelopadas (por exemplo, partículas de vetor retroviral), as proteínas ligadas à membrana que são exibidas na superfície do envelope viral podem ser reconhecidas e a própria partícula viral pode ser neutralizada. Além disso, na infecção de uma célula-alvo, o envelope viral torna-se integrado à membrana celular e, como resultado, as proteínas do envelope viral podem ser exibidas ou permanecer em associação próxima com a superfície da célula. Portanto, o sistema imunológico também pode alvejar as células que as partículas do vetor viral infectaram. Ambos os efeitos podem causar uma redução na eficácia da entrega do agente exógeno por vetores virais.

[00439] Um envelope de partícula viral normalmente se origina em uma membrana da célula de origem. Portanto, proteínas de membrana que são expressas na membrana celular a partir das quais os botões de partículas virais podem ser incorporados ao envelope viral.

A PROTEÍNA IMUNOMODULADORA CD47

[00440] A internalização de material extracelular em células é comumente realizada por um processo denominado endocitose (Rabinovitch, 1995, Trends Cell Biol. 5(3):85 a 87; Silverstein, 1995, Trends Cell Biol. 5(3):141 a 142; Swanson et al., 1995, Trends Cell Biol. 5(3):89 a 93; Allen et al., 1996, J. Exp. Med. 184(2):627 a 637). A endocitose pode se enquadrar em duas categorias gerais: fagocitose, que envolve a absorção de partículas, e pinocitose, que envolve a absorção de líquido e solutos.

[00441] Foi demonstrado que os fagócitos profissionais se diferenciam dos não próprios e dos próprios, com base em estudos com camundongos knockout sem o receptor de membrana CD47 (Oldenborg et al., 2000, Science 288(5473):2.051 a 2.054). O CD47 é um membro ubíquo da superfamília Ig que interage com o receptor inibidor imunológico SIRPα (proteína regulatória de sinal) encontrado em macrófagos (Fujioka et al., 1996, Mol. Cell. Biol. 16(12):6.887 a 6.899; Veillette et al., 1998, J. Biol. Chem. 273(35):22.719 a 22.728; Jiang et al., 1999, J. Biol. Chem. 274(2):559 a 562). Embora as interações CD47-SIRPα pareçam desativar macrófagos autólogos em camundongos, reduções graves de CD47 (talvez 90%) são encontradas em células sanguíneas humanas de alguns genótipos Rh que mostram pouca ou nenhuma evidência de anemia (Mouro-Chanteloup et al., 2003, Blood 101 (1): 338 a 344) e também pouca ou nenhuma evidência de interações celulares intensificadas com monócitos fagocíticos (Arndt et al., 2004, Br. J. Haematol. 125(3): 412 a 414).

[00442] Em algumas modalidades, um vetor retroviral ou VLP (por exemplo, uma partícula viral com um raio inferior a aproximadamente 1 μm, inferior a de aproximadamente 400 nm ou inferior a de aproximadamente 150 nm) compreende pelo menos uma porção biologicamente ativa de CD47, por exemplo, em uma superfície exposta do vetor

retroviral ou VLP. Em algumas modalidades, o vetor retroviral (por exemplo, lentivírus) ou VLP inclui um revestimento lipídico. Nas modalidades, a quantidade de CD47 biologicamente ativo no vetor retroviral ou VLP está entre de aproximadamente 20 a 250, 20 a 50, 50 a 100, 100 a 150, 150 a 200 ou 200 a 250 moléculas/μm^2. Em algumas modalidades, o CD47 é CD47 humano.

[00443] Um método descrito no presente documento pode compreender a evasão da fagocitose de uma partícula por uma célula fagocítica. O método pode incluir a expressão de pelo menos um peptídeo que incluir pelo menos uma porção biologicamente ativa de CD47 em um vetor retroviral ou VLP de modo que, quando o vetor retroviral ou VLP que compreende o CD47 é exposto a uma célula fagocítica, a partícula viral evita a facocitose pela célula fagocítica ou mostra uma fagocitose diminuída em comparação com um vetor retroviral ou VLP de outra forma similar não modificado. Em algumas modalidades, a meia-vida do vetor retroviral ou VLP em um indivíduo é estendida em comparação com um vetor retroviral ou VLP de outra forma similar não modificado.

DELEÇÃO DE MHC

[00444] O principal complexo de histocompatibilidade de classe I (MHC-I) é uma proteína da membrana da célula hospedeira que pode ser incorporada aos envelopes virais e, por ser altamente polimórfica por natureza, é um alvo principal da resposta imunológica do corpo (McDevitt H. O. (2000) *Annu. Rev. Immunol.* 18: 1 a 17). As moléculas MHC-I expostas na membrana plasmática das células de origem podem ser incorporadas no envelope da partícula viral durante o processo de brotamento do vetor. Essas moléculas MHC-I derivadas das células de origem e incorporadas nas partículas virais podem, por sua vez, ser transferidas para a membrana plasmática das células-alvo. Alternativamente, as moléculas MHC-I podem permanecer em estreita

associação à membrana da célula-alvo como resultado da tendência das partículas virais para absorver e permanecer ligadas à membrana da célula-alvo.

[00445] A presença de moléculas MHC-I exógenas na membrana plasmática de células submetidas à transdução ou próximo à mesma pode desencadear uma resposta imunológica alorreativa em indivíduos. Isso pode levar à morte mediada por imunidade ou fagocitose de células submetidas à transdução, seja por transferência de genes *ex vivo* seguida pela administração das células submetidas à transdução ao indivíduo, seja por administração direta *in vivo* das partículas virais. Além disso, no caso de administração *in vivo* de partículas virais contendo MHC-I na corrente sanguínea, as partículas virais podem ser neutralizadas por anticorpos específicos de MHC-I pré-existentes antes de atingirem suas células-alvo.

[00446] Consequentemente, em algumas modalidades, uma célula de origem é modificada (por exemplo, geneticamente modificada) para diminuir a expressão de MHC-I na superfície da célula. Nas modalidades, a fonte compreende uma ruptura geneticamente modificada de um gene que codifica β2-microglobulina (β2M). Nas modalidades, a célula de origem compreende uma ruptura geneticamente modificada de um ou mais genes que codificam uma cadeia MHC-I α. A célula pode compreender interrupções geneticamente modificadas em todas as cópias do gene que codifica β2-microglobulina. A célula pode compreender interrupções geneticamente modificadas em todas as cópias dos genes que codificam uma cadeia MHC-I α. A célula pode compreender tanto interrupções por engenharia genética de genes que codificam β2-microglobulina quanto interrupções por engenharia genética de genes que codificam uma cadeia MHC-I α. Em algumas modalidades, o vetor retroviral ou VLP compreende um número reduzido de moléculas MHC-I expostas à superfície. O número de moléculas MHC-I expostas

à superfície pode ser diminuído de modo que a resposta imunológica ao MHC-I seja reduzida a um grau terapeuticamente relevante. Em algumas modalidades, a partícula de vetor viral envelopada é substancialmente desprovida de moléculas MHC-I expostas à superfície.

SUPEREXPRESSÃO DE HLA-G/E

[00447] Em algumas modalidades, um vetor retroviral ou VLP exibe em seu envelope uma proteína tolerogênica, por exemplo, um agonista ILT-2 ou ILT-4, por exemplo, HLA-E ou HLA-G ou qualquer outro agonista ILT-2 ou ILT-4. Em algumas modalidades, um vetor retroviral ou VLP aumentou a expressão de HLA-E, HLA-G, ILT-2 ou ILT-4 em comparação a um retrovírus de referência, por exemplo, um retrovírus não modificado de outra forma similar ao retrovírus.

[00448] Em algumas modalidades, uma composição de retrovírus diminuiu a Classe I de MHC em comparação a um retrovírus não modificado e aumentou HLA-G em comparação a um retrovírus não modificado.

[00449] Em algumas modalidades, o vetor retroviral ou VLP tem um aumento na expressão de HLA-G ou HLA-E, por exemplo, um aumento na expressão de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de HLA-G ou HLA-E, em comparação com um retrovírus de referência, por exemplo, um retrovírus não modificado semelhante ao retrovírus, em que a expressão de HLA-G ou HLA- E é testada *in vitro* usando citometria de fluxo, por exemplo, FACS.

[00450] Em algumas modalidades, o retrovírus com expressão aumentada de HLA-G demonstra imunogenicidade reduzida, por exemplo, conforme medido pela infiltração de células imunes reduzida, em um ensaio de formação de teratoma.

PROTEÍNAS REGULATÓRIAS DO COMPLEMENTO

[00451] A atividade do complemento é normalmente controlada por várias proteínas regulatórias do complemento (CRPs). Essas proteínas

evitam a inflamação espúria e os danos ao tecido do hospedeiro. Um grupo de proteínas, incluindo CD55/fator de aceleração de decaimento (DAF) e CD46/proteína cofator de membrana (MCP), inibe as enzimas de conversão C3/C5 da via clássica e alternativa. Outro conjunto de proteínas, incluindo CD59, regula a montagem do MAC. Os CRPs têm sido usados para prevenir a rejeição de tecidos xenotransplantados e também mostraram proteger os vírus e vetores virais da inativação do complemento.

[00452] Os fatores de controle do complemento residente na membrana incluem, por exemplo, fator de aceleração do decaimento (DAF) ou CD55, proteína-1 semelhante ao fator H (FH) (FHL-1), proteína de ligação a C4b (C4BP), receptor de complemento 1 (CD35), proteína cofator de membrana (MCP) ou CD46 e CD59 (protectina) (por exemplo, para prevenir a formação do complexo de ataque à membrana (MAC) e proteger as células da lise).

PROTEÍNA DE LIGAÇÃO À ALBUMINA

[00453] Em algumas modalidades, o lentivírus se liga à albumina. Em algumas modalidades, o lentivírus compreende em sua superfície uma proteína que se liga à albumina. Em algumas modalidades, o lentivírus compreende em sua superfície uma proteína de ligação à albumina. Em algumas modalidades, a proteína de ligação à albumina é a proteína de ligação à albumina estreptocócica. Em algumas modalidades, a proteína de ligação à albumina é o domínio de ligação à albumina estreptocócica.

EXPRESSÃO DE PROTEÍNAS NÃO FUSOGÊNICAS NO ENVELOPE LENTIVIRAL

[00454] Em algumas modalidades, o lentivírus é projetado para compreender uma ou mais proteínas em sua superfície. Em algumas modalidades, as proteínas afetam as interações imunológicas com um indivíduo. Em algumas modalidades, as proteínas afetam a farmacolo-

gia do lentivírus no indivíduo. Em algumas modalidades, a proteína é um receptor. Em algumas modalidades, a proteína é um agonista. Em algumas modalidades, a proteína é uma molécula de sinalização. Em algumas modalidades, a proteína na superfície lentiviral compreende um anticorpo anti-CD3 (por exemplo, OKT3) ou IL7.

[00455] Em algumas modalidades, uma proteína transmembranar mitogênica e/ou uma proteína transmembranar à base de citocina está presente na célula de origem, que pode ser incorporada ao retrovírus quando brota da membrana da célula de origem. A proteína transmembranar mitogênica e/ou uma proteína transmembranar à base de citocina pode ser expressa como uma molécula de superfície celular separada na célula de origem em vez de ser parte da glicoproteína do envelope viral.

[00456] Em algumas modalidades de qualquer um dos aspectos descritos no presente documento, o vetor retroviral, VLP ou composição farmacêutica é substancialmente não imunogênico. A imunogenicidade pode ser quantificada, por exemplo, conforme descrito no presente documento.

[00457] Em algumas modalidades, um vetor retroviral ou VLP é fundido com uma célula-alvo para produzir uma célula receptora. Em algumas modalidades, uma célula receptora que se fundiu a um ou mais vetores retrovirais ou VLPs é avaliada quanto à imunogenicidade. Nas modalidades, uma célula receptora é analisada quanto à presença de anticorpos na superfície da célula, por exemplo, por marcação com um anticorpo anti-IgM. Em outras modalidades, a imunogenicidade é avaliada por um ensaio de lise de células PBMC. Nas modalidades, uma célula receptora é incubada com células mononucleares de sangue periférico (PBMCs) e, em seguida, avaliada quanto à lise das células pelas PBMCs. Em outras modalidades, a imunogenicidade é avaliada por um ensaio de lise de células exterminadoras naturais (NK).

Nas modalidades, uma célula receptora é incubada com células NK e, em seguida, avaliada quanto à lise das células pelas células NK. Em outras modalidades, a imunogenicidade é avaliada por um ensaio de lise de células T CD8+. Nas modalidades, uma célula receptora é incubada com células T CD8+ e, em seguida, avaliada quanto à lise das células pelas células T CD8+.

[00458] Em algumas modalidades, o vetor retroviral ou VLP compreende níveis elevados de um agente imunossupressor (por exemplo, proteína imunossupressora) em comparação com um vetor retroviral ou VLP de referência, por exemplo, um produzido a partir de uma célula de origem não modificada de outra forma similar à célula de origem, ou uma célula HEK293. Em algumas modalidades, o nível elevado é de pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes. Em algumas modalidades, o vetor retroviral ou VLP compreende um agente imunossupressor que está ausente da célula de referência. Em algumas modalidades, o vetor retroviral ou VLP compreende níveis reduzidos de um agente imunoestimulador (por exemplo, proteína imunoestimuladora) em comparação a um vetor retroviral ou VLP de referência, por exemplo, um produzido a partir de uma célula de origem não modificada de outra forma similar à célula de origem, ou uma célula HEK293. Em algumas modalidades, o nível reduzido é de pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% ou 99% em comparação ao vetor retroviral ou VLP de referência. Em algumas modalidades, o agente imunoestimulador está substancialmente ausente do vetor retroviral ou VLP.

[00459] Em algumas modalidades, o vetor retroviral ou VLP ou a célula de origem da qual o vetor retroviral ou VLP é derivado, tem um, dois, três, quatro, cinco, seis, sete, oito, nove, dez, onze, doze ou mais das seguintes características:

[00460] a. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de classe I de MHC ou classe II de MHC em comparação a um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de origem de outra forma similar à célula de origem, ou uma célula HeLa, ou uma célula HEK293;

[00461] b. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de uma ou mais proteínas coestimuladoras incluindo, porém sem limitação: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3 ou B7-H4, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um retroviral não modificado vetor ou VLP de uma célula de outra forma similar à célula de origem, ou uma célula HEK, ou uma célula de referência descrita no presente documento;

[00462] c. expressão de proteínas de superfície que suprimem o envolvimento de macrófagos, por exemplo, CD47, por exemplo, expressão detectável por um método descrito no presente documento, por exemplo, mais de 1,5 vezes, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, ou mais expressão da proteína de superfície que suprime o envolvimento de macrófagos, por exemplo, CD47, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula Jurkat ou uma célula HEK293;

[00463] d. expressão de citocinas imunossupressoras solúveis, por exemplo, IL-10, por exemplo, expressão detectável por um método descrito no presente documento, por exemplo, mais de 1,5 vezes, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, ou mais expressão de citocinas imunossupressoras solúveis, por exemplo, IL-10, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um

vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, ou uma célula HEK293;

[00464] e. expressão de proteínas imunossupressoras solúveis, por exemplo, PD-L1, por exemplo, expressão detectável por um método descrito no presente documento, por exemplo, mais de 1,5 vezes, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, ou mais expressão de proteínas imunossupressoras solúveis, por exemplo, PD-L1, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, ou uma célula HEK293;

[00465] f. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de citocinas estimulantes da imunidade solúvel, por exemplo, IFN-gama ou TNF-a, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, ou uma célula HEK293 ou uma célula U-266;

[00466] g. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor do antígeno imunoestimulante endógeno, por exemplo, Zg16 ou Hormad1, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, ou uma célula HEK293 ou uma célula A549, ou uma célula SK-BR-3;

[00467] h. expressão de, por exemplo, expressão detectável por um método descrito no presente documento, HLA-E ou HLA-G, em comparação a um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat;

[00468] i. perfil de glicosilação de superfície, por exemplo, que contém ácido siálico, que atua para, por exemplo, suprimir a ativação de células NK;

[00469] j. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de TCRα/β, em comparação a um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat;

[00470] k. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de grupos sanguíneos ABO, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula HeLa;

[00471] l. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de Antígeno de Histocompatibilidade Menor (MHA), em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat; ou

[00472] m. tem menos de 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% ou menos de MHAs mitocondriais em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat, ou não ter MHAs mitocondriais detectáveis.

[00473] Nas modalidades, a proteína coestimuladora é 4-1BB, B7, SLAM, LAG3, HVEM ou LIGHT, e a célula de referência é HDLM-2. Em algumas modalidades, a proteína coestimuladora é BY-H3, e a célula de referência é HeLa. Em algumas modalidades, a proteína coestimuladora é ICOSL ou B7-H4, e a célula de referência é SK-BR-3. Em algumas modalidades, a proteína coestimuladora é ICOS ou OX40, e a célula de referência é MOLT-4. Em algumas modalidades, a proteína coestimuladora é CD28, e a célula de referência é U-266. Em algumas

modalidades, a proteína coestimuladora é CD30L ou CD27, e a célula de referência é Daudi.

[00474] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica não elicia substancialmente uma resposta imunogênica pelo sistema imunológico, por exemplo, sistema imunológico inato. Nas modalidades, uma resposta imunogênica pode ser quantificada, por exemplo, conforme descrito no presente documento. Em algumas modalidades, a resposta imunogênica do sistema imune inato compreende uma resposta de células imunes inatas incluindo, porém sem limitação, células NK, macrófagos, neutrófilos, basófilos, eosinófilos, células dendríticas, mastócitos ou células T gama/delta. Em algumas modalidades, uma resposta imunogênica do sistema imune inato compreende uma resposta do sistema complemento que inclui componentes sanguíneos solúveis e componentes ligados à membrana.

[00475] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica não elicia substancialmente uma resposta imunogênica pelo sistema imunológico, por exemplo, sistema imunológico adaptativo. Em algumas modalidades, uma resposta imunogênica pelo sistema imune adaptativo compreende uma resposta imunogênica por uma célula imune adaptativa incluindo, porém sem limitação, uma mudança, por exemplo, aumento, no número ou atividade de linfócitos T (por exemplo, células T CD4, T CD8 células e ou células T gama-delta) ou linfócitos B. Em algumas modalidades, uma resposta imunogênica pelo sistema imune adaptativo inclui níveis aumentados de componentes do sangue solúveis incluindo, porém sem limitação, uma mudança, por exemplo, aumento, no número ou atividade de citocinas ou anticorpos (por exemplo, IgG, IgM, IgE, IgA ou IgD).

[00476] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é modificado para ter imunogenicidade reduzida. Em algumas modalidades, o vetor retroviral, VLP ou composição far-

macêutica tem uma imunogenicidade menor que 5%, 10%, 20%, 30%, 40% ou 50% menor do que a imunogenicidade de um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat.

[00477] Em algumas modalidades de qualquer um dos aspectos descritos no presente documento, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de origem, por exemplo, uma célula de mamífero, que tem um genoma modificado, por exemplo, modificado com o uso de um método descrito no presente documento, para reduzir, por exemplo, diminuir, imunogenicidade. A imunogenicidade pode ser quantificada, por exemplo, conforme descrito no presente documento.

[00478] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de mamífero exaurida, por exemplo, com um knockout de um, dois, três, quatro, cinco, seis, sete ou mais dos seguintes:

[00479] a. MHC classe I, MHC classe II ou MHA;

[00480] b. uma ou mais proteínas coestimuladoras incluindo, porém sem limitação: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3 ou B7-H4;

[00481] c. citocinas imunoestimulantes solúveis, por exemplo, IFN-gama ou TNF-a;

[00482] d. antígeno imunoestimulador endógeno, por exemplo, Zg16 ou Hormad1;

[00483] e. receptores de células T (TCR);

[00484] f. os genes que codificam grupos sanguíneos ABO, por exemplo, gene *ABO*;

[00485] g. fatores de transcrição que conduzem a ativação imuno-

lógica, por exemplo, NFkB;

[00486] h. fatores de transcrição que controlam a expressão de MHC, por exemplo, transativador de classe II (CIITA), fator regulatório do Xbox 5 (RFX5), proteína associada a RFX (RFXAP) ou repetições de anquirina RFX (RFXANK; também conhecido como RFXB); ou

[00487] i. Proteínas TAP, por exemplo, TAP2, TAP1 ou TAPBP, que reduzem a expressão de MHC de classe I.

[00488] Em algumas modalidades, o vetor retroviral ou VLP é derivado de uma célula de origem com uma modificação genética que resulta no aumento da expressão de um agente imunossupressor, por exemplo, um, dois, três ou mais dos seguintes (por exemplo, em que antes da modificação genética a célula não expressou o fator):

[00489] a. proteínas de superfície que suprimem o envolvimento de macrófagos, por exemplo, CD47; por exemplo, expressão aumentada de CD47 em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat;

[00490] b. citocinas imunossupressoras solúveis, por exemplo, IL-10, por exemplo, expressão aumentada de IL-10 em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat;

[00491] c. proteínas imunossupressoras solúveis, por exemplo, PD-1, PD-L1, CTLA4 ou BTLA; por exemplo, expressão aumentada de proteínas imunossupressoras em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à fonte de célula, uma célula HEK293 ou uma célula Jurkat;

[00492] d. uma proteína tolerogênica, por exemplo, um agonista de

ILT-2 ou ILT-4, por exemplo, HLA-E ou HLA-G ou qualquer outro agonista de ILT-2 ou ILT-4 endógeno, por exemplo, expressão aumentada de HLA-E, HLA-G, ILT-2 ou ILT-4 em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat; ou

[00493] e. proteínas de superfície que suprimem a atividade do complemento, por exemplo, proteínas regulatórias do complenento, por exemplo, proteínas que se ligam ao fator acelerador do decaímento (DAF, CD55), por exemplo, proteína-1 semelhante ao fator H (FH) (FHL-1), por exemplo, proteína de ligação a C4b (C4BP), por exemplo, receptor 1 do complemento (CD35), por exemplo, proteína de cofator membranar (MCP, CD46), por exemplo, Profectina (CD59), por exemplo, proteínas que inibem enzimas convertase CD/C5 de trajetória de complemento clássico e alternativo, por exemplo, proteínas que regulam a montagem de MAC; por exemplo, expressão aumentada de uma proteína regulatória do complemento em comparação a um vetor retroviral de referência ou VLP, por exemplo, um vetor retroviral não modificado ou VLP de uma célula, de outro modo, semelhante para a célula de origem, uma célula HEK293 ou uma célula Jurkat.

[00494] Em algumas modalidades, o nível de expressão aumentado é de pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes maior em comparação com um vetor retroviral ou VLP de referência.

[00495] Em algumas modalidades, o vetor retroviral ou VLP é derivado de uma célula de origem modificada para ter expressão diminuída de um agente imunoestimulador, por exemplo, um, dois, três, quatro, cinco, seis, sete, oito ou mais dos seguintes:

[00496] a. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou

expressão menor de classe I de MHC ou classe II de MHC, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula HeLa;

[00497] b. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de uma ou mais proteínas coestimuladoras incluindo, porém sem limitação: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3 ou B7-H4, em comparação a um vetor retroviral ou VLP de referência, por exemplo, um retroviral não modificado vetor ou VLP de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula de referência descrita no presente documento;

[00498] c. menos de 50%, 40%, 30%, 20%, 15%, 10%, ou 5% ou expressão menor de citocinas imunoestimulantes solúveis, por exemplo, IFN-gama ou TNF-a, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula U-266;

[00499] d. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de antígeno imunoestimulante endógeno, por exemplo, Zg16 ou Hormad1, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293, ou uma célula A549, ou uma célula SK-BR-3;

[00500] e. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de receptores de células T (TCR) em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat;

[00501] f. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou ex-

pressão menor de grupos sanguíneos ABO, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula HeLa;

[00502] g. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de fatores de transcrição que conduzem a ativação imunológica, por exemplo, NFkB; em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat

[00503] h. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de fatores de transcrição que controlam a expressão de MHC, por exemplo, classe II de transativador (CIITA), fator regulatório do Xbox 5 (RFX5), proteína associada a RFX (RFXAP) ou repetições de anquirina RFX (RFXANK; também conhecido como RFXB) em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula Jurkat; ou

[00504] i. menos de 50%, 40%, 30%, 20%, 15%, 10% ou 5% ou expressão menor de proteínas TAP, por exemplo, TAP2, TAP1 ou TAPBP, que reduzem a expressão de classe I de MHC em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, uma célula HEK293 ou uma célula HeLa.

[00505] Em algumas modalidades, um vetor retroviral, VLP ou composição farmacêutica derivada de uma célula de mamífero, por exemplo, uma HEK293, modificada usando shRNA expressando lentivírus para diminuir a expressão de Classe I de MHC, tem menor expressão de Classe I de MHC em comparação com um vetor retroviral ou VLP

não modificado, por exemplo, um vetor retroviral ou VLP de uma célula (por exemplo, célula-tronco mesenquimal) que não foi modificada. Em algumas modalidades, um vetor retroviral ou VLP derivado de uma célula de mamífero, por exemplo, uma HEK293, modificada usando lentivírus que expressa HLA-G para aumentar a expressão de HLA-G, aumentou a expressão de HLA-G em comparação a um vetor retroviral ou VLP não modificado, por exemplo, de uma célula (por exemplo, uma HEK293) que não foi modificada.

[00506] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de origem, por exemplo, uma célula de mamífero, que não é substancialmente imunogênica, em que as células de origem estimulam, por exemplo, induzem a secreção de IFN-gama de células T, em um nível de 0 pg/ml a > 0 pg/ml, por exemplo, conforme testado *in vitro,* pelo ensaio ELISPOT de IFN-gama.

[00507] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de origem, por exemplo, uma célula de mamífero, em que a célula de mamífero é de uma cultura de células tratada com um agente imunossupressor, por exemplo, um glucocorticoide (por exemplo, dexametasona), citostático (por exemplo, metotrexato), anticorpo (por exemplo, Muromonab (OKT3)-CD3) ou modulador de imunofilina (por exemplo, Ciclosporina ou rapamicina).

[00508] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de origem, por exemplo, uma célula de mamífero, em que a célula de mamífero compreende um agente exógeno, por exemplo, um agente terapêutico.

[00509] Em algumas modalidades, o vetor retroviral, VLP ou composição farmacêutica é derivado de uma célula de origem, por exemplo, uma célula de mamífero, em que a célula de mamífero é uma cé-

lula recombinante.

[00510] Em algumas modalidades, o vetor retroviral, VLP ou produto farmacêutico é derivado de uma célula de mamífero geneticamente modificada para expressar imunoevasinas virais, por exemplo, hCMV US2 ou US11.

[00511] Em algumas modalidades, a superfície do vetor retroviral ou VLP, ou a superfície da célula de origem, é modificada de maneira covalente ou não covalente com um polímero, por exemplo, um polímero biocompatível que reduz a imunogenicidade e a depuração mediada imunologicamente, por exemplo, PEG.

[00512] Em algumas modalidades, a superfície do vetor retroviral ou VLP, ou a superfície da célula de origem é covalentemente ou não covalentemente modificada com um ácido siálico, por exemplo, um ácido siálico que compreende glicopolímeros, que contêm epítopos de glicano supressores de NK.

[00513] Em algumas modalidades, a superfície do vetor retroviral ou VLP ou a superfície da célula de origem é tratada enzimaticamente, por exemplo, com enzimas glicosidase, por exemplo, α-N-acetilgalactosaminidases, para remover grupos sanguíneos ABO

[00514] Em algumas modalidades, a superfície do vetor retroviral ou VLP, ou a superfície da célula de origem é tratada enzimaticamente, para dar origem a, por exemplo, induzir a expressão de grupos sanguíneos ABO que correspondem ao tipo de sangue do receptor.

PARÂMETROS PARA AVALIAR A IMUNOGENICIDADE

[00515] Em algumas modalidades, o vetor retroviral ou VLP é derivado de uma célula de origem, por exemplo, uma célula de mamífero que não é substancialmente imunogênica, ou modificada, por exemplo, modificada com o uso de um método descrito no presente documento, para ter uma redução na imunogenicidade. A imunogenicidade da célula de origem e do vetor retroviral ou VLP pode ser determinada por

qualquer um dos ensaios descritos no presente documento.

[00516] Em algumas modalidades, o vetor retroviral ou VLP tem um aumento, por exemplo, um aumento de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, ou mais, na sobrevivência do enxerto *in vivo* em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem.

[00517] Em algumas modalidades, o vetor retroviral ou VLP tem uma redução na imunogenicidade medida por uma redução na resposta humoral após uma ou mais implantações do vetor retroviral ou VLP em um modelo animal apropriado, por exemplo, um modelo animal descrito no presente documento, em comparação com um resposta humoral após uma ou mais implantações de um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em um modelo animal apropriado, por exemplo, um modelo animal descrito no presente documento. Em algumas modalidades, a redução na resposta humoral é medida em uma amostra de soro por um título de anticorpo anti-célula, por exemplo, título de anticorpo antirretroviral ou anti-VLP, por exemplo, por ELISA. Em algumas modalidades, a amostra de soro de animais administrados com o vetor retroviral ou VLP tem uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, ou mais de um título de anticorpo antirretroviral ou anti-VLP em comparação com a amostra de soro de animais administrados com um vetor retroviral ou VLP não modificado. Em algumas modalidades, a amostra de soro de animais administrados com o vetor retroviral ou VLP tem um título de anticorpo antirretroviral ou anti-VLP aumentado, por exemplo, aumentado em 1%, 2%, 5%, 10%, 20%, 30%, ou 40% da linha de base, por exemplo, em que a linha de base se refere à amostra de soro dos mesmos animais antes da administração do vetor re-

troviral ou VLP.

[00518] Em algumas modalidades, o vetor retroviral vector ou VLP tem uma redução em fagocitose de macrófagos, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, ou mais em fagocitose de macrófagos em comparação a um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a redução na fagocitose de macrófagos é determinada por ensaio do índice de fagocitose *in vitro*, por exemplo, conforme descrito no Exemplo 8. Em algumas modalidades, o vetor retroviral ou VLP tem um índice de fagocitose de 0, 1, 10, 100 ou mais, por exemplo, conforme medido por um ensaio do Exemplo 8, quando incubado com macrófagos em um ensaio *in vitro* de fagocitose de macrófagos.

[00519] Em algumas modalidades, a célula de origem ou célula receptora tem uma redução na lise celular mediada por citotoxicidade por PBMCs, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na lise celular em comparação com uma célula de referência, por exemplo, uma célula não modificada semelhante à célula de origem, ou uma célula receptora que recebeu um vetor retroviral ou VLP não modificado, ou células-tronco mesenquimais, por exemplo, com o uso de um ensaio do Exemplo 17. Nas modalidades, a célula de origem expressa HLA-G exógeno.

[00520] Em algumas modalidades, a célula de origem ou célula receptora tem uma redução na lise celular mediada por NK, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na lise celular mediada por NK em comparação com uma célula de referência, por exemplo, uma célula não modificada semelhante à célula de origem, ou uma célula receptora que recebeu um vetor retroviral ou VLP não modificado, em que a lise celular mediada por NK é avaliada *in vitro,* por um ensaio de liberação de

cromo ou ensaio de liberação de európio, por exemplo, com o uso de um ensaio do Exemplo 18.

[00521] Em algumas modalidades, a célula de origem ou célula receptora tem uma redução na lise de células mediada por células T CD8+, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na lise de células mediadas por células T CD8 em comparação com uma célula de referência, por exemplo, uma célula não modificada semelhante à célula de origem, ou uma célula receptora que recebeu um vetor retroviral ou VLP não modificado em que a lise de células mediada por células T CD8 é avaliada *in vitro,* por um ensaio do Exemplo 19.

[00522] Em algumas modalidades, a célula de origem ou célula receptora tem uma redução na proliferação e/ou ativação de células T CD4+, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais em comparação com uma célula de referência, por exemplo, uma célula não modificada de outra forma similar à célula de origem, ou uma célula receptora que recebeu um vetor retroviral ou VLP não modificado, em que a proliferação de células T CD4 é testada *in vitro* (por exemplo, ensaio de cocultura de célula de origem de mamífero modificada ou não modificada e células T CD4+ com Dynabeads CD3/CD28).

[00523] Em algumas modalidades, o vetor retroviral ou VLP causa uma redução na secreção de IFN-gama de células T, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na secreção de IFN-gama de células T em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a secreção de IFN-gama de célula T é avaliada *in vitro*, por exemplo, por ELISPOT de IFN-gama.

[00524] Em algumas modalidades, o vetor retroviral ou VLP causa

uma redução na secreção de citocinas imunogênicas, por exemplo, uma redução de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na secreção de citocinas imunogênicas em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a secreção de citocinas imunogênicas é testada *in vitro* usando ELISA ou ELISPOT.

[00525] Em algumas modalidades, o vetor retroviral ou VLP resulta no aumento da secreção de uma citocina imunossupressora, por exemplo, um aumento de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na secreção de uma citocina imunossupressora em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a secreção da citocina imunossupressora é avaliada *in vitro* usando ELISA ou ELISPOT.

[00526] Em algumas modalidades, o vetor retroviral ou VLP tem um aumento na expressão de HLA-G ou HLA-E, por exemplo, um aumento na expressão de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de HLA-G ou HLA-E, em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula semelhante à célula de origem, em que a expressão de HLA-G ou HLA-E é testada *in vitro* usando citometria de fluxo, por exemplo, FACS. Em algumas modalidades, o vetor retroviral ou VLP é derivado de uma célula de origem que é modificada para ter uma expressão aumentada de HLA-G ou HLA-E, por exemplo, em comparação com uma célula não modificada, por exemplo, uma expressão aumentada de 1%, 5 %, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de HLA-G ou HLA-E, em que a expressão de HLA-G ou HLA-E é testada *in vitro* usando citometria de fluxo, por exemplo, FACS. Em algumas modalidades, o vetor retroviral ou VLP

derivado de uma célula modificada com expressão aumentada de HLA-G demonstra imunogenicidade reduzida.

[00527] Em algumas modalidades, o vetor retroviral ou VLP tem ou causa um aumento na expressão de ligantes inibidores de células T (por exemplo, CTLA4, PD1, PD-L1), por exemplo, um aumento na expressão de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de ligantes inibidores de células T em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a expressão de ligantes inibidores de células T é testada *in vitro* usando citometria de fluxo, por exemplo, FACS.

[00528] Em algumas modalidades, o vetor retroviral ou VLP tem uma diminuição na expressão de ligantes coestimuladores, por exemplo, uma diminuição de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais na expressão de ligantes coestimuladores em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula de outra forma similar à célula de origem, em que a expressão de ligantes é testada *in vitro* usando citometria de fluxo, por exemplo, FACS.

[00529] Em algumas modalidades, o vetor retroviral ou VLP tem uma diminuição na expressão de classe I de MHC ou classe II de MHC, por exemplo, uma diminuição na expressão de 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais de Classe I de MHC ou Classe II de MHC em comparação com um vetor retroviral ou VLP de referência, por exemplo, um vetor retroviral ou VLP não modificado de uma célula semelhante à célula de origem ou uma célula HeLa, em que a expressão de MHC de Classe I ou II é testada *in vitro* usando citometria de fluxo, por exemplo, FACS.

[00530] Em algumas modalidades, o vetor retroviral ou VLP é derivado de uma fonte de células, por exemplo, uma fonte de células de

mamíferos, que é substancialmente não imunogênica. Em algumas modalidades, a imunogenicidade pode ser quantificada, por exemplo, conforme descrito no presente documento. Em algumas modalidades, a fonte de células de mamífero compreende qualquer um, todos ou uma combinação dos seguintes recursos:

[00531] a. em que a célula de origem é obtida a partir de uma fonte de células autólogas; por exemplo, uma célula obtida de um receptor que receberá, por exemplo, ao qual será administrado, o vetor retroviral ou VLP;

[00532] b. em que a célula de origem é obtida de uma fonte de célula alogênica que é de gênero compatível, por exemplo, semelhante a um receptor, por exemplo, um receptor descrito no presente documento que receberá, por exemplo, ao qual será administrado; o vetor retroviral ou VLP;

[00533] c. em que a célula de origem é obtida de uma fonte de célula alogênica que é compatível em HLA com o HLA do receptor, por exemplo, em um ou mais alelos;

[00534] d. em que a célula de origem é obtida de uma fonte celular alogênica que é um homozigoto HLA;

[00535] e. em que a célula de origem é obtida de uma fonte de células alogênicas que não possui (ou tem níveis reduzidos em comparação com uma célula de referência) classe I e II de MHC; ou

[00536] f. em que a célula de origem é obtida a partir de uma fonte celular que é conhecida por ser substancialmente não imunogênica, incluindo, porém sem limitação, uma célula-tronco, uma célula-tronco mesenquimal, uma célula-tronco pluripotente induzida, uma célula-tronco embrionária, uma célula de Sertoli ou uma célula epitelial pigmentar da retina.

[00537] Em algumas modalidades, o indivíduo ao qual será administrado o vetor retroviral ou VLP tem, ou é conhecido por ter, ou é tes-

tado para, um anticorpo pré-existente (por exemplo, IgG ou IgM) reativo com um vetor retroviral ou VLP. Em algumas modalidades, o indivíduo ao qual será administrado o vetor retroviral ou VLP não tem níveis detectáveis de um anticorpo pré-existente reativo com o vetor retroviral ou VLP. Os testes para o anticorpo são descritos, por exemplo, no Exemplo 13.

[00538] Em algumas modalidades, um indivíduo que recebeu o vetor retroviral ou VLP tem, ou é conhecido por ter, ou é testado para, um anticorpo (por exemplo, IgG ou IgM) reativo com um vetor retroviral ou VLP. Em algumas modalidades, o indivíduo que recebeu o vetor retroviral ou VLP (por exemplo, pelo menos uma, duas, três vezes, quatro vezes, cinco vezes ou mais) não tem níveis detectáveis de anticorpo reativo com o vetor retroviral ou VLP. Nas modalidades, os níveis de anticorpo não aumentam mais que 1%, 2%, 5%, 10%, 20% ou 50% entre dois instantes do tempo, em que o primeiro instante do tempo ocorre antes da primeira administração do vetor retroviral ou VLP e o segundo instante do tempo ocorre após uma ou mais administrações do vetor retroviral ou VLP. Os testes para o anticorpo são descritos, por exemplo, no Exemplo 14.

AGENTES EXÓGENOS

[00539] Em algumas modalidades, um vetor retroviral, VLP ou composição farmacêutica descrito no presente documento codifica um agente exógeno.

AGENTES EXÓGENOS DE PROTEÍNA

[00540] Em algumas modalidades, o agente exógeno compreende uma proteína citosólica, por exemplo, uma proteína que é produzida na célula receptora e localizada no citoplasma da célula receptora. Em algumas modalidades, o agente exógeno compreende uma proteína secretada, por exemplo, uma proteína que é produzida e secretada pela célula receptora. Em algumas modalidades, o agente exógeno

compreende uma proteína nuclear, por exemplo, uma proteína que é produzida na célula receptora e é importada para o núcleo da célula receptora. Em algumas modalidades, o agente exógeno compreende uma proteína organelar (por exemplo, uma proteína mitocondrial), por exemplo, uma proteína que é produzida na célula receptora e é importada para uma organela (por exemplo, uma mitocondrial) da célula receptora. Em algumas modalidades, a proteína é uma proteína de tipo selvagem ou uma proteína mutante. Em algumas modalidades, a proteína é uma proteína de fusão ou quimérica.

[00541] Em algumas modalidades, o agente exógeno é codificado por um gene entre OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAH, PAL, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LNBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTD, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL,GRHPR, HOGA1, LDLR, ACAD8, ACADSB, ACAT1, ACSF3, ASPA, AUH, DNAJC19, ETHE1, FBP1, FTCD, GSS, HIBCH, IDH2, L2HGDH, MLYCD, OPA3, OPLAH, OXCT1, POLG, PPM1K, SERAC1, SLC25A1, SUCLA2, SUCLG1, TAZ, AGK, CLPB, TMEM70, ALDH18A1, OAT, CA5A, GLUD1, GLUL, UMPS, SLC22A5, CPT1A, HADHA, HADH, SLC52A1, SLC52A2, SLC52A3, HADHB, GYS2, PYGL, SLC2A2, ALG1, ALG2, ALG3, ALG6, ALG8, ALG9, ALG11, ALG12, ALG13, ATP6V0A2, B3GLCT, CHST14, COG1, COG2, COG4, COG5, COG6, COG7, COG8, DOLK, DHDDS, DPAGT1, DPM1, DPM2, DPM3, G6PC3, GFPT1, GMPPA, GMPPB, MAGT1, MAN1B1,

MGAT2, MOGS, MPDU1, MPI, NGLY1, PGM1, PGM3, RFT1, SEC23B, SLC35A1, SLC35A2, SLC35C1, SSR4, SRD5A3, TMEM165, TRIP11, TUSC3, ALG14, B4GALT1, DDOST, NUS1, RPN2, SEC23A, SLC35A3, ST3GAL3, STT3A, STT3B, AGA, ARSA, ARSB, ASAH1, ATP13A2, CLN3, CLN5, CLN6, CLN8, CTNS, CTSA, CTSD, CTSF, CTSK, DNAJC5, FUCA1, GAA, GALC, GALNS, GLA, GLB1, GM2A, GNPTAB, GNPTG, GNS, GRN, GUSB, HEXA, HEXB, HGSNAT, HYAL1, IDS, IDUA, KCTD7, LAMP2, MAN2B1, MANBA, MCOLN1, MFSD8, NAGA, NAGLU, NEU1, NPC1, NPC2, SGSH, PPT1, PSAP, SLC17A5, SMPD1, SUMF1, TPP1, AHCY, GNMT, MAT1A, GCH1, PCBD1, PTS, QDPR, SPR, DNAJC12, ALDH4A1, PRODH, HPD, GBA, HGD, AMN, CD320, CUBN, GIF, TCN1, TCN2, PREPL, PHGDH, PSAT1, PSPH, AMT, GCSH, GLDC, LIAS, NFU1, SLC6A9, SLC2A1, ATP7A, AP1S1, CP, SLC33A1, PEX7, PHYH, AGPS, GNPAT, ABCD1, ACOX1, PEX1, PEX2, PEX3, PEX5, PEX6, PEX10, PEX12, PEX13, PEX14, PEX16, PEX19, PEX26, AMACR, ADA, ADSL, AMPD1, GPHN, MOCOS, MOCS1, PNP, XDH, SUOX, OGDH, SLC25A19, DHTKD1, SLC13A5, FH, DLAT, MPC1, PDHA1, PDHB, PDHX, PDP1, ABCC2, SLCO1B1, SLCO1B3, HFE2, ADAMTS13, PYGM, COL1A2, TNFRSF11B, TSC1, TSC2, DHCR7, PGK1, VLDLR, KYNU, F5, C3, COL4A1, CFH, SLC12A2, GK, SFTPC, CRTAP, P3H1, COL7A1, PKLR, TALDO1, TF, EPCAM, VHL, GC, SERPINA1, ABCC6, F8, F9, ApoB, PCSK9, LDLRAP1,ABCG5, ABCG8, LCAT, SPINK5 ou GNE.

[00542] Em algumas modalidades, o agente exógeno é codificado por um gene entre OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAL, PAH, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LMBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTD, HLCS, PC, SLC7A7, CPT2,

ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1 ou LDLR. Em algumas modalidades, o agente exógeno é a enzima fenilalanina amônia-liase (PAL).

[00543] Em algumas modalidades, o agente exógeno compreende uma proteína da Tabela 5 a seguir. Em algumas modalidades, o agente exógeno compreende a sequência humana de tipo selvagem de qualquer uma das proteínas da Tabela 5, um fragmento funcional da mesma (por exemplo, um fragmento enzimaticamente ativo da mesma) ou uma variante funcional da mesma. Em algumas modalidades, o agente exógeno compreende uma sequência de aminoácidos com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com relação a uma sequência de aminoácidos da Tabela 5, por exemplo, uma sequência de número de acesso de proteína Uniprot da coluna 4 da Tabela 5 ou uma sequência de aminoácidos da coluna 5 da Tabela 5. Em algumas modalidades, o gene de carga útil que codifica um agente exógeno codifica uma sequência de aminoácidos com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com relação a uma sequência de aminoácidos da Tabela 5. Em algumas modalidades, o gene de carga útil que codifica um agente exógeno tem uma sequência de ácidos nucleicos com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com relação a uma sequência de ácidos nucleicos da Tabela 5, por exemplo, um Número de Acesso ao Gene Ensembl da coluna 3 da Tabela 5.

[00544] Em algumas modalidades, o agente exógeno compreende uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 161 a 518. Em algumas modalidades, o agente exógeno

compreende a sequência humana de tipo selvagem apresentada em qualquer uma das SEQ ID NOS: 161 a 518, um fragmento funcional do mesmo (por exemplo, um fragmento enzimaticamente ativo do mesmo) ou uma variante funcional do mesmo. Em algumas modalidades, o agente exógeno compreende uma sequência de aminoácidos com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com relação a uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 161 a 518. Em algumas modalidades, o gene de carga útil que codifica um agente exógeno codifica uma sequência de aminoácidos com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com qualquer uma das SEQ ID NOS: 161 a 518. Em algumas modalidades, o gene de carga útil que codifica um agente exógeno codifica a sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 161 a 518.

TABELA 5. A primeira coluna lista os agentes exógenos que podem ser administrados para tratar as indicações na sexta coluna, de acordo com os métodos e usos descritos no presente documento. Cada número de acesso uniprot da tabela 5 é incorporado no presente documento a título de referência em sua totalidade.

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
OTC	5009	0036473	P00480	MLFNLRILLNNAAFRNGHNFMVRNFRCGQPLQNKVQLKGRDLLTLKNFTG EEIKYMLWLSADLKFRIKQKGEYLPLLQGKSLGMIFEKRSTRTRLSTETGFA LLGGHPCFLTTQDIHLGVNESLTDTARVLSSMADAVLARVYKQSDLDTLAK EASIPIINGLSDLYHPIQILADYLTLQEHYSSLKGLTLSWIGDGNNILHSIMMS AAKFGMHLQAATPKGYEPDASVTKLAEQYAKENGTKLLLTNDPLEAAHGG NVLITDTWISMGQEEEKKKRLQAFQGYQVTMKTAKVAASDWTFLHCLPRK PEEVDDEVFYSPRSLVFPEAENRKWTIMAVMVSLLTDYSPQLQKPKF [SEQ ID NO:161]	deficiência de ornitina transcar-bamilase (OTC)	Desordem do ciclo da ureia
CPS1	1373	0021826	P31327, Q6PEK7, B7ZAW0, A0A024R454	MTRILTAFKVVRTLKTGFGFTNVTAHQKWKFSRPGIRLLSVKAQTAHIVLED GTKMKGYSFGHPSSVAGEVVFNTGLGGYPEAITDPAYKGQILTMANPIIGN GGAPDTTALDELGLSKYLESNGIKVSGLLVLDYSKDYNHWLATKSLGQWL QEEKVPAIYGVDTRMLTKIIRDKGTMLGKIEFEGQPVDFVDPNKQNLIAEVS TKDVKVYGKGNPTKVVAVDCGIKNNVIRLLVKRGAEVHLVPWNHDFTKME YDGILIAGGPGNPALAEPLIQNVRKILESDRKEPLFGISTGNLITGLAAGAKT YKMSMANRGQNQPVLNITNKQAFITAQNHGYALDNTLPAGWKPLFVNVN DQTNEGIMHESKPFFAVQFHPEVTPGPIDTEYLFDSFFSLIKKGKATTITSV LPKPALVASRVEVSKVLILGSGGLSIGQAGEFDYSGSQAVKAMKEENVKTV LMNPNIASVQTNEVGLKQADTVYFLPITPQFVTEVIKAEQPDGLILGMGGQ TALNCGVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFSDKLNEINEKIAP SFAVESIEDALKAADTIGYPVMIRSAYALGGLGSGICPNRETLMDLSTKAFA MTNQILVEKSVTGWKEIEYEVVRDADDNCVTVCNMENVDAMGVHTGDSV VVAPAQTLSNAEFQMLRRTSINVVRHLGIVGECNIQFALHPTSMEYCIIEVN ARLSRSSALASKATGYPLAFIAAKIALGIPLPEIKNVVSGKTSACFEPSLDYM VTKIPRWDLDRFHGTSSRIGSSMKSVGEVMAIGRTFEESFQKALRMCHPSI EGFTPRLPMNKEWPSNLDLRKELSEPSSTRIYAIAKAIDDNMSLDEIEKLTYI	deficiência de carbamil-fosfato sintetase I (CPSI)	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DKWFLYKMRDILNMEKTLKGLNSESMTEETLKRAKEIGFSDKQISKCLGLT EAQTRELRLKKNIHPWVKQIDTLAAEYPSVTNYLYVTYNGQEHDVNFDDH GMMVLGCGPYHIGSSVEFDWCAVSSIRTLRQLGKKTVVVNCNPETVSTDF DECDKLYFEELSLERILDIYHQEACGGCIISVGGQIPNNLAVPLYKNGVKIM GTSPLQIDRAEDRSIFSAVLDELKVAQAPWKAVNTLNEALEFAKSVDYPCL LRPSYVLSGSAMNVVFSEDEMKKFLEEATRVSQEHPVVLTKFVEGAREVE MDAVGKDGRVISHAISEHVEDAGVHSGDATLMLPTQTISQGAIEKVKDATR KIAKAFAISGPFNVQFLVKGNDVLVIECNLRASRSFPFVSKTLGVDFIDVAT KVMIGENVDEKHLPTLDHPIIPADYVAIKAPMFSWPRLRDADPILRCEMAST GEVACFGEGIHTAFLKAMLSTGFKIPQKGILIGIQQSFRPRFLGVAEQLHNE GFKLFATEATSDWLNANNVPATPVAWPSQEGQNPSLSSIRKLIRDGSIDLV INLPNNNTKFVHDNYVIRRTAVDSGIPLLTNFQVTKLFAEAVQKSRKVDSKS LFHYRQYSAGKAA [SEQ ID NO:162]		
NAGS	162417	0161653	Q8N159, Q2NKP2	MATALMAVVLRAAAVAPRLRGRGGTGGARRLSCGARRRAARGTSPGRR LSTAWSQPQPPPEEYAGADDVSQSPVAEEPSWVPSPRPPVPHESPEPPS GRSLVQRDIQAFLNQCGASPGEARHWLTQFQTCHHSADKPFAVIEVDEEV LKCQQGVSSLAFALAFLQRMDMKPLVVLGLPAPTAPSGCLSFWEAKAQLA KSCKVLVDALRHNAAAAVPFFGGGSVLRAAEPAPHASYGGIVSVETDLLQ WCLESGSIPILCPIGETAARRSVLLDSLEVTASLAKALRPTKIIFLNNTGGLR DSSHKVLSNVNLPADLDLVCNAEWVSTKERQQMRLIVDVLSRLPHHSSAV ITAASTLLTELFSNKGSGTLFKNAERMLRVRSLDKLDQGRLVDLVNASFGK KLRDDYLASLRPRLHSIYVSEGYNAAAILTMEPVLGGTPYLDKFVVSSSRQ GQGSGQMLWECLRRDLQTLFWRSRVTNPINPWYFKHSDGSFSNKQWIFF WFGLADIRDSYELVNHAKGLPDSFHKPASDPGS [SEQ ID NO:163]	Deficiência de N-acetilglutamato sintase (NAGS)	Desordem do ciclo da ureia
BCK-DHA	593	0248098	A0A024R0K3, P12694, Q59EI3	MAVAIAAARVWRLNRGLSQAALLLLRQPGARGLARSHPPRQQQQFSSLD DKPQFPGASAEFIDKLEFIQPNVISGIPIYRVMDRQGQIINPSEDPHLPKEKV LKLYKSMTLLNTMDRILYESQRQGRISFYMTNYGEEGTHVGSAAALDNTD LVFGQYREAGVLMYRDYPLELFMAQCYGNISDLGKGRQMPVHYGCKERH FVTISSPLATQIPQAVGAAYAAKRANANRVVICYFGEGAASEGDAHAGFNF AATLECPIIFFCRNNGYAISTPTSEQYRGDGIAARGPGYGIMSIRVDGNDVF AVYNATKEARRRAVAENQPFLIEAMTYRIGHHSTSDDSSAYRSVDEVNYW	doença da urina em xarope de bordo (MSUD); Doença da Urina em Xarope de Bordo Clássica (CMSUD)	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DKQDHPISRLRHYLLSQGWWDEEQEKAWRKQSRRKVMEAFEQAERKPK PNPNLLFSDVYQEMPAQLRKQQESLARHLQTYGEHYPLDHFDK [SEQ ID NO:164]		
BCK-DHB	594	0083123	A0A140VKB3, P21953, B4E2N3, B7ZB80	MAVVAAAAGWLLRLRAAGAEGHWRRLPGAGLARGFLHPAATVEDAAQR RQVAHFTFQPDPEPREYGGQTQKMNLFQSVTSALDNSLAKDPTAVIFGEDV AFGGVFRCTVGLRDKYGKDRVFNTPLCEQGIVGFGIGIAVTGATAIAEIQFA DYIFPAFDQIVNEAAKYRYRSGDLFNCGSLTIRSPWGCVGHGALYHSQSP EAFFAHCPGIKVVIPRSPFQAKGLLLSCIEDKNPCIFFEPKILYRAAAEEVPIE PYNIPLSQAEVIQEGSDVTLVAWGTQVHVIREVASMAKEKLGVSCEVIDLR TIIPWDVDTICKSVIKTGRLLISHEAPLTGGFASEISSTVQEECFLNLEAPISR VCGYDTPFPHIFEPFYIPDKWKCYDALRKMINY [SEQ ID NO:165]	doença da urina em xarope de bordo (MSUD); Doença da Urina em Xarope de Bordo Clássica (CMSUD)	Acidemia orgânica
DBT	1629	0137992	P11182	MAAVRMLRTWSRNAGKLICVRYFQTCGNVHVLKPNYVCFFGYPSFKYSH PHHFLKTTAALRGQVVQFKLSDIGEGIREVTVKEWYVKEGDTVSQFDSICE VQSDKASVTITSRYDGVIKKLYYNLDDIAYVGKPLVDIETEALKDSEEDVVE TPAVSHDEHTHQEIKGRKTLATPAVRRLAMENNIKLSEVVGSGKDGRILKE DILNYLEKQTGAILPPSPKVEIMPPPPKPKDMTVPILVSKPPVFTGKDKTEPI KGFQKAMVKTMSAALKIPHFGYCDEIDLTELVKLREELKPIAFARGIKLSFM PFFLKAASLGLLQFPILNASVDENCQNITYKASHNIGIAMDTEQGLIVPNVK NVQICSIFDIATELNRLQKLGSVGQLSTTDLTGGTFTLSNIGSIGGTFAKPVI MPPEVAIGALGSIKAIPRFNQKGEVYKAQIMNVSWSADHRVIDGATMSRFS NLWKSYLENPAFMLLDLK [SEQ ID NO:166]	doença da urina em xarope de bordo (MSUD); Doença da Urina em Xarope de Bordo Clássica (CMSUD)	Acidemia orgânica
DLD	1738	0091140	A0A024R713, P09622, E9PEX6	MQSWSRVYCSLAKRGHFNRISHGLQGLSAVPLRTYADQPIDADVTVIGSG PGGYVAAIKAAQLGFKTVCIEKNETLGGTCLNVGCIPSKALLNNSHYYHMA HGKDFASRGIEMSEVRLNLDKMMEQKSTAVKALTGGIAHLFKQNKVVHVN GYGKITGKNQVTATKADGGTQVIDTKNILIATGSEVTPFPGITIDEDTIVSST GALSLKKVPEKMVVIGAGVIGVELGSVWQRLGADVTAVEFLGHVGGVGID MEISKNFQRILQKQGFKFKLNTKVTGATKKSDGKIDVSIEAASGGKAEVITC DVLLVCIGRRPFTKNLGLEELGIELDPRGRIPVNTRFQTKIPNIYAIGDVVAG PMLAHKAEDEGIICVEGMAGGAVHIDYNCVPSVIYTHPEVAWVGKSEEQL KEEGIEYKVGKFPFAANSRAKTNADTDGMVKILGQKSTDRVLGAHILGPGA	doença da urina em xarope de bordo (MSUD) Deficiência de di-idrolipoamida desidrogenase	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GEMVNEAALALEYGASCEDIARVCHAHPTLSEAFREANLAASFGKSINF [SEQ ID NO:167]		
MUT	4594	0146085	A0A024RD82, B2R6K1, P22033	MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQQQPLHPEWAALAKKQ LKGKNPEDLIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYT FRPWTIRQYAGFSTVEESNKFYKDNIKAGQQGLSVAFDLATHRGYDSDNP RVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVSMTMNGAVIPVLANFIVT GEEQGVPKEKLTGTIQNDILKEFMVRNTYIFPPEPSMKIIADIFEYTAKHMPK FNSISISGYHMQEAGADAILELAYTLADGLEYSRTGLQAGLTIDEFAPRLSF FWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQTSGW SLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIARN TQIIIQEESGIPKVADPWGGSYMMECLTNDVYDAALKLINEIEEMGGMAKA VAEGIPKLRIEECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAIDNTSV RNRQIEKLKKIKSSRDQALAERCLAALTECAASGDGNILALAVDASRARCT VGEITDALKKVFGEHKANDRMVSGAYRQEFGESKEITSAIKRVHKFMERE GRRPRLLVAKMGQDGHDRGAKVIATGFADLGFDVDIGPLFQTPREVAQQ AVDADVHAVGISTLAAGHKTLVPELIKELNSLGRPDILVMCGGVIPPQDYEF LFEVGVSNVFGPGTRIPKAAVQVLDDIEKCLEKKQQSV [SEQ ID NO:168]	deficiência de cobalamina A (acidemia metilmalônica)	Acidemia orgânica
MMAA	166785	0151611	Q8IVH4	MPMLLPHPHQHFLKGLLRAPFRCYHFIFHSSTHLGSGIPCAQPFNSLGLHC TKWMLLSDGLKRKLCVQTTLKDHTEGLSDKEQRFVDKLYTGLIQGQRACL AEAITLVESTHSRKKELAQVLLQKVLLYHREQEQSNKGKPLAFRVGLSGPP GAGKSTFIEYFGKMLTERGHKLSVLAVDPSSCTSGGSLLGDKTRMTELSR DMNAYIRPSPTRGTLGGVTRTTNEAILLCEGAGYDIILIETVGVGQSEFAVA DMVDMFVLLLPPAGGDELQGIKRGIIEMADLVAVTKSDGDLIVPARRIQAEY VSALKLLRKRSQVWKPKVIRISARSGEGISEMWDKMKDFQDLMLASGELT AKRRKQQKVWMWNLIQESVLEHFRTHPTVREQIPLLEQKVLIGALSPGLAA DFLLKAFKSRD [SEQ ID NO:169]	deficiência de cobalamina A (acidemia metilmalônica)	Acidemia orgânica
MMAB	326625	0139428	Q96EY8	MAVCGLGSRLGLGSRLGLRGCFGAARLLYPRFQSRGPQGVEDGDRPQP SSKTPRIPKIYTKTGDKGFSSTFTGERRPKDDQVFEAVGTTDELSSAIGFAL ELVTEKGHTFAEELQKIQCTLQDVGSALATPCSSAREAHLKYTTFKAGPILE LEQWIDKYTSQLPPLTAFILPSGGKISSALHFCRAVCRRAERRVVPLVQMG	deficiência de cobalamina B (acidemia metilmalônica)	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ETDANVAKFLNRLSDYLFTLARYAAMKEGNQEKIYMKNDPSAESEGL [SEQ ID NO:170]		
MMA-CHC	25974	0132763	A0A0C4DGU2, Q9Y4U1	MFDRALKPFLQSCHLRMLTDPVDQCVAYHLGRVRESLPELQIEIIADYEVH PNRRPKILAQTAAHVAGAAYYYQRQDVEADPWGNQRISGVCIHPRFGGW FAIRGVVLLPGIEVPDLPPRKPHDCVPTRADRIALLEGFNFHWRDWTYRDA VTPQERYSEEQKAYFSTPPAQRLALLGLAQPSEKPSSPSPDLPFTTPAPK KPGNPSRARSWLSPRVSPPASPGP [SEQ ID NO:171]	deficiência de cobalamina C (acidemia metilmalônica); Acidemia Metilmalônica com Homocistinúria	Acidemia orgânica
MMADHC	27249	0168288	Q9H3L0	MANVLCNRARLVSYLPGFCSLVKRVVNPKAFSTAGSSGSDESHVAAAPP DICSRTVWPDETMGPFGPQDQRFQLPGNIGFDCHLNGTASQKKSLVHKTL PDVLAEPLSSERHEFVMAQYVNEFQGNDAPVEQEINSAETYFESARVECA IQTCPELLRKDFESLFPEVANGKLMILTVTQKTKNDMTVWSEEVEIEREVLL EKFINGAKEICYALRAEGYWADFIDPSSGLAFFGPYTNNTLFETDERYRHL GFSVDDLGCCKVIRHSLWGTHVVVGSIFTNATPDSHIMKKLSGN [SEQ ID NO:172]	deficiência de cobalamina D (acidemia metilmalônica); Acidemia Metilmalônica com Homocistinúria; Homocistinúria; Deficiência de cobalamina C	Acidemia orgânica
MCEE	84693	0124370	Q96PE7	MARVLKAAAANAVGLFSRLQAPIPTVRASSTSQPLDQVTGSVWNLGRLNH VAIAVPDLEKAAAFYKNILGAQVSEAVPLPEHGVSVVFVNLGNTKMELLHP LGRDSPIAGFLQKNKAGGMHHICIEVDNINAAVMDLKKKKIRSLSEEVKIGA HGKPVIFLHPKDCGGVLVELEQA [SEQ ID NO:173]	acidemia metilmalônica; Deficiência de Cobalamina D	Acidemia orgânica
PCCA	5095	0175198	P05165	MAGFWVGTAPLVAAGRRGRWPPQQLMLSAALRTLKHVLYYSRQCLMVS RNLGSVGYDPNEKTFDKILVANRGEIACRVIRTCKKMGIKTVAIHSDVDASS VHVKMADEAVCVGPAPTSKSYLNMDAIMEAIKKTRAQAVHPGYGFLSENK EFARCLAAEDVVFIGPDTHAIQAMGDKIESKLLAKKAEVNTIPGFDGVVKDA	acidemia propiônica	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				EEAVRIAREIGYPVMIKASAGGGGKGMRIAWDDEETRDGFRLSSQEAASS FGDDRLLIEKFIDNPRHIEIQVLGDKHGNALWLNERECSIQRRNQKVVEEA PSIFLDAETRRAMGEQAVALARAVKYSSAGTVEFLVDSKKNFYFLEMNTRL QVEHPVTECITGLDLVQEMIRVAKGYPLRHKQADIRINGWAVECRVYAEDP YKSFGLPSIGRLSQYQEPLHLPGVRVDSGIQPGSDISIYYDPMISKLITYGS DRTEALKRMADALDNYVIRGVTHNIALLREVIINSRFVKGDISTKFLSDVYPD GFKGHMLTKSEKNQLLAIASSLFVAFQLRAQHFQENSRMPVIKPDIANWEL SVKLHDKVHTVVASNNGSVFSVEVDGSKLNVTSTWNLASPLLSVSVDGTQ RTVQCLSREAGGNMSIQFLGTVYKVNILTRLAAELNKFMLEKVTEDTSSVL RSPMPGVVVAVSVKPGDAVAEGQEICVIEAMKMQNSMTAGKTGTVKSVH CQAGDTVGEGDLLVELE [SEQ ID NO:174]		
PCCB	5096	0114054	P05166	MAAALRVAAVGARLSVLASGLRAAVRSLCSQATSVNERIENKRRTALLGG GQRRIDAQHKRGKLTARERISLLLDPGSFVESDMFVEHRCADFGMAADKN KFPGDSVVTGRGRINGRLVYVFSQDFTVFGGSLSGAHAQKICKIMDQAITV GAPVIGLNDSGGARIQEGVESLAGYADIFLRNVTASGVIPQISLIMGPCAGG AVYSPALTDFTFMVKDTSYLFITGPDVVKSVTNEDVTQEELGGAKTHTTMS GVAHRAFENDVDALCNLRDFFNYLPLSSQDPAPVRECHDPSDRLVPELDT IVPLESTKAYNMVDIIHSVVDEREFFEIMPNYAKNIIVGFARMNGRTVGIVGN QPKVASGCLDINSSVKGARFVRFCDAFNIPLITFVDVPGFLPGTAQEYGGII RHGAKLLYAFAEATVPKVTVITRKAYGGAYDVMSSKHLCGDTNYAWPTAE IAVMGAKGAVEIIFKGHENVEAAQAEYIEKFANPFPAAVRGFVDDIIQPSST RARICCDLDVLASKKVQRPWRKHANIPL [SEQ ID NO:175]	acidemia propiônica	Acidemia orgânica
UGT1A1	54658	0241635	P22309, Q5DT03	MAVESQGGRPLVLGLLLCVLGPVVSHAGKILLIPVDGSHWLSMLGAIQQLQ QRGHEIVVLAPDASLYIRDGAFYTLKTYPVPFQREDVKESFVSLGHNVFEN DSFLQRVIKTYKKIKKDSAMLLSGCSHLLHNKELMASLAESSFDVMLTDPF LPCSPIVAQYLSLPTVFFLHALPCSLEFEATQCPNPFSYVPRPLSSHSDHM TFLQRVKNMLIAFSQNFLCDVVYSPYATLASEFLQREVTVQDLLSSASVWL FRSDFVKDYPRPIMPNMVFVGGINCLHQNPLSQEFEAYINASGEHGIVVFS LGSMVSEIPEKKAMAIADALGKIPQTVLWRYTGTRPSNLANNTILVKWLPQ NDLLGHPMTRAFITHAGSHGVYESICNGVPMVMMPLFGDQMDNAKRMET KGAGVTLNVLEMTSEDLENALKAVINDKSYKENIMRLSSLHKDRPVEPLDL	Síndrome de Crigler-Najjar tipo 1 Síndrome de Crigler-Najjar tipo 2, síndrome de Gilbert	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				AVFWVEFVMRHKGAPHLRPAAHDLTWYQYHSLDVIGFLLAVVLTVAFITFK CCAYGYRKCLGKKGRVKKAHKSKTH [SEQ ID NO:176]		
ASS1	445	0130707	P00966, Q5T6L4	MSSKGSVVLAYSGGLDTSCILVWLKEQGYDVIAYLANIGQKEDFEEARKKA LKLGAKKVFIEDVSREFVEEFIWPAIQSSALYEDRYLLGTSLARPCIARKQV EIAQREGAKYVSHGATGKGNDQVRFELSCYSLAPQIKVIAPWRMPEFYNR FKGRNDLMEYAKQHGIPIPVTPKNPWSMDENLMHISYEAGILENPKNQAP PGLYTKTQDPAKAPNTPDILEIEFKKGVPVKVTNVKDGTTHQTSLELFMYL NEVAGKHGVGRIDIVENRFIGMKSRGIYETPAGTILYHAHLDIEAFTMDREV RKIKQGLGLKFAELVYTGFWHSPECEFVRHCIAKSQERVEGKVQVSVLKG QVYILGRESPLSLYNEELVSMNVQGDYEPTDATGFININSLRLKEYHRLQS KVTAK [SEQ ID NO:177]	citrulinemia tipo I	Desordem do ciclo da ureia
PAH	5053	0171759	A0A024RBG4, P00439	MSTAVLENPGLGRKLSDFGQETSYIEDNCNQNGAISLIFSLKEEVGALAKV LRLFEENDVNLTHIESRPSRLKKDEYEFFTHLDKRSLPALTNIIKILRHDIGAT VHELSRDKKKDTVPWFPRTIQELDRFANQILSYGAELDADHPGFKDPVYR ARRKQFADIAYNYRHGQPIPRVEYMEEEKKTWGTVFKTLKSLYKTHACYE YNHIFPLLEKYCGFHEDNIPQLEDVSQFLQTCTGFRLRPVAGLLSSRDFLG GLAFRVFHCTQYIRHGSKPMYTPEPDICHELLGHVPLFSDRSFAQFSQEIG LASLGAPDEYIEKLATIYWFTVEFGLCKQGDSIKAYGAGLLSSFGELQYCLS EKPKLLPLELEKTAIQNYTVTEFQPLYYVAESFNDAKEKVRNFAATIPRPFS VRYDPYTQRIEVLDNTQQLKILADSINSEIGILCSALQKIK [SEQ ID NO:178]	Deficiência de fenilalanina hidro-xilase	Aminoacidopatia
PAL				MAKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSL TNNTDILQGIQASCDYINNAVESGEPIYGVTSGFGGMANVAISREQASELQ TNLVWFLKTGAGNKLPLADVRAAMLLRANSHMRGASGIRLELIKRMEIFLN AGVTPYVYEFGSIGASGDLVPLSYITGSLIGLDPSFKVDFNGKEMDAPTAL RQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCVYDTQILTAIAMGVHALDI QALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLANS QLVRDELDGKHDYRDHELIQDRYSLRCLPQYLGPIVDGISQIAKQIEIEINSV TDNPLIDVDNQASYHGGNFLGQYVGMGMDHLRYYIGLLAKHLDVQIALLA SPEFSNGLPPSLLGNRERKVNMGLKGLQICGNSIMPLLTFYGNSIADRFPT HAEQFNQNINSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTG	Deficiência de fenilalanina hidro-xilase	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				HYDARASLSPATERLYSAVRHVVGQKPTSDRPYIWNDNEQGLDEHIARIS ADIAAGGVIVQAVQDILPSLH [SEQ ID NO:179]		
ATP8B 1	5205	0081923	O43520	MSTERDSETTFDEDSQPNDEVVPYSDDETEDELDDQGSAVEPEQNRVNR EAEENREPFRKECTWQVKANDRKYHEQPHFMNTKFLCIKESKYANNAIKT YKYNAFTFIPMNLFEQFKRAANLYFLALLILQAVPQISTLAWYTTLVPLLVVL GVTAIKDLVDDVARHKMDKEINNRTCEVIKDGRFKVAKWKEIQVGDVIRLK KNDFVPADILLLSSSEPNSLCYVETAELDGETNLKFKMSLEITDQYLQREDT LATFDGFIECEEPNNRLDKFTGTLFWRNTSFPLDADKILLRGCVIRNTDFCH GLVIFAGADTKIMKNSGKTRFKRTKIDYLMNYMVYTIFVVLILLSAGLAIGHA YWEAQVGNSSWYLYDGEDDTPSYRGFLIFWGYIIVLNTMVPISLYVSVEVI RLGQSHFINWDLQMYYAEKDTPAKARTTTLNEQLGQIHYIFSDKTGTLTQN IMTFKKCCINGQIYGDHRDASQHNHNKIEQVDFSWNTYADGKLAFYDHYLI EQIQSGKEPEVRQFFFLLAVCHTVMVDRTDGQLNYQAASPDEGALVNAA RNFGFAFLARTQNTITISELGTERTYNVLAILDFNSDRKRMSIIVRTPEGNIK LYCKGADTVIYERLHRMNPTKQETQDALDIFANETLRTLCLCYKEIEEKEFT EWNKKFMAASVASTNRDEALDKVYEEIEKDLILLGATAIEDKLQDGVPETIS KLAKADIKIWVLTGDKKETAENIGFACELLTEDTTICYGEDINSLLHARMEN QRNRGGVYAKFAPPVQESFFPPGGNRALIITGSWLNEILLEKKTKRNKILKL KFPRTEEERRMRTQSKRRLEAKKEQRQKNFVDLACECSAVICCRVTPKQ KAMVVDLVKRYKKAITLAIGDGANDVNMIKTAHIGVGISGQEGMQAVMSSD YSFAQFRYLQRLLLVHGRWSYIRMCKFLRYFFYKNFAFTLVHFWYSFFNG YSAQTAYEDWFITLYNVLYTSLPVLLMGLLDQDVSDKLSLRFPGLYIVGQR DLLFNYKRFFVSLLHGVLTSMILFFIPLGAYLQTVGQDGEAPSDYQSFAVTI ASALVITVNFQIGLDTSYWTFVNAFSIFGSIALYFGIMFDFHSAGIHVLFPSA FQFTGTASNALRQPYIWLTIILAVAVCLLPVVAIRFLSMTIWPSESDKIQKHR KRLKAEEQWQRRQQVFRRGVSTRRSAYAFSHQRGYADLISSGRSIRKKR SPLDAIVADGTAEYRRTGDS [SEQ ID NO:180]		
ABCB 11	8647	0073734,0276582	O95342	MSDSVILRSIKKFGEENDGFESDKSYNNDKKSRLQDEKKGDGVRVGFFQL FRFSSSTDIWLMFVGSLCAFLHGIAQPGVLLIFGTMTDVFIDYDVELQELQI PGKACVNNTIVWTNSSLNQNMTNGTRCGLLNIESEMIKFASYYAGIAVAVLI TGYIQICFWVIAAARQIQKMRKFYFRRIMRMEIGWFDCNSVGELNTRFSDD	Colestase intra-hepática familiar progressiva Tipo 2;	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				INKINDAIADQMALFIQRMTSTICGFLLGFFRGWKLTLVIISVSPLIGIGAATIG LSVSKFTDYELKAYAKAGVVADEVISSMRTVAAFGGEKREVERYEKNLVF AQRWGIRKGIVMGFFTGFVWCLIFLCYALAFWYGSTLVLDEGEYTPGTLV QIFLSVIVGALNLGNASPCLEAFATGRAAATSIFETIDRKPIIDCMSEDGYKL DRIKGEIEFHNVTFHYPSRPEVKILNDLNMVIKPGEMTALVGPSGAGKSTAL QLIQRFYDPCEGMVTVDGHDIRSLNIQWLRDQIGIVEQEPVLFSTTIAENIR YGREDATMEDIVQAAKEANAYNFIMDLPQQFDTLVGEGGGQMSGGQKQ RVAIARALIRNPKILLLDMATSALDNESEAMVQEVLSKIQHGHTIISVAHRLS TVRAADTIIGFEHGTAVERGTHEELLERKGVYFTLVTLQSQGNQALNEEDI KDATEDDMLARTFSRGSYQDSLRASIRQRSKSQLSYLVHEPPLAVVDHKS TYEEDRKDKDIPVQEEVEPAPVRRILKFSAPEWPYMLVGSVGAAVNGTVT PLYAFLFSQILGTFSIPDKEEQRSQINGVCLLFVAMGCVSLFTQFLQGYAFA KSGELLTKRLRKFGFRAMLGQDIAWFDDLRNSPGALTTRLATDASQVQGA AGSQIGMIVNSFTNVTVAMIIAFSFSWKLSLVILCFFPFLALSGATQTRMLTG FASRDKQALEMVGQITNEALSNIRTVAGIGKERRFIEALETELEKPFKTAIQK ANIYGFCFAFAQCIMFIANSASYRYGGYLISNEGLHFSYVFRVISAVVLSAT ALGRAFSYTPSYAKAKISAARFFQLLDRQPPISVYNTAGEKWDNFQGKIDF VDCKFTYPSRPDSQVLNGLSVSISPGQTLAFVGSSGCGKSTSIQLLERFYD PDQGKVMIDGHDSKKVNVQFLRSNIGIVSQEPVLFACSIMDNIKYGDNTKEI PMERVIAAAKQAQLHDFVMSLPEKYETNVGSQGSQLSRGEKQRIAIARAIV RDPKILLLDEATSALDTESEKTVQVALDKAREGRTCIVIAHRLSTIQNADIIAV MAQGVVIEKGTHEELMAQKGAYYKLVTTGSPIS [SEQ ID NO:181]	Colestase intra-hepática familiar progressiva Tipo 3	
ABCB 4	5244	0005471	P21439	MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKKTKTVKMIGVLTLFRYSD WQDKLFMSLGTIMAIAHGSGLPLMMIVFGEMTDKFVDTAGNFSFPVNFSL SLLNPGKILEEEMTRYAYYYSGLGAGVLVAAYIQVSFWTLAAGRQIRKIRQ KFFHAILRQEIGWFDINDTTELNTRLTDDISKISEGIGDKVGMFFQAVATFFA GFIVGFIRGWKLTLVIMAISPILGLSAAVWAKILSAFSDKELAAYAKAGAVAE EALGAIRTVIAFGGQNKELERYQKHLENAKEIGIKKAISANISMGIAFLLIYAS YALAFWYGSTLVISKEYTIGNAMTVFFSILIGAFSVGQAAPCIDAFANARGA AYVIFDIIDNNPKIDSFSERGHKPDSIKGNLEFNDVHFSYPSRANVKILKGLN LKVQSGQTVALVGSSGCGKSTTVQLIQRLYDPDEGTINIDGQDIRNFNVNY LREIIGVVSQEPVLFSTTIAENICYGRGNVTMDEIKKAVKEANAYEFIMKLPQ	Colestase intra-hepática familiar progressiva Tipo 3; Colestase intra-hepática familiar progressiva Tipo 2	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAEV QAALDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVIVEQGSHSELMKKEG VYFKLVNMQTSGSQIQSEEFELNDEKAATRMAPNGWKSRLFRHSTQKNL KNSQMCQKSLDVETDGLEANVPPVSFLKVLKLNKTEWPYFVVGTVCAIAN GGLQPAFSVIFSEIIAIFGPGDDAVKQQKCNIFSLIFLFLGIISFFTFFLQGFTF GKAGEILTRRLRSMAFKAMLRQDMSWFDDHKNSTGALSTRLATDAAQVQ GATGTRLALIAQNIANLGTGIIISFIYGWQLTLLLLAVVPIIAVSGIVEMKLLAG NAKRDKKELEAAGKIATEAIENIRTVVSLTQERKFESMYVEKLYGPYRNSV QKAHIYGITFSISQAFMYFSYAGCFRFGAYLIVNGHMRFRDVILVFSAIVFG AVALGHASSFAPDYAKAKLSAAHLFMLFERQPLIDSYSEEGLKPDKFEGNI TFNEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLER FYDPLAGTVFVDFGFQLLDGQEAKKLNVQWLRAQLGIVSQEPILFDCSIAE NIAYGDNSRVVSQDEIVSAAKAANIHPFIETLPHKYETRVGDKGTQLSGGQ KQRIAIARALIRQPQILLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHR LSTIQNADLIVVFQNGRVKEHGTHQQLLAQKGIYFSMVSVQAGTQNL [SEQ ID NO:184]		
TJP2	9414	0119139	B7Z2R3, Q9UDY2, B7Z954	MPVRGDRGFPPRRELSGWLRAPGMEELIWEQYTVTLQKDSKRGFGIAVS GGRDNPHFENGETSIVISDVLPGGPADGLLQENDRVVMVNGTPMEDVLH SFAVQQLRKSGKVAAIVVKRPRKVQVAALQASPPLDQDDRAFEVMDEFD GRSFRSGYSERSRLNSHGGRSRSWEDSPERGRPHERARSRERDLSRDR SRGRSLERGLDQDHARTRDRSRGRSLERGLDHDFGPSRDRDRDRSRGR SIDQDYERAYHRAYDPDYERAYSPEYRRGARHDARSRGPRSRSREHPHS RSPSPEPRGRPGPIGVLLMKSRANEEYGLRLGSQIFVKEMTRTGLATKDG NLHEGDIILKINGTVTENMSLTDARKLIEKSRGKLQLVVLRDSQQTLINIPSL NDSDSEIEDISEIESNRSFSPEERRHQYSDYDYHSSSEKLKERPSSREDTP SRLSRMGATPTPFKSTGDIAGTVVPETNKEPRYQEDPPAPQPKAAPRTFL RPSPEDEAIYGPNTKMVRFKKGDSVGLRLAGGNDVGIFVAGIQEGTSAEQ EGLQEGDQILKVNTQDFRGLVREDAVLYLLEIPKGEMVTILAQSRADVYRDI LACGRGDSFFIRSHFECEKETPQSLAFTRGEVFRVVDTLYDGKLGNWLAV RIGNELEKGLIPNKSRAEQMASVQNAQRDNAGDRADFWRMRGQRSGVK KNLRKSREDLTAVVSVSTKFPAYERVLLREAGFKRPVVLFGPIADIAMEKLA NELPDWFQTAKTEPKDAGSEKSTGVVRLNTVRQIIEQDKHALLDVTPKAV	Colestase intra-hepática familiar progressiva Tipo 4	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DLLNYTQWFPIVIFFNPDSRQGVKTMRQRLNPTSNKSSRKLFDQANKLKK TCAHLFTATINLNSANDSWFGSLKDTIQHQQGEAVWVSEGKMEGMDDDP EDRMSYLTAMGADYLSCDSRLISDFEDTDGEGGAYTDNELDEPAEEPLVS SITRSSEPVQHEESIRKPSPEPRAQMRRAASSDQLRDNSPPPAFKPEPPK AKTQNKEESYDFSKSYEYKSNPSAVAGNETPGASTKGYPPPVAAKPTFG RSILKPSTPIPPQEGEEVGESSEEQDNAPKSVLGKVKIFEKMDHKARLQRM QELQEAQNARIEIAQKHPDIYAVPIKTHKPDPGTPQHTSSRPPEPQKAPSR PYQDTRGSYGSDAEEEEYRQQLSEHSKRGYYGQSARYRDTEL [SEQ ID NO:183]		
IVD	3712	0128928	P26440, A0A0A0MT83	MATATRLLGWRVASWRLRPPLAGFVSQRAHSLLPVDDAINGLSEEQRQL RQTMAKFLQEHLAPKAQEIDRSNEFKNLREFWKQLGNLGVLGITAPVQYG GSGLGYLEHVLVMEEISRASGAVGLSYGAHSNLCINQLVRNGNEAQKEKY LPKLISGEYIGALAMSEPNAGSDVVSMKLKAEKKGNHYILNGNKFWITNGP DADVLIVYAKTDLAAVPASRGITAFIVEKGMPGFSTSKKLDKLGMRGSNTC ELIFEDCKIPAANILGHENKGVYVLMSGLDLERLVLAGGPLGLMQAVLDHTI PYLHVREAFGQKIGHFQLMQGKMADMYTRLMACRQYVYNVAKACDEGH CTAKDCAGVILYSAECATQVALDGIQCFGGNGYINDFPMGRFLRDAKLYEI GAGTSEVRRLVIGRAFNADFH [SEQ ID NO:184]	acidemia isovalérica (IVD)	Acidemia orgânica
GCDH	2639	0105607	A0A024R7F9, Q92947	MALRGVSVRLLSRGPGLHVLRTWVSSAAQTEKGGRTQSQLAKSSRPEFD WQDPLVLEEQLTTDEILIRDTFRTYCQERLMPRILLANRNEVFHREIISEMG ELGVLGPTIKGYGCAGVSSVAYGLLARELERVDSGYRSAMSVQSSLVMHP IYAYGSEEQRQKYLPQLAKGELLGCFGLTEPNSGSDPSSMETRAHYNSSN KSYTLNGTKTWITNSPMADLFVVWARCEDGCIRGFLLEKGMRGLSAPRIQ GKFSLRASATGMIIMDGVEVPEENVLPGASSLGGPFGCLNNARYGIAWGV LGASEFCLHTARQYALDRMQFGVPLARNQLIQKKLADMLTEITLGLHACLQ LGRLKDQDKAAPEMVSLLKRNNCGKALDIARQARDMLGGNGISDEYHVIR HAMNLEAVNTYEGTHDIHALILGRAITGIQAFTASK [SEQ ID NO:185]	acidemia glutárica tipo I	Acidemia orgânica
ETFA	2108	0140374	A0A0S2Z3L0, P13804	MFRAAAPGQLRRAASLLRFQSTLVIAEHANDSLAPITLNTITAATRLGGEVS CLVAGTKCDKVAQDLCKVAGIAKVLVAQHDVYKGLLPEELTPLILATQKQF NYTHICAGASAFGKNLLPRVAAKLEVAPISDIIAIKSPDTFVRTIYAGNALCTV	deficiência múltipla de acil-CoA desidrogenase (tam-	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KCDEKVKVFSVRGTSFDAAATSGGSASSEKASSTSPVEISEWLDQKLTKS DRPELTGAKVVVSGGRGLKSGENFKLLYDLADQLHAAVGASRAAVDAGF VPNDMQVGQTGKIVAPELYIAVGISGAIQHLAGMKDSKTIVAINKDPEAPIF QVADYGIVADLFKVVPEMTEILKKK [SEQ ID NO:186]	bém conhecida como acidúria glutárica tipo II)	
ETFB	2109	0105379	P38117	MAELRVLVAVKRVIDYAVKIRVKPDRTGVVTDGVKHSMNPFCEIAVEEAVR LKEKKLVKEVIAVSCGPAQCQETIRTALAMGADRGIHVEVPPAEAERLGPL QVARVLAKLAEKEKVDLVLLGKQAIDDDCNQTGQMTAGFLDWPQGTFAS QVTLEGDKLKVEREIDGGLETLRLKLPAVVTADLRLNEPRYATLPNIMKAKK KKIEVIKPGDLGVDLTSKLSVISVEDPPQRTAGVKVETTEDLVAKLKEIGRI [SEQID NO:187]	deficiência múltipla de acil-CoA desidrogenase (também conhecida como acidúria glutárica tipo II)	Acidemia orgânica
ET-FDH	2110	0171503	B4DEQ0, Q16134	MLVPLAKLSCLAYQCFHALKIKKNYLPLCATRWSSTSTVPRITTHYTIYPRD KDKRWEGVNMERFAEEADVVIVGAGPAGLSAAVRLKQLAVAHEKDIRVCL VEKAAQIGAHTLSGACLDPGAFKELFPDWKEKGAPLNTPVTEDRFGILTEK YRIPVPILPGLPMNNHGNYIVRLGHLVSWMGEQAEALGVEVYPGYAAAEV LFHDDGSVKGIATNDVGIQKDGAPKATFERGLELHAKVTIFAEGCHGHLAK QLYKKFDLRANCEPQTYGIGLKELWVIDEKNWKPGRVDHTVGWPLDRHT YGGSFLYHLNEGEPLVALGLVVGLDYQNPYLSPFREFQRWKHHPSIRPTL EGGKRIAYGARALNEGGFQSIPKLTFPGGLLIGCSPGFMNVPKIKGTHTAM KSGILAAESIFNQLTSENLQSKTIGLHVTEYEDNLKNSWVWKELYSVRNIRP SCHGVLGVYGGMIYTGIFYWILRGMEPWTLKHKGSDFERLKPAKDCTPIE YPKPDGQISFDLLSSVALSGTNHEHDQPAHLTLRDDSIPVNRNLSIYDGPE QRFCPAGVYEFVPVEQGDGFRLQINAQNCVHCKTCDIKDPSQNINWVVPE GGGGPAYNGM [SEQ ID NO:188]	deficiência múltipla de acil-CoA desidrogenase (também conhecida como acidúria glutárica tipo II)	Acidemia orgânica
ASL	435	0126522	A0A024RDL8, P04424, A0A0S2Z316	MASESGKLWGGRFVGAVDPIMEKFNASIAYDRHLWEVDVQGSKAYSRGL EKAGLLTKAEMDQILHGLDKVAEEWAQGTFKLNSNDEDIHTANERRLKELI GATAGKLHTGRSRNDQVVTDLRLWMRQTCSTLSGLLWELIRTMVDRAEA ERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSERLLEVRKRINVLPLG SGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWASLCMT HLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIRSKAGR VFGRCAGLLMTLKGLPSTYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTL	deficiência de argininosuccinato liase (ASL)	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QIHQENMGQALSPDMLATDLAYYLVRKGMPFRQAHEASGKAVFMAETKG VALNQLSLQELQTISPLFSGDVICVWDYGHSVEQYGALGGTARSSVDWQI RQVRALLQAQQA [SEQ ID NO:189]		
D2HG DH	728294	0180902	B3KSR6, B4E3K7, B5MCV2, Q8N465	MVGGSVPVFDEIILSTARMNRVLSFHSVSGILVCQAGCVLEELSRYVEERD FIMPLDLGAKGSCHIGGNVATNAGGLRFLRYGSLHGTVLGLEVVLADGTVL DCLTSLRKDNTGYDLKQLFIGSEGTLGIITTVSILCPPKPRAVNVAFLGCPG FAEVLQTFSTCKGMLGEILSAFEFMDAVCMQLVGRHLHLASPVQESPFYV LIETSGSNAGHDAEKLGHFLEHALGSGLVTDGTMATDQRKVKMLWALRE RITEALSRDGYVYKYDLSLPVERLYDIVTDLRARLGPHAKHVVGYGHLGDG NLHLNVTAEAFSPSLLAALEPHVYEWTAGQQGSVSAEHGVGFRKRDVLG YSKPPGALQLMQQLKALLDPKGILNPYKTLPSQA [SEQ ID NO:190]	Acidúria D-2-hidroxiglutárica tipo I	Acidemia orgânica
HMGC L	3155	0117305	P35914	MAAMRKALPRRLVGLASLRAVSTSSMGTLPKRVKIVEVGPRDGLQNEKNI VSTPVKIKLIDMLSEAGLSVIETTSFVSPKWVPQMGDHTEVLKGIQKFPGIN YPVLTPNLKGFEAAVAAGAKEVVIFGAASELFTKKNINCSIEESFQRFDAILK AAQSANISVRGYVSCALGCPYEGKISPAKVAEVTKKFYSMGCYEISLGDTI GVGTPGIMKDMLSAVMQEVPLAALAVHCHDTYGQALANTLMALQMGVSV VDSSVAGLGGCPYAQGASGNLATEDLVYMLEGLGIHTGVNLQKLLEAGNF ICQALNRKTSSKVAQATCKL [SEQ ID NO:191]	Deficiência de 3-hidroxi-3-metilglutaril-CoA liase (3HMG)	Acidemia orgânica Desordem do ciclo da ureia
MCCC 1	56922	0078070	Q68D27, Q96RQ3, A0A0S2Z693, E9PHF7	MAAASAVSVLLVAAERNRWHRLPSLLLPPRTWVWRQRTMKYTTATGRNI TKVLIANRGEIACRVMRTAKKLGVQTVAVYSEADRNSMHVDMADEAYSIG PAPSQQSYLSMEKIIQVAKTSAAQAIHPGCGFLSENMEFAELCKQEGIIFIG PPPSAIRDMGIKSTSKSIMAAAGVPVVEGYHGEDQSDQCLKEHARRIGYP VMIKAVRGGGGKGMRIVRSEQEFQEQLESARREAKKSFNDDAMLIEKFVD TPRHVEVQVFGDHHGNAVYLFERDCSVQRRHQKIIEEAPAPGIKSEVRKK LGEAAVRAAKAVNYVGAGTVEFIMDSKHNFCFMEMNTRLQVEHPVTEMIT GTDLVEWQLRIAAGEKIPLSQEEITLQGHAFEARIYAEDPSNNFMPVAGPL VHLSTPRADPSTRIETGVRQGDEVSVHYDPMIAKLVVWAADRQAALTKLR YSLRQYNIVGLHTNIDFLLNLSGHPEFEAGNVHTDFIPQHHKQLLLSRKAAA KESLCQAALGLILKEKAMTDTFTLQAHDQFSPFSSSSGRRLNISYTRNMTL KDGKNNVAIAVTYNHDGSYSMQIEDKTFQVLGNLYSEGDCTYLKCSVNGV	Deficiência de 3-metilcrotonil-CoA carboxilase (3MCC)	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ASKAKLIILENTIYLFSKEGSIEIDIPVPKYLSSVSSQETQGGPLAPMTGTIEK VFVKAGDKVKAGDSLMVMIAMKMEHTIKSPKDGTVKKVFYREGAQANRH TPLVEFEEEESDKRESE [SEQID NO:192]		
MCCC 2	64087	0131844,0281742,0275300	A0A140VK29, Q9HCC0	MWAVLRLALRPCARASPAGPRAYHGDSVASLGTQPDLGSALYQENYKQ MKALVNQLHERVEHIKLGGGEKARALHISRGKLLPRERIDNLIDPGSPFLEL SQFAGYQLYDNEEVPGGGIITGIGRVSGVECMIIANDATVKGGAYYPVTVK KQLRAQEIAMQNRLPCIYLVDSGGAYLPRQADVFPDRDHFGRTFYNQAIM SSKNIAQIAVVMGSCTAGGAYVPAMADENIIVRKQGTIFLAGPPLVKAATGE EVSAEDLGGADLHCRKSGVSDHWALDDHHALHLTRKVVRNLNYQKKLDV TIEPSEEPLFPADELYGIVGANLKRSFDVREVIARIVDGSRFTEFKAFYGDT LVTGFARIFGYPVGIVGNNGVLFSESAKKGTHFVQLCCQRNIPLLFLQNITG FMVGREYEAEGIAKDGAKMVAAVACAQVPKITLIIGGSYGAGNYGMCGRA YSPRFLYIWPNARISVMGGEQAANVLATITKDQRAREGKQFSSADEAALK EPIIKKFEEEGNPYYSSARVWDDGIIDPADTRLVLGLSFSAALNAPIEKTDF GIFRM [SEQ ID NO:193]	Deficiência de 3-metilcrotonil-CoA carboxilase (3MCC)	Acidemia orgânica
ABCD 4	5826	0119688	A0A024R6B9, O14678, A0A024R6C8	MAVAGPAPGAGARPRLDLQFLQRFLQILKVLFPSWSSQNALMFLTLLCLTL LEQFVIYQVGLIPSQYYGVLGNKDLEGFKTLTFLAVMLIVLNSTLKSFDQFT CNLLYVSWRKDLTEHLHRLYFRGRAYYTLNVLRDDIDNPDQRISQDVERF CRQLSSMASKLIISPFTLVYYTYQCFQSTGWLGPVSIFGYFILGTVVNKTLM GPIVMKLVHQEKLEGDFRFKHMQIRVNAEPAAFYRAGHVEHMRTDRRLQ RLLQTQRELMSKELWLYIGINTFDYLGSILSYVVIAIPIFSGVYGDLSPAELST LVSKNAFVCIYLISCFTQLIDLSTTLSDVAGYTHRIGQLRETLLDMSLKSQDC EILGESEWGLDTPPGWPAAEPADTAFLLERVSISAPSSDKPLIKDLSLKISE GQSLLITGNTGTGKTSLLRVLGGLWTSTRGSVQMLTDFGPHGVLFLPQKP FFTDGTLREQVIYPLKEVYPDSGSADDERILRFLELAGLSNLVARTEGLDQ QVDWNWYDVLSPGEMQRLSFARLFYLQPKYAVLDEATSALTEEVESELY RIGQQLGMTFISVGHRQSLEKFHSLVLKLCGGGRWELMRIKVE [SEQ ID NO:194]	acidemia metilmalônica com homocistinúria	Acidemia orgânica
HCFC 1	3054	0172534	P51610, A6NEM2	MASAVSPANLPAVLLQPRWKRVVGWSGPVPRPRHGHRAVAIKELIVVFG GGNEGIVDELHVYNTATNQWFIPAVRGDIPPGCAAYGFVCDGTRLLVFGG	acidemia metilmalônica com homo-	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				MVEYGKYSNDLYELQASRWEWKRLKAKTPKNGPPPCPRLGHSFSLVGNK CYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWDIPITYGVLPPPR ESHTAVVYTEKDNKKSKLVIYGGMSGCRLGDLWTLDIDTLTWNKPSLSGV APLPRSLHSATTIGNKMYVFGGWVPLVMDDVKVATHEKEWKCTNTLACL NLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSGRDGYRKAWNN QVCCKDLWYLETEKPPPPARVQLVRANTNSLEVSWGAVATADSYLLQLQ KYDIPATAATATSPTPNPVPSVPANPPKSPAPAAAAPAVQPLTQVGITLLP QAAPAPPTTTTIQVLPTVPGSSISVPTAARTQGVPAVLKVTGPQATTGTPL VTMRPASQAGKAPVTVTSLPAGVRMVVPTQSAQGTVIGSSPQMSGMAAL AAAAAATQKIPPSSAPTVLSVPAGTTIVKTMAVTPGTTTLPATVKVASSPVM VSNPATRMLKTAAAQVGTSVSSATNTSTRPIITVHKSGTVTVAQQAQVVTT VVGGVTKTITLVKSPISVPGGSALISNLGKVMSVVQTKPVQTSAVTGQAST GPVTQIIQTKGPLPAGTILKLVTSADGKPTTIITTTQASGAGTKPTILGISSVS PSTTKPGTTTIIKTIPMSAIITQAGATGVTSSPGIKSPITIITTKVMTSGTGAPA KIITAVPKIATGHGQQGVTQVVLKGAPGQPGTILRTVPMGGVRLVTPVTVS AVKPAVTTLVVKGTTGVTTLGTVTGTVSTSLAGAGGHSTSASLATPITTLG TIATLSSQVINPTAITVSAAQTTLTAAGGLTTPTITMQPVSQPTQVTLITAPS GVEAQPVHDLPVSILASPTTEQPTATVTIADSGQGDVQPGTVTLVCSNPPC ETHETGTTNTATTTVVANLGGHPQPTQVQFVCDRQEAAASLVTSTVGQQ NGSVVRVCSNPPCETHETGTTNTATTATSNMAGQHGCSNPPCETHETGT TNTATTAMSSVGANHQRDARRACAAGTPAVIRISVATGALEAAQGSKSQC QTRQTSATSTTMTVMATGAPCSAGPLLGPSMAREPGGRSPAFVQLAPLS SKVRLSSPSIKDLPAGRHSHAVSTAAMTRSSVGAGEPRMAPVCESLQGG SPSTTVTVTALEALLCPSATVTQVCSNPPCETHETGTTNTATTSNAGSAQ RVCSNPPCETHETGTTHTATTATSNGGTGQPEGGQQPPAGRPCETHQTT STGTTMSVSVGALLPDATSSHRTVESGLEVAAAPSVTPQAGTALLAPFPT QRVCSNPPCETHETGTTHTATTVTSNMSSNQDPPPAASDQGEVESTQGD SVNITSSSAITTTVSSTLTRAVTTVTQSTPVPGPSVPPPEELQVSPGPRQQ LPPRQLLQSASTALMGESAEVLSASQTPELPAAVDLSSTGEPSSGQESAG SAVVATVVVQPPPPTQSEVDQLSLPQELMAEAQAGTTTLMVTGLTPEELA VTAAAEAAAQAAATEEAQALAIQAVLQAAQQAVMGTGEPMDTSEAAATVT QAELGHLSAEGQEGQATTIPIVLTQQELAALVQQQQLQEAQAQQQHHHLP	cistinúria	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				TEALAPADSLNDPAIESNCLNELAGTVPSTVALLPSTATESLAPSNTFVAPQ PVVVASPAKLQAAATLTEVANGIESLGVKPDLPPPPSKAPMKKENQWFDV GVIKGTNVMVTHYFLPPDDAVPSDDDLGTVPDYNQLKKQELQPGTAYKFR VAGINACGRGPFSEISAFKTCLPGFPGAPCAIKISKSPDGAHLTWEPPSVT SGKIIEYSVYLAIQSSQAGGELKSSTPAQLAFMRVYCGPSPSCLVQSSSLS NAHIDYTTKPAIIFRIAARNEKGYGPATQVRWLQETSKDSSGTKPANKRPM SSPEMKSAPKKSKADGQ [SEQ ID NO:195]		
LMBRD1	55788	0168216	Q9NUN5	MATSGAASAELVIGWCIFGLLLLAILAFCWIYVRKYQSRRESEVVSTITAIFS LAIALITSALLPVDIFLVSYMKNQNGTFKDWANANVSRQIEDTVLYGYYTLY SVILFCVFFWIPFVYFYYEEKDDDDTSKCTQIKTALKYTLGFVVICALLLLVG AFVPLNVPNNKNSTEWEKVKSLFEELGSSHGLAALSFSISSLTLIGMLAAIT YTAYGMSALPLNLIKGTRSAAYERLENTEDIEEVEQHIQTIKSKSKDGRPLP ARDKRALKQFEERLRTLKKRERHLEFIENSWWTKFCGALRPLKIVWGIFFIL VALLFVISLFLSNLDKALHSAGIDSGFIIFGANLSNPLNMLLPLLQTVFPLDYI LITIIIMYFIFTSMAGIRNIGIWFFWIRLYKIRRGRTRPQALLFLCMILLLIVLHTS YMIYSLAPQYVMYGSQNYLIETNITSDNHKGNSTLSVPKRCDADAPEDQC TVTRTYLFLHKFWFFSAAYYFGNWAFLGVFLIGLIVSCCKGKKSVIEGVDE DSDISDDEPSVYSA [SEQID NO:196]	acidemia metilmalônica com homocistinúria	Acidemia orgânica
ARG1	383	0118520	P05089	MSAKSRTIGIIGAPFSKGQPRGGVEEGPTVLRKAGLLEKLKEQECDVKDY GDLPFADIPNDSPFQIVKNPRSVGKASEQLAGKVAEVKKNGRISLVLGGDH SLAIGSISGHARVHPDLGVIWVDAHTDINTPLTTTSGNLHGQPVSFLLKELK GKIPDVPGFSWVTPCISAKDIVYIGLRDVDPGEHYILKTLGIKYFSMTEVDRL GIGKVMEETLSYLLGRKKRPIHLSFDVDGLDPSFTPATGTPVVGGLTYREG LYITEEIYKTGLLSGLDIMEVNPSLGKTPEEVTRTVNTAVAITLACFGLAREG NHKPIDYLNPPK [SEQ ID NO:197]	deficiência de arginase (ARG1)	Desordem do ciclo da ureia
SLC25A15	10166	0102743	Q9Y619	MKSNPAIQAAIDLTAGAAGGTACVLTGQPFDTMKVKMQTFPDLYRGLTDC CLKTYSQVGFRGFYKGTSPALIANIAENSVLFMCYGFCQQVVRKVAGLDK QAKLSDLQNAAAGSFASAFAALVLCPTELVKCRLQTMYEMETSGKIAKSQ NTVWSVIKSILRKDGPLGFYHGLSSTLLREVPGYFFFFGGYELSRSFFASG RSKDELGPVPLMLSGGVGGICLWLAVYPVDCIKSRIQVLSMSGKQAGFIRT	síndrome de hiperamonemia-hiperornitinemia-homocitrulinúria (HHH)	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FINVVKNEGITALYSGLKPTMIRAFPANGALFLAYEYSRKLMMNQLEAY [SEQ ID NO:198]		
SLC25A13	10165	0004864	Q9UJS0	MAAAKVALTKRADPAELRTIFLKYASIEKNGEFFMSPNDFVTRYLNIFGESQPNPKTVELLSGVVDQTKDGLISFQEFVAFESVLCAPDALFMVAFQLFDKAGKGEVTFEDVKQVFGQTTIHQHIPFNWDSEFVQLHFGKERKRHLTYAEFTQFLLEIQLEHAKQAFVQRDNARTGRVTAIDFRDIMVTIRPHVLTPFVEECLVAAAGGTTSHQVSFSYFNGFNSLLNNMELIRKIYSTLAGTRKDVEVTKEEFVLAAQKFGQVTPMEVDILFQLADLYEPRGRMTLADIERIAPLEEGTLPFNLAEAQRQKASGDSARPVLLQVAESAYRFGLGSVAGAVGATAVYPIDLVKTRMQNQRSTGSFVGELMYKNSFDCFKKVLRYEGFFGLYRGLLPQLLGVAPEKAIKLTVNDFVRDKFMHKDGSVPLAAEILAGGCAGGSQVIFTNPLEIVKIRLQVAGEITTGPRVSALSVVRDLGFFGIYKGAKACFLRDIPFSAIYFPCYAHVKASFANEDGQVSPGSLLLAGAIAGMPAASLVTPADVIKTRLQVAARAGQTTYSGVIDCFRKILREEGPKALWKGAGARVFRSSPQFGVTLLTYELLQRWFYIDFGGVKPMGSEPVPKSRINLPAPNPDHVGGYKLAVATFAGIENKFGLYLPLFKPSVSTSKAIGGGP [SEQ ID NO:199]	deficiência de citrina citrulinemia tipo II	Desordem do ciclo da ureia
ALAD	210	0148218	P13716	MQPQSVLHSGYFHPLLRAWQTATTTLNASNLIYPIFVTDVPDDIQPITSLPGVARYGVKRLEEMLRPLVEEGLRCVLIFGVPSRVPKDERGSAADSEESPAIEAIHLLRKTFPNLLVACDVCLCPYTSHGHCGLLSENGAFRAEESRQRLAEVALAYAKAGCQVVAPSDMMDGRVEAIKEALMAHGLGNRVSVMSYSAKFASCFYGPFRDAAKSSPAFGDRRCYQLPPGARGLALRAVDRDVREGADMLMVKPGMPYLDIVREVKDKHPDLPLAVYHVSGEFAMLWHGAQAGAFDLKAAVLEAMTAFRRAGADIIITYYTPQLLQWLKEE [SEQ ID NO:200]	Porfiria hepática aguda	Porfiria
CPOX	1371	0080819	P36551	MALQLGRLSSGPCWLVARGGCGGPRAWSQCGGGGLRAWSQRSAAGRVCRPPGPAGTEQSRGLGHGSTSRGGPWVGTGLAAALAGLVGLATAAFGHVQRAEMLPKTSGTRATSLGRPEEEEDELAHRCSSFMAPPVTDLGELRRRPGDMKTKMELLILETQAQVCQALAQVDGGANFSVDRWERKEGGGGISCVLQDGCVFEKAGVSISVVHGNLSEEAAKQMRSRGKVLKTKDGKLPFCAMGVSSVIHPKNPHAPTIHFNYRYFEVEEADGNKQWWFGGGCDLTPTYLNQEDAVHFHRTLKEACDQHGPDLYPKFKKWCDDYFFIAHRGERRGIGGIFFD	Porfiria hepática aguda	Porfiria

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DLDSPSKEEVFRFVQSCARAVVPSYIPLVKKHCDDSFTPQEKLWQQLRRG RYVEFNLLYDRGTKFGLFTPGSRIESILMSLPLTARWEYMHSPSENSKEAE ILEVLRHPRDWVR [SEQ ID NO:201]		
HMBS	3145	0256269,0281702	P08397	MSGNGNAAATAEENSPKMRVIRVGTRKSQLARIQTDSVVATLKASYPGLQ FEIIAMSTTGDKILDTALSKIGEKSLFTKELEHALEKNEVDLVVHSLKDLPTV LPPGFTIGAICKRENPHDAVVFHPKFVGKTLETLPEKSVVGTSSLRRAAQL QRKFPHLEFRSIRGNLNTRLRKLDEQQEFSAIILATAGLQRMGWHNRVGQI LHPEECMYAVGQGALGVEVRAKDQDILDLVGVLHDPETLLRCIAERAFLRH LEGGCSVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIHVPAQH EDGPEDDPQLVGITARNIPRGPQLAAQNLGISLANLLLSKGAKNILDVARQL NDAH [SEQ ID NO:202]	Porfiria hepática aguda; Porfiria aguda intermitente	Porfiria
PPOX	5498	0143224	P50336, B4DY76	MGRTVVVLGGGISGLAASYHLSRAPCPPKVVLVESSERLGGWIRSVRGPN GAIFELGPRGIRPAGALGARTLLLVSELGLDSEVLPVRGDHPAAQNRFLYV GGALHALPTGLRGLLRPSPPFSKPLFWAGLRELTKPRGKEPDETVHSFAQ RRLGPEVASLAMDSLCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLL LGAGRTPQPDSALIRQALAERWSQWSLRGGLEMLPQALETHLTSRGVSV LRGQPVCGLSLQAEGRWKVSLRDSSLEADHVISAIPASVLSELLPAEAAPL ARALSAITAVSVAVVNLQYQGAHLPVQGFGHLVPSSEDPGVLGIVYDSVAF PEQDGSPPGLRVTVMLGGSWLQTLEASGCVLSQELFQQRAQEAAATQLG LKEMPSHCLVHLHKNCIPQYTLGHWQKLESARQFLTAHRLPLTLAGASYE GVAVNDCIESGRQAAVSVLGTEPNS [SEQ ID NO:203]	Porfiria hepática aguda	Porfiria
BTD	686	0169814	P43251	MAHAHIQGGRRAKSRFVVCIMSGARSKLALFLCGCYVVALGAHTGEESVA DHHEAEYYVAAVYEHPSILSLNPLALISRQEALELMNQNLDIYEQQVMTAA QKDVQIIVFPEDGIHGFNFTRTSIYPFLDFMPSPQVVRWNPCLEPHRFNDT EVLQRLSCMAIRGDMFLVANLGTKEPCHSSDPRCPKDGRYQFNTNVVFS NNGTLVDRYRKHNLYFEAAFDVPLKVDLITFDTPFAGRFGIFTCFDILFFDP AIRVLRDYKVKHVVYPTAWMNQLPLLAAIEIQKAFAVAFGINVLAANVHHPV LGMTGSGIHTPLESFWYHDMENPKSHLIIAQVAKNPVGLIGAENATGETDP SHSKFLKILSGDPYCEKDAQEVHCDEATKWNVNAPPTFHSEMMYDNFTLV PVWGKEGYLHVCSNGLCCYLLYERPTLSKELYALGVFDGLHTVHGTYYIQ	Deficiência de biotinidase	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				VCALVRCGGLGFDTCGQEITEATGIFEFHLWGNFSTSYIFPLFLTSGMTLE VPDQLGWENDHYFLRKSRLSSGLVTAALYGRLYERD [SEQ ID NO:204]		
HLCS	3141	0159267	P50747	MEDRLHMDNGLVPQKIVSVHLQDSTLKEVKDQVSNKQAQILEPKPEPSLEI KPEQDGMEHVGRDDPKALGEEPKQRRGSASGSEPAGDSDRGGGPVEH YHLHLSSCHECLELENSTIESVKFASAENIPDLPYDYSSSLESVADETSPER EGRRVNLTGKAPNILLYVGSDSQEALGRFHEVRSVLADCVDIDSYILYHLLE DSALRDPWTDNCLLLVIATRESIPEDLYQKFMAYLSQGGKVLGLSSSFTFG GFQVTSKGALHKTVQNLVFSKADQSEVKLSVLSSGCRYQEGPVRLSPGR LQGHLENEDKDRMIVHVPFGTRGGEAVLCQVHLELPPSSNIVQTPEDFNL LKSSNFRRYEVLREILTTLGLSCDMKQVPALTPLYLLSAAEEIRDPLMQWL GKHVDSEGEIKSGQLSLRFVSSYVSEVEITPSCIPVVTNMEAFSSEHFNLEI YRQNLQTKQLGKVILFAEVTPTTMRLLDGLMFQTPQEMGLIVIAARQTEGK GRGGNVWLSPVGCALSTLLISIPLRSQLGQRIPFVQHLMSVAVVEAVRSIP EYQDINLRVKWPNDIYYSDLMKIGGVLVNSTLMGETFYILIGCGFNVTNSNP TICINDLITEYNKQHKAELKPLRADYLIARVVTVLEKLIKEFQDKGPNSVLPL YYRYWVHSGQQVHLGSAEGPKVSIVGLDDSGFLQVHQEGGEVVTVHPD GNSFDMLRNLILPKRR [SEQ ID NO:205]	Deficiência de sintetase de holo-carboxilase	Acidemia orgânica
PC	5091	0173599	P11498 A0A024R5C5	MLKFRTVHGGLRLLGIRRTSTAPAASPNVRRLEYKPIKKVMVANRGEIAIRV FRACTELGIRTVAIYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIK VAKENNVDAVHPGYGFLSERADFAQACQDAGVRFIGPSPEVVRKMGDKV EARAIAIAAGVPVVPGTDAPITSLHEAHEFSNTYGFPIIFKAAYGGGGRGMR VVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNI LHLYERDCSIQRRHQKVVEIAPAAHLDPQLRTRLTSDSVKLAKQVGYENA GTVEFLVDRHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIHVAEGRSLP DLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDN ASAFQGAVISPHYDSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIA FLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQKLLHYLGHVMVN GPTTPIPVKASPSPTDPVVPAVPIGPPPAGFRDILLREGPEGFARAVRNHP GLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGAT FDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFK FCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAGGVVEAAISYTGDV	Deficiência de Piruvato Carboxi-lase	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ADPSRTKYSLQYYMGLAEELVRAGTHILCIKDMAGLLKPTACTMLVSSLRD RFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVAADSMSGMTSQPSM GALVACTRGTPLDTEVPMERVFDYSEYWEGARGLYAAFDCTATMKSGNS DVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTP SSKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGVPHGGF PEPFRSKVLKDLPRVEGRPGASLPPLDLQALEKELVDRHGEEVTPEDVLS AAMYPDVFAHFKDFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKTLHIK ALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKALKDVKG QIGAPMPGKVIDIKVVAGAKVAKGQPLCVLSAMKMETVVTSPMEGTVRKV HVTKDMTLEGDDLILEIE [SEQ ID NO:206]		
SLC7A 7	9056	0155465	Q9UM01 A0A0S2Z502	MVDSTEYEVASQPEVETSPLGDGASPGPEQVKLKKEISLLNGVCLIVGNMI GSGIFVSPKGVLIYSASFGLSLVIWAVGGLFSVFGALCYAELGTTIKKSGAS YAYILEAFGGFLAFIRLWTSLLIIEPTSQAIIAITFANYMVQPLFPSCFAPYAA SRLLAAACICLLTFINCAYVKWGTLVQDIFTYAKVLALIAVIVAGIVRLGQGA STHFENSFEGSSFAVGDIALALYSALFSYSGWDTLNYVTEEIKNPERNLPL SIGISMPIVTIIYILTNVAYYTVLDMRDILASDAVAVTFADQIFGIFNWIIPLSVA LSCFGGLNASIVAASRLFFVGSREGHLPDAICMIHVERFTPVPSLLFNGIMA LIYLCVEDIFQLINYYSFSYWFFVGLSIVGQLYLRWKEPDRPRPLKLSVFFPI VFCLCTIFLVAVPLYSDTINSLIGIAIALSGLPFYFLIIRVPEHKRPLYLRRIVGS ATRYLQVLCMSVAAEMDLEDGGEMPKQRDPKSN [SEQ ID NO:207]	Intolerância à proteína lisinúrica	Desordem do ciclo da ureia
CPT2	1376	0157184	P23786 A0A140VK13 A0A1B0GTB8	MVPRLLLRAWPRGPAVGPGAPSRPLSAGSGPGQYLQRSIVPTMHYQDSL PRLPIPKLEDTIRRYLSAQKPLLNDGQFRKTEQFCKSFENGIGKELHEQLVA LDKQNKHTSYISGPWFDMYLSARDSVVLNFNPFMAFNPDPKSEYNDQLT RATNMTVSAIRFLKTLRAGLLEPEVFHLNPAKSDTITFKRLIRFVPSSLSWY GAYLVNAYPLDMSQYFRLFNSTRLPKPSRDELFTDDKARHLLVLRKGNFYI FDVLDQDGNIVSPSEIQAHLKYILSDSSPAPEFPLAYLTSENRDIWAELRQK LMSSGNEESLRKVDSAVFCLCLDDFPIKDLVHLSHNMLHGDGTNRWFDKS FNLIIAKDGSTAVHFEHSWGDGVAVLRFFNEVFKDSTQTPAVTPQSQPATT DSTVTVQKLNFELTDALKTGITAAKEKFDATMKTLTIDCVQFQRGGKEFLK KQKLSPDAVAQLAFQMAFLRQYGQTVATYESCSTAAFKHGRTETIRPASV YTKRCSEAFVREPSRHSAGELQQMMVECSKYHGQLTKEAAMGQGFDRH	Deficiência de carnitina palmitoil-transferase tipo II (CPT II)	Oxidação de ácidos graxos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LFALRHLAAAKGIILPELYLDPAYGQINHNVLSTSTLSSPAVNLGGFAPVVS DGFGVGYAVHDNWIGCNVSSYPGRNAREFLQCVEKALEDMFDALEGKSI KS [SEQ ID NO:208]		
ACAD M	34	0117054	P11310 A0A0S2Z366, B7Z911, Q5HYG7, Q5T4U5, B4DJE7	MAAGFGRCCRVLRSISRFHWRSQHTKANRQREPGLGFSFEFTEQQKEFQ ATARKFAREEIIPVAAEYDKTGEYPVPLIRRAWELGLMNTHIPENCGGLGL GTFDACLISEELAYGCTGVQTAIEGNSLGQMPIIIAGNDQQKKKYLGRMTE EPLMCAYCVTEPGAGSDVAGIKTKAEKKGDEYIINGQKMWITNGGKANWY FLLARSDPDPKAPANKAFTGFIVEADTPGIQIGRKELNMGQRCSDTRGIVF EDVKVPKENVLIGDGAGFKVAMGAFDKTRPVVAAGAVGLAQRALDEATKY ALERKTFGKLLVEHQAISFMLAEMAMKVELARMSYQRAAWEVDSGRRNT YYASIAKAFAGDIANQLATDAVQILGGNGFNTEYPVEKLMRDAKIYQIYEGT SQIQRLIVAREHIDKYKN [SEQID NO:209]	Deficiência de Acil-CoA Desidrogenase de Cadeia Média (MCAD)	Oxidação de ácidos graxos
ACAD S	35	0122971	P16219 E5KSD5, B4DUH1, E9PE82	MAAALLARASGPARRALCPRAWRQLHTIYQSVELPETHQMLLQTCRDFAE KELFPIAAQVDKEHLFPAAQVKKMGGLGLLAMDVPEELGGAGLDYLAYAIA MEEISRGCASTGVIMSVNNSLYLGPILKFGSKEQKQAWVTPFTSGDKIGCF ALSEPGNGSDAGAASTTARAEGDSWVLNGTKAWITNAWEASAAVVFAST DRALQNKGISAFLVPMPTPGLTLGKKEDKLGIRGSSTANLIFEDCRIPKDSIL GEPGMGFKIAMQTLDMGRIGIASQALGIAQTALDCAVNYAENRMAFGAPL TKLQVIQFKLADMALALESARLLTWRAAMLKDNKKPFIKEAAMAKLAASEA ATAISHQAIQILGGMGYVTEMPAERHYRDARITEIYEGTSEIQRLVIAGHLLR SYRS [SEQ ID NO:210]	Deficiência de Acil-CoA de Cadeia Curta (SCAD) Desidrogenase	Oxidação de ácidos graxos
ACA-DVL	37	0072778	P49748 B3KPA6	MQAARMAASLGRQLLRLGGGSSRLTALLGQPRPGPARRPYAGGAAQLAL DKSDSHPSDALTRKKPAKAESKSFAVGMFKGQLTTDQVFPYPSVLNEEQT QFLKELVEPVSRFFEEVNDPAKNDALEMVEETTWQGLKELGAFGLQVPSE LGGVGLCNTQYARLVEIVGMHDLGVGITLGAHQSIGFKGILLFGTKAQKEK YLPKLASGETVAAFCLTEPSSGSDAASIRTSAVPSPCGKYYTLNGSKLWIS NGGLADIFTVFAKTPVTDPATGAVKEKITAFVVERGFGGITHGPPEKKMGIK ASNTAEVFFDGVRVPSENVLGEVGSGFKVAMHILNNGRFGMAAALAGTM RGIIAKAVDHATNRTQFGEKIHNFGLIQEKLARMVMLQYVTESMAYMVSAN MDQGATDFQIEAAISKIFGSEAAWKVTDECIQIMGGMGFMKEPGVERVLR	Deficiência de Acil-CoA Desidrogenase de Cadeia Muito Longa (VLCAD)	Oxidação de ácidos graxos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DLRIFRIFEGTNDILRLFVALQGCMDKGKELSGLGSALKNPFGNAGLLLGE AGKQLRRRAGLGSGLSLSGLVHPELSRSGELAVRALEQFATVVEAKLIKHK KGIVNEQFLLQRLADGAIDLYAMVVVLSRASRSLSEGHPTAQHEKMLCDT WCIEAAARIREGMAALQSDPWQQELYRNFKSISKALVERGGVVTSNPLGF[SEQID NO:211]		
AGL	178	0162688	P35573 A0A0S2A4E4	MGHSKQIRILLLNEMEKLEKTLFRLEQGYELQFRLGPTLQGKAVTVYTNYP FPGETFNREKFRSLDWENPTEREDDSDKYCKLNLQQSGSFQYYFLQGNE KSGGGYIVVDPILRVGADNHVLPLDCVTLQTFLAKCLGPFDEWESRLRVAK ESGYNMIHFTPLQTLGLSRSCYSLANQLELNPDFSRPNRKYTWNDVGQLV EKLKKEWNVICITDVVYNHTAANSKWIQEHPECAYNLVNSPHLKPAWVLD RALWRFSCDVAEGKYKEKGIPALIENDHHMNSIRKIIWEDIFPKLKLWEFFQ VDVNKAVEQFRRLLTQENRRVTKSDPNQHLTIIQDPEYRRFGCTVDMNIAL TTFIPHDKGPAAIEECCNWFHKRMEELNSEKHRLINYHQEQAVNCLLGNV FYERLAGHGPKLGPVTRKHPLVTRYFTFPFEEIDFSMEESMIHLPNKACFL MAHNGWVMGDDPLRNFAEPGSEVYLRRELICWGDSVKLRYGNKPEDCP YLWAHMKKYTEITATYFQGVRLDNCHSTPLHVAEYMLDAARNLQPNLYVV AELFTGSEDLDNVFVTRLGISSLIREAMSAYNSHEEGRLVYRYGGEPVGSF VQPCLRPLMPAIAHALFMDITHDNECPIVHRSAYDALPSTTIVSMACCASG STRGYDELVPHQISVVSEERFYTKWNPEALPSNTGEVNFQSGIIAARCAIS KLHQELGAKGFIQVYVDQVDEDIVAVTRHSPSIHQSVVAVSRTAFRNPKTS FYSKEVPQMCIPGKIEEVVLEARTIERNTKPYRKDENSINGTPDITVEIREHI QLNESKIVKQAGVATKGPNEYIQEIEFENLSPGSVIIFRVSLDPHAQVAVGIL RNHLTQFSPHFKSGSLAVDNADPILKIPFASLASRLTLAELNQILYRCESEE KEDGGGCYDIPNWSALKYAGLQGLMSVLAEIRPKNDLGHPFCNNLRSGD WMIDYVSNRLISRSGTIAEVGKWLQAMFFYLKQIPRYLIPCYFDAILIGAYTT LLDTAWKQMSSFVQNGSTFVKHLSLGSVQLCGVGKFPSLPILSPALMDVP YRLNEITKEKEQCCVSLAAGLPHFSSGIFRCWGRDTFIALRGILLITGRYVE ARNIILAFAGTLRHGLIPNLLGEGIYARYNCRDAVWWWLQCIQDYCKMVPN GLDILKCPVSRMYPTDDSAPLPAGTLDQPLFEVIQEAMQKHMQGIQFRER NAGPQIDRNMKDEGFNITAGVDEETGFVYGGNRFNCGTWMDKMGESDR ARNRGIPATPRDGSAVEIVGLSKSAVRWLLELSKKNIFPYHEVTVKRHGKAI KVSYDEWNRKIQDNFEKLFHVSEDPSDLNEKHPNLVHKRGIYKDSYGASS		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				PWCDYQLRPNFTIAMVVAPELFTTEKAWKALEIAEKKLLGPLGMKTLDPDD MVYCGIYDNALDNDNYNLAKGFNYHQGPEWLWPIGYFLRAKLYFSRLMG PETTAKTIVLVKNVLSRHYVHLERSPWKGLPELTNENAQYCPFSCETQAW SIATILETLYDL [SEQ ID NO:212]		
G6PC	2538	0131482	P35575	MEEGMNVLHDFGIQSTHYLQVNYQDSQDWFILVSVIADLRNAFYVLFPIWF HLQEAVGIKLLWVAVIGDWLNLVFKWILFGQRPYWWVLDTDYYSNTSVPLI KQFPVTCETGPGSPSGHAMGTAGVYYVMVTSTLSIFQGKIKPTYRFRCLN VILWLGFWAVQLNVCLSRIYLAAHFPHQVVAGVLSGIAVAETFSHIHSIYNA SLKKYFLITFFLFSFAIGFYLLLKGLGVDLLWTLEKAQRWCEQPEWVHIDTT PFASLLKNLGTLFGLGLALNSSMYRESCKGKLSKWLPFRLSSIVASLVLLH VFDSLKPPSQVELVFYVLSFCKSAVVPLASVSVIPYCLAQVLGQPHKKSL[S EQID NO:213]	GSDIa (doença de Von Gierke)	Distúrbio de armazenamento de glicogênio hepático
GBE1	2632	0114480	Q04446 Q59ET0	MAAPMTPAARPEDYEAALNAALADVPELARLLEIDPYLKPYAVDFQRRYK QFSQILKNIGENEGGIDKFSRGYESFGVHRCADGGLYCKEWAPGAEGVFL TGDFNGWNPFSYPYKKLDYGKWELYIPPKQNKSVLVPHGSKLKVVITSKS GEILYRISPWAKYVVREGDNVNYDWIHWDPEHSYEFKHSRPKKPRSLRIY ESHVGISSHEGKVASYKHFTCNVLPRIKGLGYNCIQLMAIMEHAYYASFGY QITSFFAASSRYGTPEELQELVDTAHSMGIIVLLDVVHSHASKNSADGLNM FDGTDSCYFHSGPRGTHDLWDSRLFAYSSWEILRFLLSNIRWWLEEYRFD GFRFDGVTSMLYHHHGVGQGFSGDYSEYFGLQVDEDALTYLMLANHLVH TLCPDSITIAEDVSGMPALCSPISQGGGGFDYRLAMAIPDKWIQLLKEFKDE DWNMGDIVYTLTNRRYLEKCIAYAESHDQALVGDKSLAFWLMDAEMYTN MSVLTPFTPVIDRGIQLHKMIRLITHGLGGEGYLNFMGNEFGHPEWLDFPR KGNNESYHYARRQFHLTDDDLLRYKFLNNFDRDMNRLEERYGWLAAPQA YVSEKHEGNKIIAFERAGLLFIFNFHPSKSYTDYRVGTALPGKFKIVLDSDA AEYGGHQRLDHSTDFFSEAFEHNGRPYSLLVYIPSRVALILQNVDLPN [SEQ ID NO:214]	GSD IV (Doença de Andersen, Enzima Brancher)	Distúrbio de armazenamento de glicogênio hepático
PHKA 1	5255	0067177	P46020	MRSRSNSGVRLDGYARLVQQTILCHQNPVTGLLPASYDQKDAWVRDNVY SILAVWGLGLAYRKNADRDEDKAKAYELEQSVVKLMRGLLHCMIRQVDKV ESFKYSQSTKDSLHAKYNTKTCATVVGDDQWGHLQLDATSVYLLFLAQMT	GSD IXa	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ASGLHIIHSLDEVNFIQNLVFYIEAAYKTADFGIWERGDKTNQGISELNASSV GMAKAALEALDELDLFGVKGGPQSVIHVLADEVQHCQSILNSLLPRASTSK EVDASLLSVVSFPAFAVEDSQLVELTKQEIITKLQGRYGCCRFLRDGYKTP KEDPNRLYYEPAELKLFENIECEWPLFWTYFILDGVFSGNAEQVQEYKEAL EAVLIKGKNGVPLLPELYSVPPDRVDEEYQNPHTVDRVPMGKLPHMWGQ SLYILGSLMAEGFLAPGEIDPLNRRFSTVPKPDVVVQVSILAETEEIKTILKD KGIYVETIAEVYPIRVQPARILSHIYSSLGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVEMLRTDLSYLCSRWRMTGQP TITFPISHSMLDEDGTSLNSSILAALRKMQDGYFGGARVQTGKLSEFLTTS CCTHLSFMDPGPEGKLYSEDYDDNYDYLESGNWMNDYDSTSHARCGDE VARYLDHLLAHTAPHPKLAPTSQKGGLDRFQAAVQTTCDLMSLVTKAKEL HVQNVHMYLPTKLFQASRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQSG EVDFKALVLQLKETSSLQEQADILYMLYTMKGPDWNTELYNERSATVRELL TELYGKVGEIRHWGLIRYISGILRKKVEALDEACTDLLSHQKHLTVGLPPEP REKTISAPLPYEALTQLIDEASEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGLMNLSPSAMKNLLHHILSGKE FGVERSVRPTDSNVSPAISIHEIGAVGATKTERTGIMQLKSEIKQVEFRRLSI SAESQSPGTSMTPSSGSFPSAYDQQSSKDSRQGQWQRRRRLDGALNRV PVGFYQKVWKVLQKCHGLSVEGFVLPSSTTREMTPGEIKFSVHVESVLNR VPQPEYRQLLVEAILVLTMLADIEIHSIGSIIAVEKIVHIANDLFLQEQKTLGAD DTMLAKDPASGICTLLYDSAPSGRFGTMTYLSKAAATYVQEFLPHSICAMQ [SEQ ID NO:215]		
PHKA2	0044446 5256	5256 0044446	P46019	MRSRSNSGVRLDGYARLVQQTILCYQNPVTGLLSASHEQKDAWVRDNIY SILAVWGLGMAYRKNADRDEDKAKAYELEQNVVKLMRGLLQCMMRQVAK VEKFKHTQSTKDSLHAKYNTATCGTVVGDDQWGHLQVDATSLFLLFLAQ MTASGLRIIFTLDEVAFIQNLVFYIEAAYKVADYGMWERGDKTNQGIPELNA SSVGMAKAALEAIDELDLFGAHGGRKSVIHVLPDEVEHCQSILFSMLPRAS TSKEIDAGLLSIISFPAFAVEDVNLVNVTKNEIISKLQGRYGCCRFLRDGYKT PREDPNRLHYDPAELKLFENIECEWPVFWTYFIIDGVFSGDAVQVQEYRE ALEGILIRGKNGIRLVPELYAVPPNKVDEEYKNPHTVDRVPMGKVPHLWG QSLYILSSLLAEGFLAAGEIDPLNRRFSTSVKPDVVVQVTVLAENNHIKDLL RKHGVNVQSIADIHPIQVQPGRILSHIYAKLGRNKNMNLSGRPYRHIGVLGT	GSD IXa	Distúrbio de armazenamento de glicogênio hepático

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SKLYVIRNQIFTFTPQFTDQHHFYLALDNEMIVEMLRIELAYLCTCWRMTGR PTLTFPISRTMLTNDGSDIHSAVLSTIRKLEDGYFGGARVKLGNLSEFLTTS FYTYLTFLDPDCDEKLFDNASEGTFSPDSDSDLVGYLEDTCNQESQDELD HYINHLLQSTSLRSYLPPLCKNTEDRHVFSAIHSTRDILSVMAKAKGLEVPF VPMTLPTKVLSAHRKSLNLVDSPQPLLEKVPESDFQWPRDDHGDVDCEK LVEQLKDCSNLQDQADILYILYVIKGPSWDTNLSGQHGVTVQNLLGELYGK AGLNQEWGLIRYISGLLRKKVEVLAEACTDLLSHQKQLTVGLPPEPREKIIS APLPPEELTKLIYEASGQDISIAVLTQEIVVYLAMYVRAQPSLFVEMLRLRIG LIIQVMATELARSLNCSGEEASESLMNLSPFDMKNLLHHILSGKEFGVERS VRPIHSSTSSPTISIHEVGHTGVTKTERSGINRLRSEMKQMTRRFSADEQF FSVGQAASSSAHSSKSARSSTPSSPTGTSSSDSGGHHIGWGERQGQWL RRRRLDGAINRVPVGFYQRVWKILQKCHGLSIDGYVLPSSTTREMTPHEIK FAVHVESVLNRVPQPEYRQLLVEAIMVLTLLSDTEMTSIGGIIHVDQIVQMA SQLFLQDQVSIGAMDTLEKDQATGICHFFYDSAPSGAYGTMTYLTRAVAS YLQELLPNSGCQMQ [SEQ ID NO:216]		
PHKB	5257	0102893	Q93100	MAGAAGLTAEVSWKVLERRARTKRSGSVYEPLKSINLPRPDNETLWDKLD HYYRIVKSTLLLYQSPTTGLFPTKTCGGDQKAKIQDSLYCAAGAWALALAY RRIDDDKGRTHELEHSAIKCMRGILYCYMRQADKVQQFKQDPRPTTCLHS VFNVHTGDELLSYEEYGHLQINAVSLYLLYLVEMISSGLQIIYNTDEVSFIQN LVFCVERVYRVPDFGVWERGSKYNNGSTELHSSSVGLAKAALEAINGFNL FGNQGCSWSVIFVDLDAHNRNRQTLCSLLPRESRSHNTDAALLPCISYPA FALDDEVLFSQTLDKVVRKLKGKYGFKRFLRDGYRTSLEDPNRCYYKPAEI KLFDGIECEFPIFFLYMMIDGVFRGNPKQVQEYQDLLTPVLHHTTEGYPVV PKYYYVPADFVEYEKNNPGSQKRFPSNCGRDGKLFLWGQALYIIAKLLAD ELISPKDIDPVQRYVPLKDQRNVSMRFSNQGPLENDLVVHVALIAESQRLQ VFLNTYGIQTQTPQQVEPIQIWPQQELVKAYLQLGINEKLGLSGRPDRPIG CLGTSKIYRILGKTVVCYPIIFDLSDFYMSQDVFLLIDDIKNALQFIKQYWKM HGRPLFLVLIREDNIRGSRFNPILDMLAALKKGIIGGVKVHVDRLQTLISGAV VEQLDFLRISDTEELPEFKSFEELEPPKHSKVKRQSSTPSAPELGQQPDVN ISEWKDKPTHEILQKLNDCSCLASQAILLGILLKREGPNFITKEGTVSDHIER VYRRAGSQKLWLAVRYGAAFTQKFSSSIAPHITTFLVHGKQVTLGAFGHE EEVISNPLSPRVIQNIIYYKCNTHDEREAVIQQELVIHIGWIISNNPELFSGML	GSD IXb	Distúrbio de armazenamento de glicogênio hepático

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KIRIGWIIHAMEYELQIRGGDKPALDLYQLSPSEVKQLLLDILQPQQNGRCW LNRRQIDGSLNRTPTGFYDRVWQILERTPNGIIVAGKHLPQQPTLSDMTMY EMNFSLLVEDTLGNIDQPQYRQIVVELLMVVSIVLERNPELEFQDKVDLDR LVKEAFNEFQKDQSRLKEIEKQDDMTSFYNTPPLGKRGTCSYLTKAVMNL LLEGEVKPNNDDPCLIS [SEQ ID NO:217]		
PHKG 2	5261	0156873	P15735	MTLDVGPEDELPDWAAAKEFYQKYDPKDVIGRGVSSVVRRCVHRATGHE FAVKIMEVTAERLSPEQLEEVREATRRETHILRQVAGHPHIITLIDSYESSSF MFLVFDLMRKGELFDYLTEKVALSEKETRSIMRSLLEAVSFLHANNIVHRDL KPENILLDDNMQIRLSDFGFSCHLEPGEKLRELCGTPGYLAPEILKCSMDE THPGYGKEVDLWACGVILFTLLAGSPPFWHRRQILMLRMIMEGQYQFSSP EWDDRSSTVKDLISRLLQVDPEARLTAEQALQHPFFERCEGSQPWNLTPR QRFRVAVWTVLAAGRVALSTHRVRPLTKNALLRDPYALRSVRHLIDNCAF RLYGHWVKKGEQQNRAALFQHRPPGPFPIMGPEEEGDSAAITEDEAVLVL G [SEQ ID NO:218]	GSD IXc	Distúrbio de armazenamento de glicogênio hepático
SLC37 A4	2542	0281500 0137700	O43826 A0A024R3H9, A8K0S7, A0A024R3L1, B4DUH2	MAAQGYGYYRTVIFSAMFGGYSLYYFNRKTFSFVMPSLVEEIPLDKDDLG FITSSQSAAYAISKFVSGVLSDQMSARWLFSSGLLLVGLVNIFFAWSSTVP VFAALWFLNGLAQGLGWPPCGKVLRKWFEPSQFGTWWAILSTSMNLAG GLGPILATILAQSYSWRSTLALSGALCVVVSFLCLLLIHNEPADVGLRNLDP MPSEGKKGSLKEESTLQELLLSPYLWVLSTGYLVVFGVKTCCTDWGQFFL IQEKGQSALVGSSYMSALEVGGLVGSIAAGYLSDRAMAKAGLSNYGNPRH GLLLFMMAGMTVSMYLFRVTVTSDSPKLWILVLGAVFGFSSYGPIALFGVI ANESAPPNLCGTSHAIVGLMANVGGFLAGLPFSTIAKHYSWSTAFWVAEVI CAASTAAFFLLRNIRTKMGRVSKKAE [SEQ ID NO:219]	GSDIb. c, d	Distúrbio de armazenamento de glicogênio hepático
PMM2	5373	0140650	O15305, A0A0S2Z4J6, Q59F02	MAAPGPALCLFDVDGTLTAPRQKITKEMDDFLQKLRQKIKIGVVGGSDFEK VQEQLGNDVVEKYDYVFPENGLVAYKDGKLLCRQNIQSHLGEALIQDLINY CLSYIAKIKLPKKRGTFIEFRNGMLNVSPIGRSCSQEERIEFYELDKKENIRQ KFVADLRKEFAGKGLTFSIGGQISFDVFPDGWDKRYCLRHVENDGYKTIYF FGDKTMPGGNDHEIFTDPRTMGYSVTAPEDTRRICELLFS [SEQ ID NO:220]	PMM2-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
CBS	102724560, 875	0160200	P35520, P0DN79, Q9NTF0, B7Z2D6	MPSETPQAEVGPTGCPHRSGPHSAKGSLEKGSPEDKEAKEPLWIRPDAP SRCTWQLGRPASESPHHHTAPAKSPKILPDILKKIGDTPMVRINKIGKKFGL KCELLAKCEFFNAGGSVKDRISLRMIEDAERDGTLKPGDTIIEPTSGNTGIG LALAAAVRGYRCIIVMPEKMSSEKVDVLRALGAEIVRTPTNARFDSPESHV GVAWRLKNEIPNSHILDQYRNASNPLAHYDTTADEILQQCDGKLDMLVAS VGTGGTITGIARKLKEKCPGCRIIGVDPEGSILAEPEELNQTEQTTYEVEGI GYDFIPTVLDRTVVDKWFKSNDEEAFTFARMLIAQEGLLCGGSAGSTVAV AVKAAQELQEGQRCVVILPDSVRNYMTKFLSDRWMLQKGFLKEEDLTEKK PWWWHLRVQELGLSAPLTVLPTITCGHTIEILREKGFDQAPVVDEAGVILG MVTLGNMLSSLLAGKVQPSDQVGKVIYKQFKQIRLTDTLGRLSHILEMDHF ALVVHEQIQYHSTGKSSQRQMVFGVVTAIDLLNFVAAQERDQK [SEQ ID NO:221]	Deficiência de Cistationina Beta-Sintase (Homocistinúria Clássica); Homocistinúria	Aminoacidopatia
FAH	2184	0103876	P16930	MSFIPVAEDSDFPIHNLPYGVFSTRGDPRPRIGVAIGDQILDLSIIKHLFTGP VLSKHQDVFNQPTLNSFMGLGQAAWKEARVFLQNLLSVSQARLRDDTEL RKCAFISQASATMHLPATIGDYTDFYSSRQHATNVGIMFRDKENALMPNW LHLPVGYHGRASSVVVSGTPIRRPMGQMKPDDSKPPVYGACKLLDMELE MAFFVGPGNRLGEPIPISKAHEHIFGMVLMNDWSARDIQKWEYVPLGPFL GKSFGTTVSPWVVPMDALMPFAVPNPKQDPRPLPYLCHDEPYTFDINLSV NLKGEGMSQAATICKSNFKYMYWTMLQQLTHHSVNGCNLRPGDLLASGT ISGPEPENFGSMLELSWKGTKPIDLGNGQTRKFLLDGDEVIITGYCQGDGY RIGFGQCAGKVLPALLPS [SEQ ID NO:222]	Tirosinemia Tipo I	Aminoacidopatia
TAT	6898	0198650	P17735, A0A140VKB7	MDPYMIQMSSKGNLPSILDVHVNVGGRSSVPGKMKGRKARWSVRPSDM AKKTFNPIRAIVDNMKVKPNPNKTMISLSIGDPTVFGNLPTDPEVTQAMKD ALDSGKYNGYAPSIGFLSSREEIASYYHCPEAPLEAKDVILTSGCSQAIDLC LAVLANPGQNILVPRPGFSLYKTLAESMGIEVKLYNLLPEKSWEIDLKQLEY LIDEKTACLIVNNPSNPCGSVFSKRHLQKILAVAARQCVPILADEIYGDMVF SDCKYEPLATLSTDVPILSCGGLAKRWLVPGWRLGWILIHDRRDIFGNEIR DGLVKLSQRILGPCTIVQGALKSILCRTPGEFYHNTLSFLKSNADLCYGALA AIPGLRPVRPSGAMYLMVGIEMEHFPEFENDVEFTERLVAEQSVHCLPAT CFEYPNFIRVVITVPEVMMLEACSRIQEFCEQHYHCAEGSQEECDK [SEQ ID NO:223]	Tirosinemia Tipo II Tirosinemia Tipo III	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
GALT	2592	0213930	P07902, A0A0S2Z3Y7, B2RAT6	MSRSGTDPQQRQQASEADAAAATFRANDHQHIRYNPLQDEWVLVSAHR MKRPWQGQVEPQLLKTVPRHDPLNPLCPGAIRANGEVNPQYDSTFLFDN DFPALQPDAPSPGPSDHPLFQAKSARGVCKVMCFHPWSDVTLPLMSVPE IRAVVDAWASVTEELGAQYPWVQIFENKGAMMGCSNPHPHCQVWASSFL PDIAQREERSQQAYKSQHGEPLLMEYSRQELLRKERLVLTSEHWLVLVPF WATWPYQTLLLPRRHVRRLPELTPAERDDLASIMKKLLTKYDNLFETSFPY SMGWHGAPTGSEAGANWNHWQLHAHYYPPLLRSATVRKFMVGYEMLA QAQRDLTPEQAAERLRALPEVHYHLGQKDRETATIA [SEQ ID NO:224]	Galactosemia devido à deficiência de galactose-1-fosfato uridililtransferase (GALT)	Distúrbio de carbo-hidratos
GALK 1	2584	0108479	P51570	MAALRQPQVAELLAEARRAFREEFGAEPELAVSAPGRVNLIGEHTDYNQG LVLPMALELMTVLVGSPRKDGLVSLLTTSEGADEPQRLQFPLPTAQRSLEP GTPRWANYVKGVIQYYPAAPLPGFSAVVVSSVPLGGGLSSSASLEVATYT FLQQLCPDSGTIAARAQVCQQAEHSFAGMPCGIMDQFISLMGQKGHALLI DCRSLETSLVPLSDPKLAVLITNSNVRHSLASSEYPVRRRQCEEVARALGK ESLREVQLEELEAARDLVSKEGFRRARHVVGEIRRTAQAAAALRRGDYRA FGRLMVESHRSLRDDYEVSCPELDQLVEAALAVPGVYGSRMTGGGFGG CTVTLLEASAAPHAMRHIQEHYGGTATFYLSQAADGAKVLCL [SEQ ID NO:225]	Galactosemia	Distúrbio de carbo-hidratos
GALE	2582	0117308	Q14376	MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLPESLRR VQELTGRSVEFEEMDILDQGALQRLFKKYSFMAVIHFAGLKAVGESVQKPL DYYRVNLTGTIQLLEIMKAHGVKNLVFSSSATVYGNPQYLPLDEAHPTGGC TNPYGKSKFFIEEMIRDLCQADKTWNAVLLRYFNPTGAHASGCIGEDPQGI PNNLMPYVSQVAIGRREALNVFGNDYDTEDGTGVRDYIHVVDLAKGHIAA LRKLKEQCGCRIYNLGTGTGYSVLQMVQAMEKASGKKIPYKVVARREGDV AACYANPSLAQEELGWTAALGLDRMCEDLWRWQKQNPSGFGTQA [SEQ ID NO:226]	Galactosemia	Distúrbio de carbo-hidratos
G6PD	2539	0160211	P11413	MAEQVALSRTQVCGILREELFQGDAFHQSDTHIFIIMGASGDLAKKKIYPTI WWLFRDGLLPENTFIVGYARSRLTVADIRKQSEPFFKATPEEKLKLEDFFA RNSYVAGQYDDAASYQRLNSHMNALHLGSQANRLFYLALPPTVYEAVTK NIHESCMSQIGWNRIIVEKPFGRDLQSSDRLSNHISSLFREDQIYRIDHYLG KEMVQNLMVLRFANRIFGPIWNRDNIACVILTFKEPFGTEGRGGYFDEFGII	Deficiência de Glicose-6-fosfato Desidrogenase (G6PD)	Distúrbio de carbo-hidratos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				RDVMQNHLLQMLCLVAMEKPASTNSDDVRDEKVKVLKCISEVQANNVVL GQYVGNPDGEGEATKGYLDDPTVPRGSTTATFAAVVLYVENERWDGVPF ILRCGKALNERKAEVRLQFHDVAGDIFHQQCKRNELVIRVQPNEAVYTKM MTKKPGMFFNPEESELDLTYGNRYKNVKLPDAYERLILDVFCGSQMHFVR SDELREAWRIFTPLLHQIELEKPKPIPYIYGSRGPTEADELMKRVGFQYEGT YKWVNPHKL [SEQ ID NO:227]		
SLC3A1	6519	0138079	Q07837, A0A0S2Z4E1, B8ZZK1	MAEDKSKRDSIEMSMKGCQTNNGFVHNEDILEQTPDPGSSTDNLKHSTR GILGSQEPDFKGVQPYAGMPKEVLFQFSGQARYRIPREILFWLTVASVLVL IAATIAIIALSPKCLDWWQEGPMYQIYPRSFKDSNKDGNGDLKGIQDKLDYI TALNIKTVWITSFYKSSLKDFRYGVEDFREVDPIFGTMEDFENLVAAIHDKG LKLIIDFIPNHTSDKHIWFQLSRTRTGKYTDYYIWHDCTHENGKTIPPNNWL SVYGNSSWHFDEVRNQCYFHQFMKEQPDLNFRNPDVQEEIKEILRFWLT KGVDGFSLDAVKFLLEAKHLRDEIQVNKTQIPDTVTQYSELYHDFTTTQVG MHDIVRSFRQTMDQYSTEPGRYRFMGTEAYAESIDRTVMYYGLPFIQEAD FPFNNYLSMLDTVSGNSVYEVITSWMENMPEGKWPNWMIGGPDSSRLTS RLGNQYVNVMNMLLFTLPGTPITYYGEEIGMGNIVAANLNESYDINTLRSK SPMQWDNSSNAGFSEASNTWLPTNSDYHTVNVDVQKTQPRSALKLYQD LSLLHANELLLNRGWFCHLRNDSHYVVYTRELDGIDRIFIVVLNFGESTLLN LHNMISGLPAKMRIRLSTNSADKGSKVDTSGIFLDKGEGLIFEHNTKNLLHR QTAFRDRCFVSNRACYSSVLNILYTSC [SEQ ID NO:228]	Cistinúria	Aminoacidopatia
SLC7A9	11136	0021488	P82251	MGDTGLRKRREDEKSIQSQEPKTTSLQKELGLISGISIIVGTIIGSGIFVSPKS VLSNTEAVGPCLIIWAACGVLATLGALCFAELGTMITKSGGEYPYLMEAYG PIPAYLFSWASLIVIKPTSFAIICLSFSEYVCAPFYVGCKPPQIVVKCLAAAAI LFISTVNSLSVRLGSYVQNIFTAAKLVIVAIIIISGLVLLAQGNTKNFDNSFEG AQLSVGAISLAFYNGLWAYDGWNQLNYITEELRNPYRNLPLAIIIGIPLVTAC YILMNVSYFTVMTATELLQSQAVAVTFGDRVLYPASWIVPLFVAFSTIGAAN GTCFTAGRLIYVAGREGHMLKVLSYISVRRLTPAPAIIFYGIIATIYIIPGDINS LVNYFSFAAWLFYGLTILGLIVMRFTRKELERPIKVPVVIPVLMTLISVFLVLA PIISKPTWEYLYCVLFILSGLLFYFLFVHYKFGWAQKISKPITMHLQMLMEVV PPEEDPE [SEQ ID NO:229]	Cistinúria	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
MTHFR	4524	0177000	P42898, Q59GJ6, Q81U67	MVNEARGNSSLNPCLEGSASSGSESSKDSSRCSTPGLDPERHERLREKM RRRLESGDKWFSLEFFPPRTAEGAVNLISRFDRMAAGGPLYIDVTWHPAG DPGSDKETSSMMIASTAVNYCGLETILHMTCCRQRLEEITGHLHKAKQLGL KNIMALRGDPIGDQWEEEEGGFNYAVDLVKHIRSEFGDYFDICVAGYPKG HPEAGSFEADLKHLKEKVSAGADFIITQLFFEADTFFRFVKACTDMGITCPI VPGIFPIQGYHSLRQLVKLSKLEVPQEIKDVIEPIKDNDAAIRNYGIELAVSLC QELLASGLVPGLHFYTLNREMATTEVLKRLGMWTEDPRRPLPWALSAHP KRREEDVRPIFWASRPKSYIYRTQEWDEFPNGRWGNSSSPAFGELKDYY LFYLKSKSPKEELLKMWGEELTSEESVFEVFVLYLSGEPNRNGHKVTCLP WNDEPLAAETSLLKEELLRVNRQGILTINSQPNINGKPSSDPIVGWGPSGG YVFQKAYLEFFTSRETAEALLQVLKKYELRVNYHLVNVKGENITNAPELQP NAVTWGIFPGREIIQPTVVDPVSFMFWKDEAFALWIERWGKLYEEESPSR TIIQYIHDNYFLVNLVDNDFPLDNCLWQVVEDTLELLNRPTQNARETEAP [SEQ ID NO:230]	Homocistinúria	Aminoacidopatia
MTR	4548	0116984	Q99707	MSPALQDLSQPEGLKKTLRDEINAILQKRIMVLDGGMGTMIQREKLNEEHF RGQEFKDHARPLKGNNDILSITQPDVIYQIHKEYLLAGADIIETNTFSSTSIA QADYGLEHLAYRMNMCSAGVARKAAEEVTLQTGIKRFVAGALGPTNKTLS VSPSVERPDYRNITFDELVEAYQEQAKGLLDGGVDILLIETIFDTANAKAAL FALQNLFEEKYAPRPIFISGTIVDKSGRTLSGQTGEGFVISVSHGEPLCIGL NCALGAAEMRPFIEIIGKCTTAYVLCYPNAGLPNTFGDYDETPSMMAKHLK DFAMDGLVNIVGGCCGSTPDHIREIAEAVKNCKPRVPPATAFEGHMLLSG LEPFRIGPYTNFVNIGERCNVAGSRKFAKLIMAGNYEEALCVAKVQVEMGA QVLDVNMDDGMLDGPSAMTRFCNLIASEPDIAKVPLCIDSSNFAVIEAGLK CCQGKCIVNSISLKEGEDDFLEKARKIKKYGAAMVVMAFDEEGQATETDT KIRVCTRAYHLLVKKLGFNPNDIIFDPNILTIGTGMEEHNLYAINFIHATKVIK ETLPGARISGGLSNLSFSFRGMEAIREAMHGVFLYHAIKSGMDMGIVNAG NLPVYDDIHKELLQLCEDLIWNKDPEATEKLLRYAQTQGTGGKKVIQTDEW RNGPVEERLEYALVKGIEKHIIEDTEEARLNQKKYPRPLNIIEGPLMNGMKI VGDLFGAGKMFLPQVIKSARVMKKAVGHLIPFMEKEREETRVLNGTVEEE DPYQGTIVLATVKGDVHDIGKNIVGVVLGCNNFRVIDLGVMTPCDKILKAAL DHKADIIGLSGLITPSLDEMIFVAKEMERLAIRIPLLIGGATTSKTHTAVKIAPR	Homocistinúria	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				YSAPVIHVLDASKSVVVCSQLLDENLKDEYFEEIMEEYEDIRQDHYESLKE RRYLPLSQARKSGFQMDWLSEPHPVKPTFIGTQVFEDYDLQKLVDYIDWK PFFDVWQLRGKYPNRGFPKIFNDKTVGGEARKVYDDAHNMLNTLISQKKL RARGVVGFWPAQSIQDDIHLYAEAAVPQAAEPIATFYGLRQQAEKDSAST EPYYCLSDFIAPLHSGIRDYLGLFAVACFGVEELSKAYEDDGDDYSSIMVK ALGDRLAEAFAEELHERVRRELWAYCGSEQLDVADLRRLRYKGIRPAPGY PSQPDHTEKLTMWRLADIEQSTGIRLTESLAMAPASAVSGLYFSNLKSKYF AVGKISKDQVEDYALRKNISVAEVEKWLGPILGYDTD [SEQ ID NO:231]		
MTRR	4552	0124275	Q9UBK8	MGAASVRAGARLVEVALCSFTVTCLEVMRRFLLLYATQQGQAKAIAEEICE QAVVHGFSADLHCISESDKYDLKTETAPLVVVVSTTGTGDPPDTARKFVKE IQNQTLPVDFFAHLRYGLLGLGDSEYTYFCNGGKIIDKRLQELGARHFYDT GHADDCVGLELVVEPWIAGLWPALRKHFRSSRGQEEISGALPVASPASSR TDLVKSELLHIESQVELLRFDDSGRKDSEVLKQNAVNSNQSNVVIEDFESS LTRSVPPLSQASLNIPGLPPEYLQVHLQESLGQEESQVSVTSADPVFQVPI SKAVQLTTNDAIKTTLLVELDISNTDFSYQPGDAFSVICPNSDSEVQSLLQR LQLEDKREHCVLLKIKADTKKKGATLPQHIPAGCSLQFIFTWCLEIRAIPKKA FLRALVDYTSDSAEKRRLQELCSKQGAADYSRFVRDACACLLDLLLAFPS CQPPLSLLLEHLPKLQPRPYSCASSSLFHPGKLHFVFNIVEFLSTATTEVLR KGVCTGWLALLVASVLQPNIHASHEDSGKALAPKISISPRTTNSFHLPDDP SIPIIMVGPGTGIAPFIGFLQHREKLQEQHPDGNFGAMWLFFGCRHKDRDY LFRKELRHFLKHGILTHLKVSFSRDAPVGEEEAPAKYVQDNIQLHGQQVAR ILLQENGHIYVCGDAKNMAKDVHDALVQIISKEVGVEKLEAMKTLATLKEEK RYLQDIWS [SEQ ID NO:232]	Homocistinúria	Aminoacidopatia
ATP7B	540	0123191	P35670, A0A024RDX3, B7ZLR4, B7ZLR3, E7ET55	MPEQERQITAREGASRKILSKLSLPTRAWEPAMKKSFAFDNVGYEGGLDG LGPSSQVATSTVRILGMTCQSCVKSIEDRISNLKGIISMKVSLEQGSATVKY VPSVVCLQQVCHQIGDMGFEASIAEGKAASWPSRSLPAQEAVVKLRVEG MTCQSCVSSIEGKVRKLQGVVRVKVSLSNQEAVITYQPYLIQPEDLRDHVN DMGFEAAIKSKVAPLSLGPIDIERLQSTNPKRPLSSANQNFNNSETLGHQG SHVVTLQLRIDGMHCKSCVLNIEENIGQLLGVQSIQVSLENKTAQVKYDPS CTSPVALQRAIEALPPGNFKVSLPDGAEGSGTDHRSSSSHSPGSPPRNQV QGTCSTTLIAIAGMTCASCVHSIEGMISQLEGVQQISVSLAEGTATVLYNPS	Doença de Wilson Distúrbio do Metabolismo do Cobre	Distúrbio de Transporte de Metal

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				VISPEELRAAIEDMGFEASVVSESCSTNPLGNHSAGNSMVQTTDGTPTSV QEVAPHTGRLPANHAPDILAKSPQSTRAVAPQKCFLQIKGMTCASCVSNIE RNLQKEAGVLSVLVALMAGKAEIKYDPEVIQPLEIAQFIQDLGFEAAVMEDY AGSDGNIELTITGMTCASCVHNIESKLTRTNGITYASVALATSKALVKFDPEI IGPRDIIKIIEEIGFHASLAQRNPNAHHLDHKMEIKQWKKSFLCSLVFGIPVM ALMIYMLIPSNEPHQSMVLDHNIIPGLSILNLIFFILCTFVQLLGGWYFYVQA YKSLRHRSANMDVLIVLATSIAYVYSLVILVVAVAEKAERSPVTFFDTPPML FVFIALGRWLEHLAKSKTSEALAKLMSLQATEATVVTLGEDNLIIREEQVPM ELVQRGDIVKVVPGGKFPVDGKVLEGNTMADESLITGEAMPVTKKPGSTVI AGSINAHGSVLIKATHVGNDTTLAQIVKLVEEAQMSKAPIQQLADRFSGYF VPFIIIMSTLTLVVWIVIGFIDFGVVQRYFPNPNKHISQTEVIIRFAFQTSITVL CIACPCSLGLATPTAVMVGTGVAAQNGILIKGGKPLEMAHKIKTVMFDKTG TITHGVPRVMRVLLLGDVATLPLRKVLAVVGTAEASSEHPLGVAVTKYCKE ELGTETLGYCTDFQAVPGCGIGCKVSNVEGILAHSERPLSAPASHLNEAG SLPAEKDAVPQTFSVLIGNREWLRRNGLTISSDVSDAMTDHEMKGQTAILV AIDGVLCGMIAIADAVKQEAALAVHTLQSMGVDVVLITGDNRKTARAIATQV GINKVFAEVLPSHKVAKVQELQNKGKKVAMVGDGVNDSPALAQADMGVAI GTGTDVAIEAADVVLIRNDLLDVVASIHLSKRTVRRIRINLVLALIYNLVGIPIA AGVFMPIGIVLQPWMGSAAMAASSVSVVLSSLQLKCYKKPDLERYEAQAH GHMKPLTASQVSVHIGMDDRWRDSPRATPWDQVSYVSQVSLSSLTSDKP SRHSAAADDDGDKWSLLLNGRDEEQYI [SEQ ID NO:233]		
HPRT 1	3251	0165704	P00492, A0A140VJL3	MATRSPGVVISDDEPGYDLDLFCIPNHYAEDLERVFIPHGLIMDRTERLAR DVMKEMGGHHIVALCVLKGGYKFFADLLDYIKALNRNSDRSIPMTVDFIRL KSYCNDQSTGDIKVIGGDDLSTLTGKNVLIVEDIIDTGKTMQTLLSLVRQYN PKMVKVASLLVKRTPRSVGYKPDFVGFEIPDKFVVGYALDYNEYFRDLNH VCVISETGKAKYKA [SEQ ID NO:234]	Síndrome de Lesch-Nyhan Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
HJV	148738	0168509	Q6ZVN8	MGEPGQSPSPRSSHGSPPTLSTLTLLLLLCGHAHSQCKILRCNAEYVSSTL SLRGGGSSGALRGGGGGGRGGGVGSGGLCRALRSYALCTRRTARTCR GDLAFHSAVHGIEDLMIQHNCSRQGPTAPPPPRGPALPGAGSGLPAPDPC DYEGRFSRLHGRPPGFLHCASFGDPHVRSFHHHFHTCRVQGAWPLLDN DFLFVQATSSPMALGANATATRKLTIIFKNMQECIDQKVYQAEVDNLPVAF	Hemocromatose, Tipo 2A	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				EDGSINGGDRPGGSSLSIQTANPGNHVEIQAAYIGTTIIIRQTAGQLSFSIKV AEDVAMAFSAEQDLQLCVGGCPPSQRLSRSERNRRGAITIDTARRLCKEG LPVEDAYFHSCVFDVLISGDPNFTVAAQAALEDARAFLPDLEKLHLFPSDA GVPLSSATLLAPLLSGLFVLWLCIQ [SEQ ID NO:235]		
HAMP	57817	0105697	P81172	MALSSQIWAACLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWM PMFQRRRRRDTHFPICIFCCGCCHRSKCGMCCKT [SEQ ID NO:236]	Hemocromatose Tipo 2B: Hemo-cromatose Primá-ria	
JAG1	182	0101384	P78504, Q99740	MRSPRTRGRSGRPLSLLLALLCALRAKVCGASGQFELEILSMQNVNGELQ NGNCCGGARNPGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGS TPVIGGNTFNLKASRGNDRNRIVLPFSFAWPRSYTLLVEAWDSSNDTVQP DSIIEKASHSGMINPSRQWQTLKQNTGVAHFEYQIRVTCDDYYYGFGCNK FCRPRDDFFGHYACDQNGNKTCMEGWMGPECNRAICRQGCSPKHGSC KLPGDCRCQYGWQGLYCDKCIPHPGCVHGICNEPWQCLCETNWGGQLC DKDLNYCGTHQPCLNGGTCSNTGPDKYQCSCPEGYSGPNCEIAEHACLS DPCHNRGSCKETSLGFECECSPGWTGPTCSTNIDDCSPNNCSHGGTCQ DLVNGFKCVCPPQWTGKTCQLDANECEAKPCVNAKSCKNLIASYYCDCL PGWMGQNCDININDCLGQCQNDASCRDLVNGYRCICPPGYAGDHCERDI DECASNPCLNGGHCQNEINRFQCLCPTGFSGNLCQLDIDYCEPNPCQNG AQCYNRASDYFCKCPEDYEGKNCSHLKDHCRTTPCEVIDSCTVAMASND TPEGVRYISSNVCGPHGKCKSQSGGKFTCDCNKGFTGTYCHENINDCES NPCRNGGTCIDGVNSYKCICSDGWEGAYCETNINDCSQNPCHNGGTCRD LVNDFYCDCKNGWKGKTCHSRDSQCDEATCNNGGTCYDEGDAFKCMCP GGWEGTTCNIARNSSCLPNPCHNGGTCVVNGESFTCVCKEGWEGPICAQ NTNDCSPHPCYNSGTCVDGDNWYRCECAPGFAGPDCRININECQSSPCA FGATCVDEINGYRCVCPPGHSGAKCQEVSGRPCITMGSVIPDGAKWDDD CNTCQCLNGRIACSKVWCGPRPCLLHKGHSECPSGQSCIPILDDQCFVHP CTGVGECRSSSLQPVKTKCTSDSYYQDNCANITFTFNKEMMSPGLTTEHI CSELRNLNILKNVSAEYSIYIACEPSPSANNEIHVAISAEDIRDDGNPIKEITD KIIDLVSKRDGNSSLIAAVAEVRVQRRPLKNRTDFLVPLLSSVLTVAWICCL VTAFYWCLRKRRKPGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVP		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				IKDYENKNSKMSKIRTHNSEVEEDDMDKHQQKARFAKQPAYTLVDREEKP PNGTPTKHPNWTNKQDNRDLESAQSLNRMEYIV [SEQ ID NO:237]		
TTR	7276	0118271	P02766, E9KL36	MASHRLLLLCLAGLVFVSEAGPTGTGESKCPLMVKVLDAVRGSPAINVAV HVFRKAADDTWEPFASGKTSESGELHGLTTEEEFVEGIYKVEIDTKSYWK ALGISPFHEHAEVVFTANDSGPRRYTIAALLSPYSYSTTAVVTNPKE [SEQ ID NO:238]	Amiloidose TTR familiar; Polineuropatia amiloide familiar	
AGXT	189	0172482	P21549	MASHKLLVTPPKALLKPLSIPNQLLLGPGPSNLPPRIMAAGGLQMIGSMSK DMYQIMDEIKEGIQYVFQTRNPLTLVISGSGHCALEAALVNVLEPGDSFLV GANGIWGQRAVDIGERIGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLLF LTHGESSTGVLQPLDGFGELCHRYKCLLLVDSVASLGGTPLYMDRQGIDIL YSGSQKALNAPPGTSLISFSDKAKKKMYSRKTKPFSFYLDIKWLANFWGC DDQPRMYHHTIPVISLYSLRESLALIAEQGLENSWRQHREAAAYLHGRLQA LGLQLFVKDPALRLPTVTTVAVPAGYDWRDIVSYVIDHFDIEIMGGLGPSTG KVLRIGLLGCNATRENVDRVTEALRAALQHCPKKKL [SEQ ID NO:239]	Hiperoxalúria Primária Tipo I	
LIPA	3988	0107798	P38571 A0A0A0MT32	MKMRFLGLVVCLVLWTLHSEGSGGKLTAVDPETNMNVSEIISYWGFPSEE YLVETEDGYILCLNRIPHGRKNHSDKGPKPVVFLQHGLLADSSNWVTNLA NSSLGFILADAGFDVWMGNSRGNTWSRKHKTLSVSQDEFWAFSYDEMA KYDLPASINFILNKTGQEQVYYVGHSQGTTIGFIAFSQIPELAKRIKMFFALG PVASVAFCTSPMAKLGRLPDHLIKDLFGDKEFLPQSAFLKWLGTHVCTHVI LKELCGNLCFLLCGFNERNLNMSRVDVYTTHSPAGTSVQNMLHWSQAVK FQKFQAFDWGSSAKNYFHYNQSYPPTYNVKDMLVPTAVWSGGHDWLAD VYDVNILLTQITNLVFHESIPEWEHLDFIWGLDAPWRLYNKIINLMRKYQ [SEQ ID NO:240]	Deficiência de Lipase de Ácido Lisossomal	Distúrbio de Armazenamento Lisossômico
SERPING1	710	0149131	P05155, A0A0S2Z4J1, B2R659, E7EWE5, B3KSP2, G5E9S2	MASRLTLLTLLLLLLAGDRASSNPNATSSSSQDPESLQDRGEGKVATTVIS KMLFVEPILEVSSLPTTNSTTNSATKITANTTDEPTTQPTTEPTTQPTIQPTQ PTTQLPTDSPTQPTTGSFCPGPVTLCSDLESHSTEAVLGDALVDFSLKLYH AFSAMKKVETNMAFSPFSIASLLTQVLLGAGENTKTNLESILSYPKDFTCVH QALKGFTTKGVTSVSQIFHSPDLAIRDTFVNASRTLYSSSPRVLSNNSDAN LELINTWVAKNTNNKISRLLDSLPSDTRLVLLNAIYLSAKWKTTFDPKKTRM	Angioedma Hereditário	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				EPFHFKNSVIKVPMMNSKKYPVAHFIDQTLKAKVGQLQLSHNLSLVILVPQ NLKHRLEDMEQALSPSVFKAIMEKLEMSKFQPTLLTLPRIKVTTSQDMLSIM EKLEFFDFSYDLNLCGLTEDPDLQVSAMQHQTVLELTETGVEAAAASAISV ARTLLVFEVQQPFLFVLWDQQHKFPVFMGRVYDPRA [SEQ ID NO:241]		
HSD17B4	3295	0133835	P51659	MGSPLRFDGRVVLVTGAGAGLGRAYALAFAERGALVVVNDLGGDFKGVG KGSLAADKVVEEIRRRGGKAVANYDSVEEGEKVVKTALDAFGRIDVVVNN AGILRDRSFARISDEDWDIIHRVHLRGSFQVTRAAWEHMKKQKYGRIIMTS SASGIYGNFGQANYSAAKLGLLGLANSLAIEGRKSNIHCNTIAPNAGSRMT QTVMPEDLVEALKPEYVAPLVLWLCHESCEENGGLFEVGAGWIGKLRWE RTLGAIVRQKNHPMTPEAVKANWKKICDFENASKPQSIQESTGSIIEVLSKI DSEGGVSANHTSRATSTATSGFAGAIGQKLPPFSYAYTELEAIMYALGVGA SIKDPKDLKFIYEGSSDFSCLPTFGVIIGQKSMMGGGLAEIPGLSINFAKVLH GEQYLELYKPLPRAGKLKCEAVVADVLDKGSGVVIIMDVYSYSEKELICHN QFSLFLVGSGGFGGKRTSDKVKVAVAIPNRPPDAVLTDTTSLNQAALYRLS GDWNPLHIDPNFASLAGFDKPILHGLCTFGFSARRVLQQFADNDVSRFKAI KARFAKPVYPGQTLQTEMWKEGNRIHFQTKVQETGDIVISNAYVDLAPTS GTSAKTPSEGGKLQSTFVFEEIGRRLKDIGPEVVKKVNAVFEWHITKGGNI GAKWTIDLKSGSGKVYQGPAKGAADTTIILSDEDFMEVVLGKLDPQKAFFS GRLKARGNIMLSQKLQMILKDYAKL [SEQ ID NO:242]	Deficiência de Proteína D-bifuncional Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
UROD	7389	0126088	P06132	MEANGLGPQGFPELKNDTFLRAAWGEETDYTPVWCMRQAGRYLPEFRE TRAAQDFFSTCRSPEACCELTLQPLRRFPLDAAIIFSDILVVPQALGMEVTM VPGKGPSFPEPLREEQDLERLRDPEVVASELGYVFQAITLTRQRLAGRVP LIGFAGAPWTLMTYMVEGGGSSTMAQAKRWLYQRPQASHQLLRILTDAL VPYLVGQVVAGAQALQLFESHAGHLGPQLFNKFALPYIRDVAKQVKARLR EAGLAPVPMIIFAKDGHFALEELAQAGYEVVGLDWTVAPKKARECVGKTV TLQGNLDPCALYASEEEIGQLVKQMLDDFGPHRYIANLGHGLYPDMDPEH VGAFVDAVHKHSRLLRQN [SEQ ID NO:243]	Porfiria Cutanea Tardia	
HFE	3077	0010704	Q30201	MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGLSLFEAL GYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMF TVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE	Porfiria Cutanea Tardia	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLE LGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQP MDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIV IWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQGSRGAMGHYVLAERE [SEQ ID NO:244]		
LPL	4023	0175445	P06858, A0A1B1RVA9	MESKALLVLTLAVWLQSLTASRGGVAAADQRRDFIDIESKFALRTPEDTAE DTCHLIPGVAESVATCHFNHSSKTFMVIHGWTVTGMYESWVPKLVAALYK REPDSNVIVVDWLSRAQEHYPVSAGYTKLVGQDVARFINWMEEEFNYPL DNVHLLGYSLGAHAAGIAGSLTNKKVNRITGLDPAGPNFEYAEAPSRLSPD DADFVDVLHTFTRGSPGRSIGIQKPVGHVDIYPNGGTFQPGCNIGEAIRVIA ERGLGDVDQLVKCSHERSIHLFIDSLLNEENPSKAYRCSSKEAFEKGLCLS CRKNRCNNLGYEINKVRAKRSSKMYLKTRSQMPYKVFHYQVKIHFSGTES ETHTNQAFEISLYGTVAESENIPFTLPEVSTNKTYSFLIYTEVDIGELLMLKL KWKSDSYFSWSDWWSSPGFAIQKIRVKAGETQKKVIFCSREKVSHLQKG KAPAVFVKCHDKSLNKKSG [SEQ ID NO:245]	Deficiência de Lipoproteína Lipase (“hiperlipoproteinemia tipo Ia; Síndrome de Buerger-Gruetz ou Hipercilomicronemia familiar)	
GRHPR	9380	0137106	Q9UBQ7	MRPVRLMKVFVTRRIPAEGRVALARAADCEVEQWDSDEPIPAKELERGVA GAHGLLCLLSDHVDKRILDAAGANLKVISTMSVGIDHLALDEIKKRGIRVGY TPDVLTDTTAELAVSLLLTTCRRLPEAIEEVKNGGWTSWKPLWLCGYGLT QSTVGIIGLGRIGQAIARRLKPFGVQRFLYTGRQPRPEEAAEFQAEFVSTP ELAAQSDFIVVACSLTPATEGLCNKDFFQKMKETAVFINISRGDVVNQDDL YQALASGKIAAAGLDVTSPEPLPTNHPLLTLKNCVILPHIGSATHRTRNTMS LLAANNLLAGLRGEPMPSELKL [SEQ ID NO:246]	Hiperoxalúria Primária Tipo II	
HOGA1	112817	0241935	Q86XE5	MLGPQVWSSVRQGLSRSLSRNVGVWASGEGKKVDIAGIYPPVTTPFTAT AEVDYGKLEENLHKLGTFPFRGFVVQGSNGEFPFLTSSERLEVVSRVRQA MPKNRLLLAGSGCESTQATVEMTVSMAQVGADAAMVVTPCYYRGRMSS AALIHHYTKVADLSPIPVVLYSVPANTGLDLPVDAVVTLSQHPNIVGMKDSG GDVTRIGLIVHKTRKQDFQVLAGSAGFLMASYALGAVGGVCALANVLGAQ VCQLERLCCTGQWEDAQKLQHRLIEPNAAVTRRFGIPGLKKIMDWFGYYG GPCRAPLQELSPAEEEALRMDFTSNGWL [SEQ ID NO:247]	Hiperoxalúria Primária Tipo III	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
LDLR	3949	0130164	P01130, A0A024R7D5	MGPWGWKLRWTVALLLAAAGTAVGDRCERNEFQCQDGKCISYKWVCDG SAECQDGSDESQETCLSVTCKSGDFSCGGRVNRCIPQFWRCDGQVDCD NGSDEQGCPPKTCSQDEFRCHDGKCISRQFVCDSDRDCLDGSDEASCP VLTCGPASFQCNSSTCIPQLWACDNDPDCEDGSDEWPQRCRGLYVFQG DSSPCSAFEFHCLSGECIHSSWRCDGGPDCKDKSDEENCAVATCRPDEF QCSDGNCIHGSRQCDREYDCKDMSDEVGCVNVTLCEGPNKFKCHSGECI TLDKVCNMARDCRDWSDEPIKECGTNECLDNNGGCSHVCNDLKIGYECL CPDGFQLVAQRRCEDIDECQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEYTSLIPNLRNVVALDTEVASN RIYWSDLSQRMICSTQLDRAHGVSSYDTVISRDIQAPDGLAVDWIHSNIYW TDSVLGTVSVADTKGVKRKTLFRENGSKPRAIVVDPVHGFMYWTDWGTP AKIKKGGLNGVDIYSLVTENIQWPNGITLDLLSGRLYWVDSKLHSISSIDVN GGNRKTILEDEKRLAHPFSLAVFEDKVFWTDIINEAIFSANRLTGSDVNLLA ENLLSPEDMVLFHNLTQPRGVNWCERTTLSNGGCQYLCLPAPQINPHSPK FTCACPDGMLLARDMRSCLTEAEAAVATQETSTVRLKVSSTAVRTQHTTT RPVPDTSRLPGATPGLTTVEIVTMSHQALGDVAGRGNEKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNPVYQKTTEDEVHICHNQDGY SYPSRQMVSLEDDVA [SEQ ID NO:248]	Hipercolesterolemia Familiar Homozigótica	
ACAD8	27034	0151498	Q9UKU7	MLWSGCRRFGARLGCLPGGLRVLVQTGHRS LTSCIDPSMGLNEEQKEFQKVAFDFAAREM APNMAEWDQKELFPVDVMRKAAQLGFGGVY IQTDVGGSGLSRLDTSVIFEALATGCTSTT AYISIHNMCAWMIDSFGNEE QRHKFCPPLCTMEKFASYCLTEPGSGSDAA SLLTSAKKQGDHYILNGSKAFISGAGESDI YVVMCRTGGPGPKGISCIVVEKGTPGLSFG KKEKKVGWNSQPTRAVIFEDCAVPVANRIG SEGQGFLIAVRGLNGGRINIASCSLGAAHA SVILTRDHLNVRKQFGEPLASNQYLQFTLADMATRLVAARLMVRNAAVAL QEERKDAVALCSMAKLFATDECFAICNQALQMHGGYGYLKDYAVQQYVR DSRVHQILEGSNEVMRILISRSLLQE [SEQ ID NO:249]	deficiência de Isobutiril-CoA Desidrogenase (IBD)	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
ACA-DSB	36	0196177	P45954, A0A0S2Z3P9	MEGLAVRLLRGSRLLRRNFLTCLSSWKIPPHVSKSSQSEALLNITNNGIHF APLQTFTDEEMMIKSSVKKFAQEQIAPLVSTMDENSKMEKSVIQGLFQQGL MGIEVDPEYGGTGASFLSTVLVIEELAKVDASVAVFCEIQNTLINTLIRKHGT EEQKATYLPQLTTEKVGSFCLSEAGAGSDSFALKTRADKEGDYYVLNGSK MWISSAEHAGLFLVMANVDPTIGYKGITSFLVDRDTPGLHIGKPENKLGLR ASSTCPLTFENVKVPEANILGQIGHGYKYAIGSLNEGRIGIAAQMLGLAQGC FDYTIPYIKERIQFGKRLFDFQGLQHQVAHVATQLEAARLLTYNAARLLEAG KPFIKEASMAKYYASEIAGQTTSKCIEWMGGVGYTKDYPVEKYFRDAKIGT IYEGASNIQLNTIAKHIDAEY [SEQ ID NO:250]	deficiência de acil-CoA Desidrogenase (SBCAD) de Cadeia Curta	Acidemia orgânica
ACAT 1	38	0075239	A0A140VJX1, P24752	MAVLAALLRSGARSRSPLLRRLVQEIRYVERSYVSKPTLKEVVIVSATRTPI GSFLGSLSLLPATKLGSIAIQGAIEKAGIPKEEVKEAYMGNVLQGGEGQAP TRQAVLGAGLPISTPCTTINKVCASGMKAIMMASQSLMCGHQDVMVAGG MESMSNVPYVMNRGSTPYGGVKLEDLIVKDGLTDVYNKIHMGSCAENTA KKLNIARNEQDAYAINSYTRSKAAWEAGKFGNEVIPVTVTVKGQPDVVVKE DEEYKRVDFSKVPKLKTVFQKENGTVTAANASTLNDGAAALVLMTADAAK RLNVTPLARIVAFADAAVEPIDFPIAPVYAASMVLKDVGLKKEDIAMWEVNE AFSLVVLANIKMLEIDPQKVNINGGAVSLGHPIGMSGARIVGHLTHALKQGE YGLASICNGGGGASAMLIQKL [SEQ ID NO:251]	deficiência de beta-Cetotiolase	Acidemia orgânica
ACSF 3	197322	0176715	Q4G176, F5H5A1	MLPHVVLTFRRLGCALASCRLAPARHRGSGLLHTAPVARSDRSAPVFTRA LAFGDRIALDQHGRHTYRELYSRSLRLSQEICRLCGCVGGDLREERVSFL CANDASYVVAQWASWMSGGVAVPLYRKHPAAQLEYVICDSQSSVVLASQ EYLELLSPVVRKLGVPLLPLTPAIYTGAVEEPAEVPVPEQGWRNKGAMIIYT SGTTGRPKGVLSTHQNIRAVVTGLVHKWAWTKDDVILHVLPLHHVHGVVN ALLCPLWVGATCVMMPEFSPQQVWEKFLSSETPRINVFMAVPTIYTKLME YYDRHFTQPHAQDFLRAVCEEKIRLMVSGSAALPLPVLEKWKNITGHTLLE RYGMTEIGMALSGPLTTAVRLPGSVGTPLPGVQVRIVSENPQREACSYTIH AEGDERGTKVTPGFEEKEGELLVRGPSVFREYWNKPEETKSAFTLDGWF KTGDTVVFKDGQYWIRGRTSVDIIKTGGYKVSALEVEWHLLAHPSITDVAVI	acidúria malônica e metilmalônica combinada	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GVPDMTWGQRVTAVVTLREGHSLSHRELKEWARNVLAPYAVPSELVLVE EIPRNQMGKIDKKALIRHFHPS [SEQ ID NO:252]		
ASPA	443	0108381	P45381, Q6FH48	MTSCHIAEEHIQKVAIFGGTHGNELTGVFLVKHWLENGAEIQRTGLEVKPFI TNPRAVKKCTRYIDCDLNRIFDLENLGKKMSEDLPYEVRRAQEINHLFGPK DSEDSYDIIFDLHNTTSNMGCTLILEDSRNNFLIQMFHYIKTSLAPLPCYVYL IEHPSLKYATTRSIAKYPVGIEVGPQPQGVLRADILDQMRKMIKHALDFIHH FNEGKEFPPCAIEVYKIIEKVDYPRDENGEIA AIIHPNLQDQDWKPLHPGDPMFLTLDGKTIPLGGDCTVYPVFVNEAAYYEK KEAFAKTTKLTLNAKSIRCCLH [SEQ ID NO:253]	Doença de Canavan	Acidemia orgânica
AUH	549	0148090	Q13825, B4DYI6	MAAAVAAAPGALGSLHAGGARLVAACSAWLCPGLRLPGSLAGRRAGPAI WAQGWVPAAGGPAPKRGYSSEMKTEDELRVRHLEEENRGIVVLGINRAY GKNSLSKNLIKMLSKAVDALKSDKKVRTIIIRSEVPGIFCAGADLKERAKMS SSEVGPFVSKIRAVINDIANLPVPTIAAIDGLALGGGLELALACDIRVAASSA KMGLVETKLAIIPGGGGTQRLPRAIGMSLAKELIFSARVLDGKEAKAVGLIS HVLEQNQEGDAAYRKALDLAREFLPQGPVAMRVAKLAINQGMEVDLVTGL AIEEACYAQTIPTKDRLEGLLAFKEKRPPRYKGE [SEQ ID NO:254]	acidemia 3-metilglutacônica tipo I	Acidemia orgânica
DNAJC19	131118	0205981	Q96DA6, A0A0S2Z5X1	MASTVVAVGLTIAAAGFAGRYVLQAMKHMEPQVKQVFQSLPKSAFSGGY YRGGFEPKMTKREAALILGVSPTANKGKIRDAHRRIMLLNHPDKGGSPYIA AKINEAKDLLEGQAKK [SEQ ID NO:255]	cardiomiopatia dilatada com síndrome de ataxia (causa acidúria 3-metilglutacônica)	Acidemia orgânica
ETHE1	23474	0105755	A0A0S2Z580, O95571, A0A0S2Z5N8, A0A0S2Z5B3,	MAEAVLRVARRQLSQRGGSGAPILLRQMFEPVSCTFTYLLGDRESREAVL IDPVLETAPRDAQLIKELGLRLLYAVNTHCHADHITGSGLLRSLLPGCQSVI SRLSGAQADLHIEDGDSIRFGRFALETRASPGHTPGCVTFVLNDHSMAFT GDALLIRGCGRTDFQQGCAKTLYHSVHEKIFTLPGDCLIYPAHDYHGFTVS	encefalopatia etilmalônica	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
			B2RCZ7	TVEEERTLNPRLTLSCEEFVKIMGNLNLPKPQQIDFAVPANMRCGVQTPTA [SEQ ID NO:256]		
FBP1	2203	0165140	P09467, Q2TU34	MADQAPFDTDVNTLTRFVMEEGRKARGTGELTQLLNSLCTAVKAISSAVR KAGIAHLYGIAGSTNVTGDQVKKLDVLSNDLVMNMLKSSFATCVLVSEEDK HAIIVEPEKRGKYVVCFDPLDGSSNIDCLVSVGTIFGIYRKKSTDEPSEKDA LQPGRNLVAAGYALYGSATMLVLAMDCGVNCFMLDPAIGEFILVDKDVKIK KKGKIYSLNEGYARDFDPAVTEYIQRKKFPPDNSAPYGARYVGSMVADVH RTLVYGGIFLYPANKKSPNGKLRLLYECNPMAYVMEKAGGMATTGKEAVL DVIPTDIHQRAPVILGSPDDVLEFLKVYEKHSAQ[SEQ ID NO:257]	deficiência de frutose 1,6-bisfosfatase	Acidemia orgânica
FTCD	10841	0160282,0281775	O95954	MSQLVECVPNFSEGKNQEVIDAISGAITQTPGCVLLDVDAGPSTNRTVYTF VGPPECVVEGALNAARVASRLIDMSRHQGEHPRMGALDVCPFIPVRGVS VDECVLCAQAFGQRLAEELDVPVYLYGEAARMDSRRTLPAIRAGEYEALP KKLQQADWAPDFGPSSFVPSWGATATGARKFLIAFNINLLGTKEQAHRIAL NLREQGRGKDQPGRLKKVQGIGWYLDEKNLAQVSTNLLDFEVTALHTVYE ETCREAQELSLPVVGSQLVGLVPLKALLDAAAFYCEKENLFILEEEQRI RLVVSRLGLDSLCPFSPKERIIEYLVPERGPERGLGSKSLRAFVGEVGARS AAPGGGSVAAAAAAMGAALGSMVGLMTYGRRQFQSLDTTMRRLIPPFRE ASAKLTTLVDADAEAFTAYLEAMRLPKNTPEEKDRRTAALQEGLRRAVSV PLTLAETVASLWPALQELARCGNLACRSDLQVAAKALEMGVFGAYFNVLI NLRDITDEAFKDQIHHRVSSLLQEAKTQAALVLDCLETRQE [SEQ ID NO:258]	deficiência de glutamato formiminotransferase (FIGLU	Acidemia orgânica
GSS	2937	0100983	P48637, V9HWJ1	MATNWGSLLQDKQQLEELARQAVDRALAEGVLLRTSQEPTSSEVVSYAP FTLFPSLVPSALLEQAYAVQMDFNLLVDAVSQNAAFLEQTLSSTIKQDDFT ARLFDIHKQVLKEGIAQTVFLGLNRSDYMFQRSADGSPALKQIEINTISASF GGLASRTPAVHRHVLSVLSKTKEAGKILSNNPSKGLALGIAKAWELYGSPN ALVLLIAQEKERNIFDQRAIENELLARNIHVIRRTFEDISEKGSLDQDRRLFV DGQEIAVVYFRDGYMPRQYSLQNWEARLLLERSHAAKCPDIATQLAGTKK	deficiência de glutationa sintetase	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				VQQELSRPGMLEMLLPGQPEAVARLRATFAGLYSLDVGEEGDQAIAEALA APSRFVLKPQREGGGNNLYGEEMVQALKQLKDSEERASYILMEKIEPEPF ENCLLRPGSPARVVQCISELGIFGVYVRQEKTLVMNKHVGHLLRTKAIEHA DGGVAAGVAVLDNPYPV [SEQ ID NO:259]		
HIBCH	26275	0198130	A0A140VJL0, Q6NVY1	MGQREMWRLMSRFNAFKRTNTILHHLRMSKHTDAAEEVLLEKKGCTGVIT LNRPKFLNALTLNMIRQIYPQLKKWEQDPETFLIIIKGAGGKAFCAGGDIRVI SEAEKAKQKIAPVFFREEYMLNNAVGSCQKPYVALIHGITMGGGVGLSVH GQFRVATEKCLFAMPETAIGLFPDVGGGYFLPRLQGKLGYFLALTGFRLK GRDVYRAGIATHFVDSEKLAMLEEDLLALKSPSKENIASVLENYHTESKIDR DKSFILEEHMDKINSCFSANTVEEIIENLQQDGSSFALEQLKVINKMSPTSL KITLRQLMEGSSKTLQEVLTMEYRLSQACMRGHDFHEGVRAVLIDKDQSP KWKPADLKEVTEEDLNNHFKSLGSSDLKF [SEQ ID NO:260]	Deficiência de 3-hidroxiisobutiril-CoA hidrolase	Acidemia orgânica
IDH2	3418	0182054	P48735, B4DSZ6	MAGYLRVVRSLCRASGSRPAWAPAALTAPTSQEQPRRHYADKRIKVAKP VVEMDGDEMTRIIWQFIKEKLILPHVDIQLKYFDLGLPNRDQTDDQVTIDSA LATQKYSVAVKCATITPDEARVEEFKLKKMWKSPNGTIRNILGGTVFREPII CKNIPRLVPGWTKPITIGRHAHGDQYKATDFVADRAGTFKMVFTPKDGSG VKEWEVYNFPAGGVGMGMYNTDESISGFAHSCFQYAIQKKWPLYMSTKN TILKAYDGRFKDIFQEIFDKHYKTDFDKNKIWYEHRLIDDMVAQVLKSSGGF VWACKNYDGDVQSDILAQGFGSLGLMTSVLVCPDGKTIEAEAAHGTVTRH YREHQKGRPTSTNPIASIFAWTRGLEHRGKLDGNQDLIRFAQMLEKVCVE TVESGAMTKDLAGCIHGLSNVKLNEHFLNTTDFLDTIKSNLDRALGRQ [SEQ ID NO:261]	Acidúria D-2-hidroxiglutárica tipo II	Acidemia orgânica
L2HGDH	79944	0087299	Q9H9P8	MVPALRYLVGACGRARGLFAGGSPGACGFASGRPRPLCGGSRSASTSSF DIVIVGGGIVGLASARALILRHPSLSIGVLEKEKDLAVHQTGHNSGVIHSGIY YKPESLKAKLCVQGAALLYEYCQQKGISYKQCGKLIVAVEQEEIPRLQALY EKGLQNGVPGLRLIQQEDIKKKEPYCRGLMAIDCPHTGIVDYRQVALSFAQ DFQEAGGSVLTNFEVKGIEMAKESPSRSIDGMQYPIVIKNTKGEEIRCQYV VTCAGLYSDRISELSGCTPDPRIVPFRGDYLLLKPEKCYLVKGNIYPVPDSR	Acidúria L-2-hidroxiglutárica	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FPFLGVHFTPRMDGSIWLGPNAVLAFKREGYRPFDFSATDVMDIIINSGLIK LASQNFSYGVTEMYKACFLGATVKYLQKFIPEITISDILRGPAGVRAQALDR DGNLVEDFVFDAGVGDIGNRILHVRNAPSPAATSSIAISGMIADEVQQRFEL [SEQ ID NO:262]		
MLYCD	23417	0103150	O95822	MRGFGPGLTARRLLPLRLPPRPPGPRLASGQAAGALERAMDELLRRAVP PTPAYELREKTPAPAEGQCADFVSFYGGLAETAQRAELLGRLARGFGVDH GQVAEQSAGVLHLRQQQREAAVLLQAEDRLRYALVPRYRGLFHHISKLDG GVRFLVQLRADLLEAQALKLVEGPDVREMNGVLKGMLSEWFSSGFLNLE RVTWHSPCEVLQKISEAEAVHPVKNWMDMKRRVGPYRRCYFFSHCSTP GEPLVVLHVALTGDISSNIQAIVKEHPPSETEEKNKITAAIFYSISLTQQGLQ G VELGTFLIKRVVKELQREFPHLGVFSSLSPIPGFTKWLLGLLNSQTKEHGR NELFTDSECKEISEITGGPINETLKLLLSSSEWVQSEKLVRALQTPLMRLCA WYLYGEKHRGYALNPVANFHLQNGAVLWRINWMADVSLRGITGSCGLMA NYRYFLEETGPNSTSYLGSKIIKASEQVLSLVAQFQKNSKL [SEQ ID NO:263]	acidemia malônica	Acidemia orgânica
OPA3	80207	0125741	Q9H6K4, B4DK77	MVVGAFPMAKLLYLGIRQVSKPLANRIKEAARRSEFFKTYICLPPAQLYHW VEMRTKMRIMGFRGTVIKPLNEEAAAELGAELLGEATIFIVGGGCLVLEYW RHQAQQRHKEEEQRAAWNALRDEVGHLALALEALQAQVQAAPPQGALE ELRTELQEVRAQLCNPGRSASHAVPASKK [SEQ ID NO:264]	Síndrome de Costeff/acidúria 3-metilglutacônica tipo III	Acidemia orgânica
OPLAH	26873	0178814	O14841	MGSPEGRFHFAIDRGGTFTDVFAQCPGGHVRVLKLLSEDPANYADAPTE GIRRILEQEAGMLLPRDQPLDSSHIASIRMGTTVATNALLERKGERVALLVT RGFRDLLHIGTQARGDLFDLAVPMPEVLYEEVLEVDERVVLHRGEAGTGT PVKGRTGDLLEVQQPVDLGALRGKLEGLLSRGIRSLAVVLMHSYTWAQHE QQVGVLARELGFTHVSLSSEAMPMVRIVPRGHTACADAYLTPAIQRYVQG FCRGFQGQLKDVQVLFMRSDGGLAPMDTFSGSSAVLSGPAGGVVGYSA TTYQQEGGQPVIGFDMGGTSTDVSRYAGEFEHVFEASTAGVTLQAPQLDI NTVAAGGGSRLFFRSGLFVVGPESAGAHPGPACYRKGGPVTVTDANLVL GRLLPASFPCIFGPGENQPLSPEASRKALEAVATEVNSFLTNGPCPASPLS LEEVAMGFVRVANEAMCRPIRALTQARGHDPSAHVLACFGGAGGQHACA	Deficiência de 5-oxoprolinase	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				IARALGMDTVHIHRHSGLLSALGLALADVVHEAQEPCSLLYAPETFVQLDQRLSRLEEQCVDALQAQGFPRSQISTESFLHLRYQGTDCALMVSAHQHPATA RSPRAGDFGAAFVERYMREFGFVIPERPVVVDDVRVRGTGRSGLRLEDAPKAQTGPPRVDKMTQCYFEGGYQETPVYLLAELGYGHKLHGPCLIIDSNSTILVEPGCQAEVTKTGDICISVGAEVPGTVGPQLDPIQLSIFSHRFMSIAEQMGRILQRTAISTNIKERLDFSCALFGPDGGLVSNAPHIPVHLGAMQETVQFQIQHLGADLHPGDVLLSNHPSAGGSHLPDLTVITPVFWPGQTRPVFYVASRGHHADIGGITPGSMPPHSTMLQQEGAVFLSFKLVQGGVFQEEAVTEALRAPGKVPNCSGTRNLHDNLSDLRAQVAANQKGIQLVGELIGQYGLDVVQAYMGHIQANAELAVRDMLRAFGTSRQARGLPLEVSSEDHMDDGSPIRLRVQISLSQGSAVFDFSGTGPEVFGNLNAPRAVTLSALIYCLRCLVGRDIPLNQGCLAPVRVVIPRGSILDPSPEAAVVGGNVLTSQRVVDVILGAFGACAASQGCMNNVTLGNAHMGYYETVAGGAGAGPSWHGRSGVHSHMTNTRITDPEILESRYPVILRRFELRRGSGGRGRFRGGDGVTRELLFREEALLSVLTERRAFRPYGLHGGEPGARGLNLLIRKNGRTVNLGGKTSVTVYPGDVFCLHTPGGGGYGDPEDPAPPPGSPPQALAFPEHGSVYEYRRAQEAV [SEQ ID NO:265]		
OXCT1	5019	0083720	A0A024R040, P55809	MAALKLLSSGLRLCASARGSGATWYKGCVCSFSTSAHRHTKFYTDPVEAVKDIPDGATVLVGGFGLCGIPENLIDALLKTGVKGLTAVSNNAGVDNFGLGLLLRSKQIKRMVSSYVGENAEFERQYLSGELEVELTPQGTLAERIRAGGAGVPAFYTPTGYGTLVQEGGSPIKYNKDGSVAIASKPREVREFNGQHFILEEAITGDFALVKAWKADRAGNVIFRKSARNFNLPMCKAAETTVVEVEEIVDIGAFAPEDIHIPQIYVHRLIKGEKYEKRIERLSIRKEGDGEAKSAKPGDDVRERIIKRAALEFEDGMYANLGIGIPLLASNFISPNITVHLQSENGVLGLGPYPRQHEADADLINAGKETVTILPGASFFSSDESFAMIRGGHVDLTMLGAMQVSKYGDLANWMIPGKMVKGMGGAMDLVSSAKTKVVVTMEHSAKGNAHKIMEKCTLPLTGKQCVNRIITEKAVFDVDKKKGLTLIELWEGLTVDDVQKSTGCDFAVSPKLMPMQQIAN [SEQ ID NO: 266]	Deficiência de SCOT	Acidemia orgânica
POLG	5428	0140521	E5KNU5, P54098	MSRLLWRKVAGATVGPGPVPAPGRWVSSSVPASDPSDGQRRRQQQQQ	acidúria 3-	Acidemia orgâni-

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QQQQQQQQPQQPQVLSSEGGQLRHNPLDIQMLSRGLHEQIFGQGGEMP GEAAVRRSVEHLQKHGLWGQPAVPLPDVELRLPPLYGDNLDQHFRLLAQ KQSLPYLEAANLLLQAQLPPKPPAWAWAEGWTRYGPEGEAVPVAIPEER ALVFDVEVCLAEGTCPTLAVAISPSAWYSWCSQRLVEERYSWTSQLSPAD LIPLEVPTGASSPTQRDWQEQLVVGHNVSFDRAHIREQYLIQGSRMRFLD TMSMHMAISGLSSFQRSLWIAAKQGKHKVQPPTKQGQKSQRKARRGPAI SSWDWLDISSVNSLAEVHRLYVGGPPLEKEPRELFVKGTMKDIRENFQDL MQYCAQDVWATHEVFQQQLPLFLERCPHPVTLAGMLEMGVSYLPVNQN WERYLAEAQGTYEELQREMKKSLMDLANDACQLLSGERYKEDPWLWDL EWDLQEFKQKKAKKVKKEPATASKLPIEGAGAPGDPMDQEDLGPCSEEE EFQQDVMARACLQKLKGTTELLPKRPQHLPGHPGWYRKLCPRLDDPAWT PGPSLLSLQMRVTPKLMALTWDGFPLHYSERHGWGYLVPGRRDNLAKLP TGTTLESAGVVCPYRAIESLYRKHCLEQGKQQLMPQEAGLAEEFLLTDNS AIWQTVEELDYLEVEAEAKMENLRAAVPGQPLALTARGGPKDTQPSYHH GNGPYNDVDIPGCWFFKLPHKDGNSCNVGSPFAKDFLPKMEDGTLQAGP GGASGPRALEINKMISFWRNAHKRISSQMVVWLPRSALPRAVIRHPDYDE EGLYGAILPQVVTAGTITRRAVEPTWLTASNARPDRVGSELKAMVQAPPG YTLVGADVDSQELWIAAVLGDAHFAGMHGCTAFGWMTLQGRKSRGTDLH SKTATTVGISREHAKIFNYGRIYGAGQPFAERLLMQFNHRLTQQEAAEKAQ QMYAATKGLRWYRLSDEGEWLVRELNLPVDRTEGGWISLQDLRKVQRET ARKSQWKKWEVVAERAWKGGTESEMFNKLESIATSDIPRTPVLGCCISRA LEPSAVQEEFMTSRVNWVVQSSAVDYLHLMLVAMKWLFEEFAIDGRFCISI HDEVRYLVREEDRYRAALALQITNLLTRCMFAYKLGLNDLPQSVAFFSAVD IDRCLRKEVTMDCKTPSNPTGMERRYGIPQGEALDIYQIIELTKGSLEKRSQ PGP [SEQ ID NO:267]	metilglutacônica	ca
PPM1K	152926	0163644	Q8N3J5	MSTAALITLVRSGGNQVRRRVLLSSRLLQDDRRVTPTCHSSTSEPRCSRF DPDGSGSPATWDNFGIWDNRIDEPILLPPSIKYGKPIPKISLENVGCASQIG KRKENEDRFDFAQLTDEVLYFAVYDGHGGPAAADFCHTHMEKCIMDLLPK EKNLETLLTLAFLEIDKAFSSHARLSADATLLTSGTTATVALLRDGIELVVAS VGDSRAILCRKGKPMKLTIDHTPERKDEKERIKKCGGFVAWNSLGQPHVN GRLAMTRSIGDLDLKTSGVIAEPETKRIKLHHADDSFLVLTTDGINFMVNSQ EICDFVNQCHDPNEAAHAVTEQAIQYGTEDNSTAVVVPFGAWGKYKNSEI	doença da urina em xarope de bordo (MSUD), tipo de variante	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				NFSFSRSFASSGRWA [SEQ ID NO:268]		
SE-RAC1	84947	0122335	Q96JX3	MSLAAYCVICCRRIGTSTSPPKSGTHWRDIRNIIKFTGSLILGGSLFLTYEVL ALKKAVTLDTQVVEREKMKSYIYVHTVSLDKGENHGIAWQARKELHKAVR KVLATSAKILRNPFADPFSTVDIEDHECAVWLLLRKSKSDDKTTRLEAVRE MSETHHWHDYQYRIIAQACDPKTLIGLARSEESDLRFFLLPPPLPSLKEDS STEEELRQLLASLPQTELDECIQYFTSLALSESSQ SLAAQKGGLWCFGGNGLPYAESFGEVPSATVEMFCLEAIVKHSEISTHCD KIEANGGLQLLQRLYRLHKDCPKVQRNIMRVIGNMALNEHLHSSIVRSGW VSIMAEAMKSPHIMESSHAARILANLDRETVQEKYQDGVYVLHPQYRTSQ PIKADVLFIHGLMGAAFKTWRQQDSEQAVIEKPMEDEDRYTTCWPKTWLA KDCPALRIISVEYDTSLSDWRARCPMERKSIAFRSNELLRKLRAAGVGDRP VVWISHSMGGLLVKKMLLEASTKPEMSTVINNTRGIIFYSVPHHGSRLAEY SVNIRYLLFPSLEVKELSKDSPALKTLQDDFLEFAKDKNFQVLNFVETLPTYI GSMIKLHVVPVESADLGIGDLIPVDVNHLNICKPKKKDAFLYQRTLQFIREAL AKDLEN [SEQ ID NO:269]	Síndrome de Megdel	Acidemia orgânica
SLC25A1	6576	0100075	D9HTE9, B4DP62, P53007	MPAPRAPRALAAAAPASGKAKLTHPGKAILAGGLAGGIEICITFPTEYVKTQ LQLDERSHPPRYRGIGDCVRQTVRSHGVLGLYRGLSSLLYGSIPKAAVRF GMFEFLSNHMRDAQGRLDSTRGLLCGLGAGVAEAVVVVCPMETIKVKFIH DQTSPNPKYRGFFHGVREIVREQGLKGTYQGLTATVLKQGSNQAIRFFVM TSLRNWYRGDNPNKPMNPLITGVFGAIAGAASVFGNTPLDVIKTRMQGLE AHKYRNTWDCGLQILKKEGLKAFYKGTVPRLGRVCLDVAIVFVIYDEV VKLLNKVWKTD [SEQ ID NO:270]	Acidúria D, L-2-hidroxiglutárica	Acidemia orgânica
SUCLA2	8803	0136143	E5KS60, Q9P2R7, Q9Y4T0	MAASMFYGRLVAVATLRNHRPRTAQRAAAQVLGSSGLFNNHGLQVQQQ QQRNLSLHEYMSMELLQEAGVSVPKGYVAKSPDEAYAIAKKLGSKDVVIK AQVLAGGRGKGTFESGLKGGVKIVFSPEEAKAVSSQMIGKKLFTKQTGEK GRICNQVLVCERKYPRREYYFAITMERSFQGPVLIGSSHGGVNIEDVAAES PEAIIKEPIDIEEGIKKEQALQLAQKMGFPPNIVESAAENMVKLYSLFLKYDA TMIEINPMVEDSDGAVLCMDAKINFDSNSAYRQKKIFDLQDWTQEDERDK DAAKANLNYIGLDGNIGCLVNGAGLAMATMDIIKLHGGTPANFLDVGGGAT VHQVTEAFKLITSDKKVLAILVNIFGGIMRCDVIAQGIVMAVKDLEIKIPVVVR	deficiência de succinato-CoA ligase, acidúria metilmalônica	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LQGTRVDDAKALIADSGLKILACDDLDEAARMVVKLSEIVTLAKQAHVDVKF QLPI [SEQ ID NO:271]		
SUCLG1	8802	0163541	P53597	MTATLAAAADIATMVSGSSGLAAARLLSRSFLLPQNGIRHCSYTASRQHLY VDKNTKIICQGFTGKQGTFHSQQALEYGTKLVGGTTPGKGGQTHLGLPVF NTVKEAKEQTGATASVIYVPPPFAAAAINEAIEAEIPLVVCITEGIPQQDMVR VKHKLLRQEKTRLIGPNCPGVINPGECKIGIMPGHIHKKGRIGIVSRSGTLTY EAVHQTTQVGLGQSLCVGIGGDPFNGTDFIDCLEIFLNDSATEGIILIGEIGG NAEENAAEFLKQHNSGPNSKPVVSFIAGLTAPPGRRMGHAGAIIAGGKGG AKEKISALQSAGVVVSMSPAQLGTTIYKEFEKRKML [SEQ ID NO:272]	deficiência de succinato-CoA ligase, acidúria metilmalônica	Acidemia orgânica
TAZ	6901	0102125	A0A0S2Z4K0, Q16635, A6XNE1, A0A0S2Z4E6, A0A0S2Z4K9, A0A0S2Z4F4	MPLHVKWPFPAVPPLTWTLASSVVMGLVGTYSCFWTKYMNHLTVHNREV LYELIEKRGPATPLITVSNHQSCMDDPHLWGILKLRHIWNLKLMRWTPAAA DICFTKELHSHFFSLGKCVPVCRGAEFFQAENEGKGVLDTGRHMPGAGK RREKGDGVYQKGMDFILEKLNHGDWVHIFPEGKVNMSSEFLRFKWGIGR LIAECHLNPIILPLWHVGMNDVLPNSPPYFPRFGQKITVLIGKPFSALPVLER LRAENKSAVEMRKALTDFIQEEFQHLKTQAEQLHNHLQPGR [SEQ ID NO:273]	Síndrome de Barth	Acidemia orgânica
AGK	55750	0006530,0262327	A4D1U5, Q53H12	MTVFFKTLRNHWKKTTAGLCLLTWGGHWLYGKHCDNLLRRAACQEAQVF GNQLIPPNAQVKKATVFLNPAACKGKARTLFEKNAAPILHLSGMDVTIVKT DYEGQAKKLLELMENTDVIIVAGGDGTLQEVVTGVLRRTDEATFSKIPIGFI PLGETSSLSHTLFAESGNKVQHITDATLAIVKGETVPLDVLQIKGEKEQPVF AMTGLRWGSFRDAGVKVSKYWYLGPLKIKAAHFFSTLKEWPQTHQASISY TGPTERPPNEPEETPVQRPSLYRRILRRLASYWAQPQDALSQEVSPEVW KDVQLSTIELSITTRNNQLDPTSKEDFLNICIEPDTISKGDFITIGSRKVRNPK LHVEGTECLQASQCTLLIPEGAGGSFSIDSEEYEAMPVEVKLLPRKLQFFC DPRKREQMLTSPTQ [SEQ ID NO:274]	acidúria 3-metilglutacônica	Acidemia orgânica
CLPB	81570	0162129	Q9H078, A0A140VK11	MLGSLVLRRKALAPRLLLRLLRSPTLRGHGGASGRNVTTGSLGEPQWLRV ATGGRPGTSPALFSGRGAATGGRQGGRFDTKCLAAATWGRLPGPEETLP GQDSWNGVPSRAGLGMCALAAALVVHCYSKSPSNKDAALLEAARANNM QEVSRLLSEGADVNAKHRLGWTALMVAAINRNNSVVQVLLAAGADPNLG	acidúria 3-metilglutacônica	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DDFSSVYKTAKEQGIHSLEDGGQDGASRHITNQWTSALEFRRWLGLPAG VLITREDDFNNRLNNRASFKGCTALHYAVLADDYRTVKELLDGGANPLQR NEMGHTPLDYAREGEVMKLLRTSEAKYQEKQRKREAEERRRFPLEQRLK EHIIGQESAIATVGAA IRRKENGWYDEEHPLVFLFLGSSGIGKTELAKQTAKYMHKDAKKGFIRLDM SEFQERHEVAKFIGSPPGYVGHEEGGQLTKKLKQCPNAVVLFDEVDKAHP DVLTIMLQLFDEGRLTDGKGKTIDCKDAIFIMTSNVASDEIAQHALQLRQEA LEMSRNRIAENLGDVQISDKITISKNFKENVIRPILKAHFRRDEFLGRINEIVY FLPFCHSELIQLVNKELNFWAKRAKQRHNITLLWDREVADVLVDGYNVHY GARSIKHEVERRVVNQLAAAYEQDLLPGGCTLRITVEDSDKQLLKSPELPS PQAEKRLPKLRLEIIDKDSKTRRLDIRAPLHPEKVCNTI [SEQ ID NO:275]		
TMEM 70	54968	0175606	Q9BUB7	MLFLALGSPWAVELPLCGRRTALCAAAALRGPRASVSRASSSSGPSGPV AGWSTGPSGAARLLRRPGRAQIPVYWEGYVRFLNTPSDKSEDGRLIYTG NMARAVFGVKCFSYSTSLIGLTFLPYIFTQNNAISESVPLPIQIIFYGIMGSFT VITPVLLHFITKGYVIRLYHEATTDTYKAITYNAMLAETSTVFHQNDVKIPDA KHVFTTFYAKTKSLLVNPVLFPNREDYIHLMGYDKEEFILYMEETSEEKRH KDDK [SEQ ID NO:276]	acidúria 3-metilglutacônica	Acidemia orgânica
ALDH 18A1	5832	0059573	P54886	MLSQVYRCGFQPFNQHLLPWVKCTTVFRSHCIQPSVIRHVRSWSNIPFITV PLSRTHGKSFAHRSELKHAKRIVVKLGSAVVTRGDECGLALGRLASIVEQV SVLQNQGREMMLVTSGAVAFGKQRLRHEILLSQSVRQALHSGQNQLKEM AIPVLEARACAAAGQSGLMALYEAMFTQYSICAAQILVTNLDFHDEQKRRN LNGTLHELLRMNIVPIVNTNDAVVPPAEPNSDLQGVNVISVKDNDSLAARL AVEMKTDLLIVLSDVEGLFDSPPGSDDAKLIDIFYPGDQQSVTFGTKSRVG MGGMEAKVKAALWALQGGTSVVIANGTHPKVSGHVITDIVEGKKVGTFFS EVKPAGPTVEQQGEMARSGGRMLATLEPEQRAEIIHHLADLLTDQRDEILL ANKKDLEEAEGRLAAPLLKRLSLSTSKLNSLAIGLRQIAASSQDSVGRVLR RTRIAKNLELEQVTVPIGVLLVIFESRPDCLPQVAALAIASGNGLLLKGGKEA AHSNRILHLLTQEALSIHGVKEAVQLVNTREEVEDLCRLDKMIDLIIPRGSS QLVRDIQKAAKGIPVMGHSEGICHMYVDSEASVDKVTRLVRDSKCEYPAA	Cútis laxa relacionada a ALDH18A1	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				CNALETLLIHRDLLRTPLFDQIIDMLRVEQVKIHAGPKFASYLTFSPSEVKSL RTEYGDLELCIEVVDNVQDAIDHIHKYGSSHTDVIVTEDENTAEFFLQHVDS ACVFWNASTRFSDGYRFGLGAEVGISTSRIHARGPVGLEGLLTTKWLLRG KDHVVSDFSEHGSLKYLHENLPIPQRNTN [SEQ ID NO:277]		
OAT	4942	0065154	A0A140VJQ4, P04181	MFSKLAHLQRFAVLSRGVHSSVASATSVATKKTVQGPPTSDDIFEREYKY GAHNYHPLPVALERGKGIYLWDVEGRKYFDFLSSYSAVNQGHCHPKIVNA LKSQVDKLTLTSRAFYNNVLGEYEEYITKLFNYHKVLPMNTGVEAGETACK LARKWGYTVKGIQKYKAKIVFAAGNFWGRTLSAISSSTDPTSYDGFGPFM PGFDIIPYNDLPALERALQDPNVAAFMVEPIQGEAGVVVPDPGYLMGVREL CTRHQVLFIADEIQTGLARTGRWLAVDYENVRPDIVLLGKALSGGLYPVSA VLCDDDIMLTIKPGEHGSTYGGNPLGCRVAIAALEVLEEENLAENADKLGII LRNELMKLPSDVVTAVRGKGLLNAIVIKETKDWDAWKVCLRLRDNGLLAK PTHGDIIRFAPPLVIKEDELRESIEIINKTILSF [SEQ ID NO:278]	atrofia giratória (OAT)	Desordem do ciclo da ureia
CA5A	763	0174990	P35218	MLGRNTWKTSAFSFLVEQMWAPLWSRSMRPGRWCSQRSCAWQTSNNT LHPLWTVPVSVPGGTRQSPINIQWRDSVYDPQLKPLRVSYEAASCLYIWN TGYLFQVEFDDATEASGISGGPLENHYRLKQFHFHWGAVNEGGSEHTVD GHAYPAELHLVHWNSVKYQNYKEAVVGENGLAVIGVFLKLGAHHQTLQRL VDILPEIKHKDARAAMRPFDPSTLLPTCWDYWTYAGSLTTPPLTESVTWIIQ KEPVEVAPSQLSAFRTLLFSALGEEEKMMVNNYRPLQPLMNRKVWASFQ ATNEGTRS [SEQ ID NO:279]	deficiência de anidrase carbônica	Desordem do ciclo da ureia
GLUD 1	2746	0148672	P00367, E9KL48	MYRYLGEALLLSRAGPAALGSASADSAALLGWARGQPAAAPQPGLALAA RRHYSEAVADREDDPNFFKMVEGFFDRGASIVEDKLVEDLRTRESEEQKR NRVRGILRIIKPCNHVLSLSFPIRRDDGSWEVIEGYRAQHSQHRTPCKGGI RYSTDVSVDEVKALASLMTYKCAVVDVPFGGAKAGVKINPKNYTDNELEKI TRRFTMELAKKGFIGPGIDVPAPDMSTGEREMSWIADTYASTIGHYDINAH ACVTGKPISQGGIHGRISATGRGVFHGIENFINEASYMSILGMTPGFG DKTFVVQGFGNVGLHSMRYLHRFGAKCIAVGESDGSIWNPDGIDPKELED FKLQHGSILGFPKAKPYEGSILEADCDILIPAASEKQLTKSNAPRVKAKIIAE GANGPTTPEADKIFLERNIMVIPDLYLNAGGVTVSYFEWLKNLNHVSYGRL TFKYERDSNYHLLMSVQESLERKFGKHGGTIPIVPTAEFQDRISGASEKDIV	deficiência de glutamato desidrogenase	Desordem do ciclo da ureia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				HSGLAYTMERSARQIMRTAMKYNLGLDLRTAAYVNAIEKVFKVYNEAGVT FT[SEQ ID NO:280]		
GLUL	2752	0135821	A8YXX4, P15104	MTTSASSHLNKGIKQVYMSLPQGEKVQAMYIWIDGTGEGLRCKTRTLDSE PKCVEELPEWNFDGSSTLQSEGSNSDMYLVPAAMFRDPFRKDPNKLVLC EVFKYNRRPAETNLRHTCKRIMDMVSNQHPWFGMEQEYTLMGTDGHPF GWPSNGFPGPQGPYYCGVGADRAYGRDIVEAHYRACLYAGVKIAGTNAE VMPAQWEFQIGPCEGISMGDHLWVARFILHRVCEDFGVIATFDPKPIPGN WNGAGCHTNFSTKAMREENGLKYIEEAIEKLSKRHQYHIRAYDPKGGLDN ARRLTGFHETSNINDFSAGVANRSASIRIPRTVGQEKKGYFEDRRPSANC DPFSVTEALIRTCLLNETGDEPFQYKN [SEQ ID NO:281]	deficiência de glutamina sinte-tase	Desordem do ciclo da ureia
UMPS	7372	0114491	A8K5J1, P11172	MAVARAALGPLVTGLYDVQAFKFGDFVLKSGLSSPIYIDLRGIVSRPRLLSQ VADILFQTAQNAGISFDTVCGVPYTALPLATVICSTNQIPMLIRRKETKDYGT KRLVEGTINPGETCLIIEDVVTSGSSVLETVEVLQKEGLKVTDAIVLLDREQ GGKDKLQAHGIRLHSVCTLSKMLEILEQQKKVDAETVGRVKRFIQENVFVA ANHNGSPLSIKEAPKELSFGARAELPRIHPVA SKLLRLMQKKETNLCLSADVSLARELLQLADALGPSICMLKTHVDILNDFTL DVMKELITLAKCHEFLIFEDRKFADIGNTVKKQYEGGIFKIASWADLVNAHV VPGSGVVKGLQEVGLPLHRGCLLIAEMSSTGSLATGDYTRAAVRMAEEHS EFVVGFISGSRVSMKPEFLHLTPGVQLEAGGDNLGQQYNSPQEVIGKRGS DIIIVGRGIISAADRLEAAEMYRKAAWEAYLSRLGV [SEQ ID NO:282]	Acidúria Orótica	Desordem do ciclo da ureia
SLC22 A5	6584	0197375	O76082	MRDYDEVTAFLGEWGPFQRLIFFLLSASIIPNGFTGLSSVFLIATPEHRCRV PDAANLSSAWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALGLEPGRD VDLGQLEQESCLDGWEFSQDVYLSTIVTEWNLVCEDDWKAPLTISLFFVG VLLGSFISGQLSDRFGRKNVLFVTMGMQTGFSFLQIFSKNFEMFVVLFVLV GMGQISNYVAAFVLGTEILGKSVRIIFSTLGVCIFYAFGYMVLPLFAYFIRDW RMLLVALTMPGVLCVALWWFIPESPRWLISQGRFEEAEVIIRKAAKANGIV VPSTIFDPSELQDLSSKKQQSHNILDLLRTWNIRMVTIMSIMLWMTISVGYF GLSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLLQYLPRRYSMATALFLG GSVLLFMQLVPPDLYYLATVLVMVGKFGVTAAFSMVYVYTAELYPTVVRN MGVGVSSTASRLGSILSPYFVYLGAYDRFLPYILMGSLTILTAILTLFLPESF	deficiência de carnitina-acilcarnitina trans-locase (CACT)	Oxidação de ácidos graxos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GTPLPDTIDQMLRVKGMKHRKTPSHTR MLKDGQERPTILKSTAF [SEQ ID NO:283]		
CPT1A	1374	0110090	P50416, A0A024R5F4, B2RAQ8, Q8WZ48	MAEAHQAVAFQFTVTPDGIDLRLSHEALRQIYLSGLHSWKKKFIRFKNGIIT GVYPASPSSWLIVVVGVMTTMYAKIDPSLGIIAKINRTLETANCMSSQTKNV VSGVLFGTGLWVALIVTMRYSLKVLLSYHGWMFTEHGKMSRATKIWMGM VKIFSGRKPMLYSFQTSLPRLPVPAVKDTVNRYLQSVRPLMKEEDFKRMT ALAQDFAVGLGPRLQWYLKLKSWWATNYVSDWWEEYIYLRGRGPLMVN SNYYAMDLLYILPTHIQAARAGNAIHAILLYRRKLDREEIKPIRLLGSTIPLCS AQWERMFNTSRIPGEETDTIQHMRDSKHIVVYHRGRYFKVWLYHDGRLLK PREMEQQMQRILDNTSEPQPGEARLAALTAGDRVPWARCRQAYFGRGK NKQSLDAVEKAAFFVTLDETEEGYRSEDPDTSMDSYAKSLLHGRCYDRW FDKSFTFVVFKNGKMGLNAEHSWADAPIVAHLWEYVMSIDSLQLGYAEDG HCKGDINPNIPYPTRLQWDIPGECQEVIETSLNTANLLANDVDFHSFPFVAF GKGIIKKCRTSPDAFVQLALQLAHYKDMGKFCLTYEASMTRLFREGRTETV RSCTTESCDFVRAMVDPAQTVEQRLKLFKLASEKHQHMYRLAMTGSGID RHLFCLYVVSKYLAVESPFLKEVLSEPWRLSTSQTPQQQVELFDLENNPE YVSSGGGFGPVADDGYGVSYILVGENLINFHISSKFSCPETDSHRFGRHLK EAMTDIITLFGLSSNSKK [SEQ ID NO:284]	deficiência de carnitina palmitoil-transferase tipo I (CPT I)	Oxidação de ácidos graxos
HAD-HA	3030	0084754	E9KL44, P40939	MVACRAIGILSRFSAFRILRSRGYICRNFTGSSALLTRTHINYGVKGDVAVV RINSPNSKVNTLSKELHSEFSEVMNEIWASDQIRSAVLISSKPGCFIAGADI NMLAACKTLQEVTQLSQEAQRIVEKLEKSTKPIVAAINGSCLGGGLEVAISC QYRIATKDRKTVLGTPEVLLGALPGAGGTQRLPKMVGVPAALDMMLTGRS IRADRAKKMGLVDQLVEPLGPGLKPPEERTIEYLEEVAITFAKGLADKKISP KRDKGLVEKLTAYAMTIPFVRQQVYKKVEEKVRKQTKGLYPAPLKIIDVVKT GIEQGSDAGYLCESQKFGELVMTKESKALMGLYHGQVLCKKNKFGAPQK DVKHLAILGAGLMGAGIAQVSVDKGLKTILKDATLTALDRGQQQVFKGLND KVKKKALTSFERDSIFSNLTGQLDYQGFEKADMVIEAVFEDLSLKHRVLKE VEAVIPDHCIFASNTSALPISEIAAVSKRPEKVIGMHYFSPVDKMQLLEIITTE KTSKDTSASAVAVGLKQGKVIIVVK DGPGFYTTRCLAPMMSEVIRILQEGVDPKKLDSLTTSFGFPVGAATLVDEV	deficiência de 3-hidroxiacil-CoA desidrogenase (LCHAD) de cadeia longa	Oxidação de ácidos graxos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GVDVAKHVAEDLGKVFGERFGGGNPELLTQMVSKGFLGRKSGKGFYIYQ EGVKRKDLNSDMDSILASLKLPPKSEVSSDEDIQFRLVTRFVNEAVMCLQE GILATPAEGDIGAVFGLGFPPCLGGPFRFVDLYGAQKIVDRLKKYEAAYGK QFTPCQLLADHANSPNKKFYQ [SEQ ID NO:285]		
HADH	3033	0138796	Q16836, B3KTT6	MAFVTRQFMRSVSSSSTASASAKKIIVKHVTVIGGGLMGAGIAQVAAATGH TVVLVDQTEDILAKSKKGIEESLRKVAKKKFAENLKAGDEFVEKTLSTIATST DAASVVHSTDLVVEAIVENLKVKNELFKRLDKFAAEHTIFASNTSSLQITSIA NATTRQDRFAGLHFFNPVPVMKLVEVIKTPMTSQKTFESLVDFSKALGKH PVSCKDTPGFIVNRLLVPYLMEAIRLYERGDASKEDIDTAMKLGAGYPMGP FELLDYVGLDTTKFIVDGWHEMDAENPLHQPSPSLNKLVAENKFGKKTGE GFYKYK [SEQ ID NO:286]	deficiência de acil-CoA desidroge-nase (M/SCHAD) de cadeia mé-dia/curta	Oxidação de ácidos graxos
SLC52 A1	55065	0132517	Q9NWF4	MAAPTLGRLVLTHLLVALFGMGSWAAVNGIWVELPVVVKDLPEGWSLPSY LSVVVALGNLGLLVVTLWRQLAPGKGEQVPIQVVQVLSVVGTALLAPLWH HVAPVAGQLHSVAFLTLALVLAMACCTSNVTFLPFLSHLPPPFLRSFFLGQ GLSALLPCVLALVQGVGRLECPPAPTNGTSGPPLDFPERFPASTFFWALT ALLVTSAAAFRGLLLLLPSLPSVTTGGSGPELQLGSPGAEEEEKEEEEALP LQEPPSQAAGTIPGPDPEAHQLFSAHGAFLLGLMAFTSAVTNGVLPSVQS FSCLPYGRLAYHLAVVLGSAANPLACFLAMGVLCRSLAGLVGLSLLGMLF GAYLMALAILSPCPPLVGTTAGVVLVVLSWVLCLCVFSYVKVAASSLLHGG GRPALLAAGVAIQVGSLLGAGAMFPPTSIYHVFQSRKDCVDPCGP [SEQ ID NO:287]	Deficiência do transportador de riboflavina	Oxidação de ácidos graxos
SLC52 A2	79581	0185803	Q9HAB3	MAAPTPARPVLTHLLVALFGMGSWAAVNGIWVELPVVVKELPEGWSLPSY VSVLVALGNLGLLVVTLWRRLAPGKDEQVPIRVVQVLGMVGTALLASLWH HVAPVAGQLHSVAFLALAFVLALACCASNVTFLPFLSHLPPRFLRSFFLGQ GLSALLPCVLALVQGVGRLECPPAPINGTPGPPLDFLERFPASTFFWALTA LLVASAAAFQGLLLLLPPPPSVPTGELGSGLQVGAPGAEEEVEESSPLQEP PSQAAGTTPGPDPKAYQLLSARSACLLGLLAATNALTNGVLPAVQSFSCL PYGRLAYHLAVVLGSAANPLACFLAMGVLCRSLAGLGGLSLLGVFCGGYL MALAVLSPCPPLVGTSAGVVLVVLSWVLCLGVFSYVKVAASSLLHGGGRP ALLAAGVAIQVGSLLGAVAMFPPTSIYHVFHSRKDCADPCDS [SEQ ID	Deficiência do transportador de riboflavina	Oxidação de ácidos graxos

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				NO:288]		
SLC52A3	113278	0101276	K0A6P4, Q9NQ40	MAFLMHLLVCVFGMGSWVTINGLWVELPLLVMELPEGWYLPSYLTVVIQLANIGPLLVTLLHHFRPSCLSEVPIIFTLLGVGTVTCIIFAFLWNMTSWVLDGHHSIAFLVLTFFLALVDCTSSVTFLPFMSRLPTYYLTTFFVGEGLSGLLPALVALAQGSGLTTCVNVTEISDSVPSPVPTRETDIAQGVPRALVSALPGMEAPLSHLESRYLPAHFSPLVFFLLLSIMMACCLVAFFV LQRQPRCWEASVEDLLNDQVTLHSIRPREENDLGPAGTVDSSQGQGYLEEKAAPCCPAHLAFIYTLVAFVNALTNGMLPSVQTYSCLSYGPVAYHLAATLSIVANPLASLVSMFLPNRSLLFLGVLSVLGTCFGGYNMAMAVMSPCPLLQGHWGGEVLIVASWVLFSGCLSYVKVMLGVVLRDLSRSALLWCGAAVQLGSLLGALLMFPLVNVLRLFSSADFCNLHCPA [SEQ ID NO:289]	Deficiência do transportador de riboflavina	Oxidação de ácidos graxos
HADHB	3032	0138029	P55084, F5GZQ3	MTILTYPFKNLPTASKWALRFSIRPLSCSSQLRAAPAVQTKTKKTLAKPNIRNVVVVDGVRTPFLLSGTSYKDLMPHDLARAALTGLLHRTSVPKEVVDYIIFGTVIQEVKTSNVAREAALGAGFSDKTPAHTVTMACISANQAMTTGVGLIASGQCDVIVAGGVELMSDVPIRHSRKMRKLMLDLNKAKSMGQRLSLISKFRFNFLAPELPAVSEFSTSETMGHSADRLAAAFAVSRLEQDEYALRSHSLAKKAQDEGLLSDVVPFKVPGKDTVTKDNGIRPSSLEQMAKLKPAFIKPYGTVTAANSSFLTDGASAMLIMAEEKALAMGYKPKAYLRDFMYVSQDPKDQLLLGPTYATPKVLEKAGLTMNDIDAFEFHEAFSGQILANFKAMDSDWFAENYMGRKTKVGLPPLEKFNNWGGSLSLGHPFGATGCRLVMAAANRLRKEGGQYGLVAACAAGGQGHAMIVEAYPK [SEQ ID NO:290]	Deficiência de proteína trifuncional	Oxidação de ácidos graxos
GYS2	2998	0111713	P54840	MLRGRSLSVTSLGGLPQWEVEELPVEELLLFEVAWEVTNKVGGIYTVIQTKAKTTADEWGENYFLIGPYFEHNMKTQVEQCEPVNDAVRRAVDAMNKHGCQVHFGRWLIEGSPYVVLFDIGYSAWNLDRWKGDLWEACSVGIPYHDREANDMLIFGSLTAWFLKEVTDHADGKYVVAQFHEWQAGIGLILSRARKLPIATIFTTHATLLGRYLCAANIDFYNHLDKFNIDKEAGERQIYHRYCMERASVHCAHVFTTVSEITAIEAEHMLKRKPDVVTPNGLNVKKFSAVHEFQNLHAMYKARIQDFVRGHFYGHLDFDLEKTLFLFIAGRYEFSNKGADIFLESLSRLNFLLRMHKSDITVMVFFIMPAKTNNFNVETLKGQAVRKQLWDVAHSVKEKFGKKLYDALLRGEIPDLNDILDRDDLTIMKRAIFSTQRQSLPPVTTHNMIDDSTDPILS	GSD 0 (glicogênio sintase, isoforma do fígado)	Distúrbio de armazenamento de glicogênio hepático

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				TIRRIGLFNNRTDRVKVILHPEFLSSTSPLLPMDYEEFVRGCHLGVFPSYYE PWGYTPAECTVMGIPSVTTNLSGFGCFMQEHVADPTAYGIYIVDRRFRSP DDSCNQLTKFLYGFCKQSRRQRIIQRNRTERLSDLLDWRYLGRYYQHARH LTLSRAFPDKFHVELTSPPTTEGFKYPRPSSVPPSPSGSQASSPQSSDVE DEVEDERYDEEEEAERDRLNIKSPFSLSHVPHGKKKLHGEYKN [SEQ ID NO:291]		
PYGL	5836	0100504	P06737	MAKPLTDQEKRRQISIRGIVGVENVAELKKSFNRHLHFTLVKDRNVATTRD YYFALAHTVRDHLVGRWIRTQQHYYDKCPKRVYYLSLEFYMGRTLQNTMI NLGLQNACDEAIYQLGLDIEELEEIEEDAGLGNGGLGRLAACFLDSMATLG LAAYGYGIRYEYGIFNQKIRDGWQVEEADDWLRYGNPWEKSRPEFMLPV HFYGKVEHTNTGTKWIDTQVVLALPYDTPVPGYMNNTVNTMRLWSARAP NDFNLRDFNVGDYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEYFVV AATLQDIIRRFKASKFGSTRGAGTVFDAFPDQVAIQLNDTHPALAIPELMRI FVDIEKL	GSD VI (doença de Hers)	Distúrbio de armazenamento de glicogênio hepático
				PWSKAWELTQKTFAYTNHTVLPEALERWPVDLVEKLLPRHLEIIYEINQKH LDRIVALFPKDVDRLRRMSLIEEEGSKRINMAHLCIVGSHAVNGVAKIHSDI VKTKVFKDFSELEPDKFQNKTNGITPRRWLLLCNPGLAELIAEKIGEDYVKD LSQLTKLHSFLGDDVFLRELAKVKQENKLKFSQFLETEYKVKINPSSMFDV QVKRIHEYKRQLLNCLHVITMYNRIKKDPKKLFVPRTVIIGGKAAPGYHMAK MIIKLITSVADVVNNDPMVGSKLKVIFLENYRVSLAEKVIPATDLSEQISTAG TEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENLFIFGMRIDDVAA LDKKGYEAKEYYEALPELKLVIDQIDNGFFSPKQPDLFKDIINMLFYHDRFK VFADYEAYVKCQDKVSQLYMNPKAWNTMVLKNIAASGKFSSDRTIKEYAQ NIWNVEPSDLKISLSNESNKVNGN [SEQ ID NO:292]		
SLC2A2	6514	0163581	P11168, Q6PAU8	MTEDKVTGTLVFTVITAVLGSFQFGYDIGVINAPQQVIISHYRHVLGVPLDD RKAINNYVINSTDELPTISYSMNPKPTPWAEEETVAAAQLITMLWSLSVSSF AVGGMTASFFGGWLGDTLGRIKAMLVANILSLVGALLMGFSKLGPSHILIA GRSISGLYCGLISGLVPMYIGEIAPTALRGALGTFHQLAIVTGILISQIIGLEFIL GNYDLWHILLGLSGVRAILQSLLLFFCPESPRYLYIKLDEEVKAKQSLKRLR GYDDVTKDINEMRKEREEASSEQKVSIIQLFTNSSYRQPILVALMLHVAQQ FSGINGIFYYSTSIFQTAGISKPVYATIGVGAVNMVFTAVSVFLVEKAGRRS	Síndrome de Fanconi-Bickel	Distúrbio de armazenamento de glicogênio hepático

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LFLIGMSGMFVCAIFMSVGLVLLNKFSWMSYVSMIAIFLFVSFFEIGPGPIP WFMVAEFFSQGPRPAALAIAAFSNWTCNFIVALCFQYIADFCGPYVFFLFA GVLLAFTLFTFFKVPETKGKSFEEIAAEFQKKSGSAHRPKAAVEMKFLGAT ETV [SEQ ID NO:293]		
ALG1	56052	0033011	Q9BT22	MAASCLVLLALCLLLPLLLLGGWKRWRRGRAARHVVAVVLGDVGRSPRM QYHALSLAMHGFSVTLLGFCNSKPHDELLQNNRIQIVGLTELQSLAVGPRV FQYGVKVVLQAMYLLWKLMWREPGAYIFLQNPPGLPSIAVCWFVGCLCG SKLVIDWHNYGYSIMGLVHGPNHPLVLLAKWYEKFFGRLSHLNLCVTNAM REDLADNWHIRAVTVYDKPASFFKETPLDLQHRLFMKLGSMHSPFRARSE PEDPVTERSAFTERDAGSGLVTRLRERPALLVSSTSWTEDEDFSILLAALE KFEQLTLDGHNLPSLVCVITGKGPLREYYSRLIHQKHFQHIQVCTPWLEAE DYPLLLGSADLGVCLHTSSSGLDLPMKVVDMFGCCLPVCAVNFKCLHELV KHEENGLVFEDSEELAAQLQMLFSNFPDPAGKLNQFRKNLRESQQLRWD ESWVQTVLPLVMDT [SEQ ID NO:294]	ALG1-CDG	Distúrbio de glicosilação
ALG2	85365	0119523	A0A024R184, Q9H553	MAEEQGRERDSVPKPSVLFLHPDLGVGGAERLVLDAALALQARGCSVKI WTAHYDPGHCFAESRELPVRCAGDWLPRGLGWGGRGAAVCAYVRMVF LALYVLFLADEEFDVVVCDQVSACIPVFRLARRRKKILFYCHFPDLLLTKRD SFLKRLYRAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLY PSLNVTSFDSVVPEKLDDLVPKGKKFLLLSINRYERKKNLTLALEALVQLRG RLTSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQYVTFLR SFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPVIAVNSGGPLESI DHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAFT EQLYRYVTKLLV [SEQ ID NO:295]	Síndrome miastênica associada a ALG2	Distúrbio de glicosilação
ALG3	10195	0214160	Q92685, C9J7S5	MAAGLRKRGRSGSAAQAEGLCKQWLQRAWQERRLLLREPRYTLLVAACL CLAEVGITFWVIHRVAYTEIDWKAYMAEVEGVINGTYDYTQLQGDTGPLVY PAGFVYIFMGLYYATSRGTDIRMAQNIFAVLYLATLLLVFLIYHQTCKVPPFV FFFMCCASYRVHSIFVLRLFNDPVAMVLLFLSINLLLAQRWGWGCCFFSLA VSVKMNVLLFAPGLLFLLLTQFGFRGALPKLGICAGLQVVLGLPFLLENPSG YLSRSFDLGRQFLFHWTVNWRFLPEALFLHRAFHLALLTAHLTL	ALG3-CDG	Distúrbio de glicosilação
				LLLFALCRWHRTGESILSLLRDPSKRKVPPQPLTPNQIVSTLFTSNFIGICFS		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				RSLHYQFYVWYFHTLPYLLWAMPARWLTHLLRLLVLGLIELSWNTYPSTS CSSAALHICHAVILLQLWLGPQPFPKSTQHSKKAH [SEQ ID NO:296]		
ALG6	29929	0088035	Q9Y672	MEKWYLMTVVVLIGLTVRWTVSLNSYSGAGKPPMFGDYEAQRHWQEITF NLPVKQWYFNSSDNNLQYWGLDYPPLTAYHSLLCAYVAKFINPDWIALHT SRGYESQAHKLFMRTTVLIADLLIYIPAVVLYCCCLKEISTKKKIANALCILLY PGLILIDYGHFQYNSVSLGFALWGVLGISCDCDLLGSLAFCLAINYKQMELY HALPFFCFLLGKCFKKGLKGKGFVLLVKLACIVVASFVLCWLPFFTEREQTL QVLRRLFPVDRGLFEDKVANIWCSFNVFLKIKDILPRHIQLIMSFCSTFLSLL PACIKLILQPSSKGFKFTLVSCALSFFLFSFQVHEKSILLVSLPVCLVLSEIPF MSTWFLLVSTFSMLPLLLKDELLMPSVVTTMAFFIACVTSFSIFEKTSEEEL QLKSFSISVRKYLPCFTFLSRIIQYLFLISVITMVLLTLMTVTLDPPQKLPDLF SVLVCFVSCLNFLFFLVYFNIIIMWDSKSGRNQKKIS [SEQ ID NO:297]	ALG6-CDG	Distúrbio de glicosilação
ALG8	79053	0159063	Q9BVK2, A0A024R5K5	MAALTIATGTGNWFSALALGVTLLKCLLIPTYHSTDFEVHRNWLAITHSLPIS QWYYEATSEWTLDYPPFFAWFEYILSHVAKYFDQEMLNVHNLNYSSSRTL LFQRFSVIFMDVLFVYAVRECCKCIDGKKVGKELTEKPKFILSVLLLWNFGL LIVDHIHFQYNGFLFGLMLLSIARLFQKRHMEGAFLFAVLLHFKHIYLYVAPA YGVYLLRSYCFTANKPDGSIRWKSFSFVRVISLGLVVFLVSALSLGPFLALN QLPQVFSRLFPFKRGLCHAYWAPNFWALYNALDKVLSVIGLKLKFLDPNNI PKASMTSGLVQQFQHTVLPSVTPLATLICTLIAILPSIFCLWFKPQGPRGFL RCLTLCALSSFMFGWHVHEKAILLAILPMSLLSVGKAGDASIFLILTTTGHYS LFPLLFTAPELPIKILLMLLFTIYSISSLKTLFRKEKPLFNWMETFYLLGLGPL EVCCEFVFPFTSWKVKYPFIPLLLTSVYCAVGITYAWFKLYVSVLIDSAIGKT KKQ [SEQ ID NO:298]	ALG8-CDG	Distúrbio de glicosilação
ALG9	79796	0086848	Q9H6U8	MASRGARQRLKGSGASSGDTAPAADKLRELLGSREAGGAEHRTELSGNK AGQVWAPEGSTAFKCLLSARLCAALLSNISDCDETFNYWEPTHYLIYGEG FQTWEYSPAYAIRSYAYLLLHAWPAAFHARILQTNKILVFYFLRCLLAFVSCI CELYFYKAVCKKFGLHVSRMMLAFLVLSTGMFCSSSAFLPSSFCMYTTLIA MTGWYMDKTSIAVLGVAAGAILGWPFSAALGLPIAFDLLVMKHRWKSFFH WSLMALILFLVPVVVIDSYYYGKLVIAPLNIVLYNVFTPHGPDLYGT EPWYFYLINGFLNFNVAFALALLVLPLTSLMEYLLQRFHVQNLGHPYWLTL	ALG9-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				APMYIWFIIFFIQPHKEERFLFPVYPLICLCGAVALSALQKCYHFVFQRYRLE HYTVTSNWLALGTVFLFGLLSFSRSVALFRGYHGPLDLYPEFYRIATDPTIH TVPEGRPVNVCVGKEWYRFPSSFLLPDNWQLQFIPSEFRGQLPKPFAEG PLATRIVPTDMNDQNLEEPSRYIDISKCHYLVDLDTMRETPREPKYSSNKE EWISLAYRPFLDASRSSKLLRAFYVPFLSDQYTVYVNYTILKPRKAKQIRKK SGG [SEQ ID NO:299]		
ALG11	440138	0253710	Q2TAA5	MAAGERSWCLCKLLRFFYSLFFPGLIVCGTLCVCLVIVLWGIRLLLQRKKKL VSTSKNGKNQMVIAFFHPYCNAGGGGERVLWCALRALQKKYPEAVYVVY TGDVNVNGQQILEGAFRRFNIRLIHPVQFVFLRKRYLVEDSLYPHFTLLGQ SLGSIFLGWEALMQCVPDVYIDSMGYAFTLPLFKYIGGCQVGSYVHYPTIS TDMLSVVKNQNIGFNNAAFITRNPFLSKVKLIYYYLFAFIYGLVGSCSDVVM VNSSWTLNHILSLWKVGNCTNIVYPPCDVQTFLDIPLHEKKMTPGHLLVSV GQFRPEKNHPLQIRAFAKLLNKKMVESPPSLKLVLIGGCRNKDDELRVNQL RRLSEDLGVQEYVEFKINIPFDELKNYLSEATIGLHTMWNEHFGIGVVECM AAGTIILAHNSGGPKLDIVVPHEGDITGFLAESEEDYAETIAHILSMSAEKRL QIRKSARASVSRFSDQEFEVTFLSSVEKLFK [SEQ ID NO:300]	ALG11-CDG	Distúrbio de glicosilação
ALG12	79087	0182858	A0A024R4V6, Q9BV10	MAGKGSSGRRPLLLGLLVAVATVHLVICPYTKVEESFNLQATHDLLYHWQ DLEQYDHLEFPGVVPRTFLGPVVIAVFSSPAVYVLSLLEMSKFYSQLIVRG VLGLGVIFGLWTLQKEVRRHFGAMVATMFCWVTAMQFHLMFYCTRTLPN VLALPVVLLALAAWLRHEWARFIWLSAFAIIVFRVELCLFLGLLLLALGNRK VSVVRALRHAVPAGILCLGLTVAVDSYFWRQLTWPEGKVLWYNTVLNKSS NWGTSPLLWYFYSALPRGLGCSLLFIPLGLVDRRTHAPTVLALGFMALYSL LPHKELRFIIYAFPMLNITAARGCSYLLNNYKKSWLYKAGSLLVIGHLVVNA AYSATALYVSHFNYPGGVAMQRLHQLVPPQTDVLLHIDVAAAQTGVSRFL QVNSAWRYDKREDVQPGTGMLAYTHILMEAAPGLLALYRDTHRVLASVV GTTGVSLNLTQLPPFNVHLQTKLVLLERLPRPS [SEQ ID NO:301]	ALG12-CDG	Distúrbio de glicosilação
ALG13	79868	0101901	Q9NP73, A0A087WX43, A0A087WT15	MKCVFVTVGTTSFDDLIACVSAPDSLQKIESLGYNRLILQIGRGTVVPEPFS TESFTLDVYRYKDSLKEDIQKADLVISHAGAGSCLETLEKGKPLVVVINEKL MNNHQLELAKQLHKEGHLFYCTCRVLTCPGQAKSIASAPGKCQDSAALTS TAFSGLDFGLLSGYLHKQALVTATHPTCTLLFPSCHAFFPLPLTPTLYKMH	ALG13-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KGWKNYCSQKSLNEASMDEYLGSLGLFRKLTAKDASCLFRAISEQLFCSQ VHHLEIRKACVSYMRENQQTFESYVEGSFEKYLERLGDPKESAGQLEIRA LSLIYNRDFILYRFPGKPPTYVTDNGYEDKILLCYSSSGHYDSVYSKQFQSS AAVCQAVLYEILYKDVFVVDEEELKTAIKLFRSGSKKNRNNAVTGSEDAHT DYKSSNQNRMEEWGACYNAENIPEGYNKGTEETKSPENPSKMPFPYKVL KALDPEIYRNVEFDVWLDSRKELQKSDYMEYAGRQYYLGDKCQVCLESE GRYYNAHIQEVGNENNSVTVFIEELAEKHVVPLANLKPVTQVMSVPAWNA MPSRKGRGYQKMPGGYVPEIVISEMDIKQQKKMFKKIRGKEVYM		
				TMAYGKGDPLLPPRLQHSMHYGHDPPMHYSQTAGNVMSNEHFHPQHPS PRQGRGYGMPRNSSRFINRHNMPGPKVDFYPGPGKRCCQSYDNFSYRS RSFRRSHRQMSCVNKESQYGFTPGNGQMPRGLEETITFYEVEEGDETAY PTLPNHGGPSTMVPATSGYCVGRRGHSSGKQTLNLEEGNGQSENGRYH EEYLYRAEPDYETSGVYSTTASTANLSLQDRKSCSMSPQDTVTSYNYPQK MMGNIAAVAASCANNVPAPVLSNGAAANQAISTTSVSSQNAIQPLFVSPPT HGRPVIASPSYPCHSAIPHAGASLPPPPPPPPPPPPPPPPPPPPPPPPP ALDVGETSNLQPPPPLPPPPYSCDPSGSDLPQDTKVLQYYFNLGLQCYYH SYWHSMVYVPQMQQQLHVENYPVYTEPPLVDQTVPQCYSEVRREDGIQ AEASANDTFPNADSSSVPHGAVYYPVMSDPYGQPPLPGFDSCLPVVPDY SCVPPWHPVGTAYGGSSQIHGAINPGPIGCIAPSPPASHYVPQGM		
				[SEQ ID NO:302]		
ATP6V0A2	23545	0185344	Q9Y487	MGSLFRSETMCLAQLFLQSGTAYECLSALGEKGLVQFRDLNQNVSSFQR KFVGEVKRCEELERILVYLVQEINRADIPLPEGEASPPAPPLKQVLEMQEQ LQKLEVELREVTKNKEKLRKNLLELIEYTHMLRVTKTFVKRNVEFEPTYEEF PSLESDSLLDYSCMQRLGAKLGFVSGLINQGKVEAFEKMLWRVCKGYTIV SYAELDESLEDPETGEVIKWYVFLISFWGEQIGHKVKKICDCYHCHVYPYP NTAEERREIQEGLNTRIQDLYTVLHKTEDYLRQVLCKAAESVYSRVIQVKK MKAIYHMLNMCSFDVTNKCLIAEVWCPEADLQDLRRALEEGSRESGATIP SFMNIIPTKETPPTRIRTNKFTEGFQNIVDAYGVGSYREVNPALFTIITFPFLF AVMFGDFGHGFVMFLFALLLVLNENHPRLNQSQEIMRMFFNGRYILLLMG LFSVYTGLIYNDCFSKSVNLFGSGWNVSAMYSSSHPPAEHKKMVLWNDS VVRHNSILQLDPSIPGVFRGPYPLGIDPIWNLATNRLTFLNSFKMKMSVILGI	Cutis laxa associada a ATP6V0A2	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				IHMTFGVILGIFNHLHFRKKFNIYLVSIPELLFMLCIFGYLIFMIFYKWLVFSAE TSRVAPSILIEFINMFLFPASKTSGLYTGQEYVQRVLLVVTALSVPVLFLGKP LFLLWLHNGRSCFGVNRSGYTLIRKDSEEEVSLLGSQDIEEGNHQVEDGC REMACEEFNFGEILMTQVIHSIEYCLGCISNTASYLRLWALSLAHAQLSDVL WAMLMRVGLRVDTTYGVLLLLPVIALFAVLTIFILLIMEGLSAFLHAIRLHWV EFQNKFYVGAGTKFVPF		
				SFSLLSSKFNNDDSVA [SEQ ID NO:303]		
B3GL CT	145173	0187676	Q6Y288	MRPPACWWLLAPPALLALLTCSLAFGLASEDTKKEVKQSQDLEKSGISRK NDIDLKGIVFVIQSQSNSFHAKRAEQLKKSILKQAADLTQELPSVLLLHQLA KQEGAWTILPLLPHFSVTYSRNSSWIFFCEEETRIQIPKLLETLRRYDPSKE WFLGKALHDEEATIIHHYAFSENPTVFKYPDFAAGWALSIPLVNKLTKRLKS ESLKSDFTIDLKHEIALYIWDKGGGPPLTPVPEF	B3GLCT-CDG	Distúrbio de glicosilação
				CTNDVDFYCATTFHSFLPLCRKPVKKKDIFVAVKTCKKFHGDRIPIVKQTW ESQASLIEYYSDYTENSIPTVDLGIPNTDRGHCGKTFAILERFLNRSQDKTA WLVIVDDDTLISISRLQHLLSCYDSGEPVFLGERYGYGLGTGGYSYITGGG GMVFSREAVRRLLASKCRCYSNDAPDDMVLGMCFSGLGIPVTHSPLFHQ ARPVDYPKDYLSHQVPISFHKHWNIDPVKVYFTWLAPSDEDKARQETQKG FREEL		
				[SEQ ID NO:304]		
CHST 14	113189	0169105	Q8NCH0	MFPRPLTPLAAPNGAEPLGRALRRAPLGRARAGLGGPPLLLPSMLMFAVI VASSGLLLMIERGILAEMKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDA DLQVRQDVRNRTLRAVCGQPGMPRDPWDLPVGQRRTLLRHILVSDRYRF LYCYVPKVACSNWKRVMKVLAGVLDSVDVRLKMDHRSDLVFLADLRPEEI RYRLQHYFKFLFVREPLERLLSAYRNKFGEIREYQQRYGAEIVRRYRAGA GPSPAGDDVTFPEFLRYLVDEDPERMNEHWMPVYHLCQPCAVHYDFVG SYERLEADANQVLEWVRAPPHVRFPARQAWYRPASPESLHYHLCSAPRA LLQDVLPKYILDFSLFAYPLPNVTKEACQQ		
				[SEQ ID NO:305]		
COG1	9382	0166685	Q8WTW3	MATAATSPALKRLDLRDPAALFETHGAEEIRGLERQVRAEIEHKKEELRQM	COG1-CDG	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				VGERYRDLIEAADTIGQMRRCAVGLVDAVKATDQYCARLRQAGSAAPRPP RAQQPQQPSQEKFYSMAAQIKLLLEIPEKIWSSMEASQCLHATQLYLLCCH LHSLLQLDSSSSRYSPVLSRFPILIRQVAAASHFRSTILHESKMLLKCQGVS DQAVAEALCSIMLLEESSPRQALTDFLLARKATIQKLLNQPHHGAGIKAQIC SLVELLATTLKQAHALFYTLPEGLLPDPALPCGLLFSTLETITGQHPAGKGT GVLQEEMKLCSWFKHLPASIVEFQPTLRTLAHPISQEYLKDTLQKWIHMCN EDIKNGITNLLMYVKSMKGLAGIRDAMWELLTNESTNHSWDVLCRRLLEK PLLFWEDMMQQLFLDRLQTLTKEGFDSISSSSKELLVSALQELESSTSNSP SNKHIHFEYNMSLFLWSESPNDLPSDAAWVSVANRGQFASSGLSMKAQAI SPCVQNFCSALDSKLKVKLDDLLAYLPSDD		
				SSLPKDVSPTQAKSSAFDRYADAGTVQEMLRTQSVACIKHIVDCIRAELQS IEEGVQGQQDALNSAKLHSVLFMARLCQSLGELCPHLKQCILGKSESSEK PAREFRALRKQGKVKTQEIIPTQAKWQEVKEVLLQQSVMGYQVWSSAVV KVLIHGFTQSLLLDDAGSVLATATSWDELEIQEEAESGSSVTSKIRLPAQPS WYVQSFLFSLCQEINRVGGHALPKVTLQEMLKSCMVQVVAAYEKLSEEKQ IKKEGAFPVTQNRALQLLYDLRYLNIVLTAKGDEVKSGRSKPDSRIEK		
				VTDHLEALIDPFDLDVFTPHLNSNLHRLVQRTSVLFGLVTGTENQLAPRSS TFNSQEPHNILPLASSQIRFGLLPLSMTSTRKAKSTRNIETKAQVVPPARST AGDPTVPGSLFRQLVSEEDNTSAPSLFKLGWLSSMTK [SEQ ID NO:306]		
COG2	22796	0135775	Q14746, B1ALW7	MEKSRMNLPKGPDTLCFDKDEFMKEDFDVDHFVSDCRKRVQLEELRDDL ELYYKLLKTAMVELINKDYADFVNLSTNLVGMDKALNQLSVPLGQLREEVL SLRSSVSEGIRAVDERMSKQEDIRKKKMCVLRLIQVIRSVEKIEKILNSQSS KETSALEASSPLLTGQILERIATEFNQLQFHAVQSKGMPLLDKVRPRIAGIT AMLQQSLEGLLLEGLQTSDVDIIRHCLRTYATIDKTRDAEALVGQVLVKPYI DEVIIEQFVESHPNGLQVMYNKLLEFVPHHCRLLREVTGGAISSEKGNTVP GYDFLVNSVWPQIVQGLEEKLPSLFNPGNPDAFHEKYTISMDFVRRLERQ CGSQASVKRLRAHPAYHSFNKKWNLPVYFQIRFREIAGSLEAALTDVLEDA PAESPYCLLASHRTWSSLRRCWSDEMFLPLLVHRLWRLTLQILARYSVFV NELSLRPISNESPKEIKKPLVTGSKEPSITQGNTEDQGSGPSETKPVVSISR TQLVYVVADLDKLQEQLPELLEIIKPKLEMIGFKNFSSISAALEDSQSSFSAC VPSLSSKIIQDLSDSCFGFLKSALEVPRLYRRTNKEVPTTASSYVDSALKPL	COG2-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FQLQSGHKDKLKQAIIQQWLEGTLSESTHKYYETVSDVLNSVKKMEESLK RLKQARKTTPANPVGPSGGMSDDDKIRLQLALDVEYLGEQIQKLGLQASDI KSFSALAELVAAAKDQATAEQP [SEQ ID NO:307]		
COG4	25839	0103051	A0A0A0MS45, Q8N8L9, Q9H9E3, J3KNI1	MADLDSPPKLSGVQQPSEGVGGGRCSEISAELIRSLTELQELEAVYERLC GEEKVVERELDALLEQQNTIESKMVTLHRMGPNLQLIEGDAKQLAGMITFT CNLAENVSSKVRQLDLAKNRLYQAIQRADDILDLKFCMDGVQTALRSEDY EQAAAHTHRYLCLDKSVIELSRQGKEGSMIDANLKLLQEAEQRLKAIVAEK FAIATKEGDLPQVERFFKIFPLLGLHEEGLRKFSEYLCKQVASKAEENLLMV LGTDMSDRRAAVIFADTLTLLFEGIARIVETHQPIVETYYGPGRLYTLIKYLQ VECDRQVEKVVDKFIKQRDYHQQFRHVQNNLMRNSTTEKIEPRELDPILTE VTLMNARSELYLRFLKKRISSDFEVGDSMASEEVKQEHQKCLDKLLNNCLL SCTMQELIGLYVTMEEYFMRETVNKAVALDTYEKGQLTSSMVDDVFYIVK KCIGRALSSSSIDCLCAMINLATTELESDFRDVLCNKLRMGFPATTFQDIQR GVTSAVNIMHSSLQQGKFDTKGIESTDEAKMSFLVTLNNVEVCSENISTLK KTLESDCTKLFSQGIGGEQAQAKFDSCLSDLAAVSNKFRDLLQEGLTELNS TAIKPQVQPWINSFFSVSHNIEEEEFNDYEANDPWVQQFILNLEQQMAEFK ASLSPVIYDSLTGLMTSLVAVELEKVVLKSTFNRLGGLQFDKELRSLIAYLT TVTTWTIRDKFARLSQMATILNLERVTEILDYWGPNSGPLTWRLTPAEVRQ VLALRIDFRSEDIKRLRL [SEQ ID NO:308]	COG4-CDG	Distúrbio de glicosilação
COG5	10466	0164597,0284369	Q9UP83	MGWVGGRRRDSASPPGRSRSAADDINPAPANMEGGGGSVAVAGLGAR GSGAAAATVRELLQDGCYSDFLNEDFDVKTYTSQSIHQAVIAEQLAKLAQ GISQLDRELHLQVVARHEDLLAQATGIESLEGVLQMMQTRIGALQGAVDRI KAKIVEPYNKIVARTAQLARLQVACDLLRRIIRILNLSKRLQGQLQGGSREIT KAAQSLNELDYLSQGIDLSGIEVIENDLLFIARARLEVENQAKRLLEQGLET QNPTQVGTALQVFYNLGTLKDTITSVVDGYCATLEENINSALDIKVLTQPSQ SAVRGGPGRSTMPTPGNTAALRASFWTNMEKLMDHIYAVCGQVQHLQK VLAKKRDPVSHICFIEEIVKDGQPEIFYTFWNSVTQALSSQFHMATNSSMF LKQAFEGEYPKLLRLYNDLWKRLQQYSQHIQGNFNASGTTDLYVDLQHM EDDAQDIFIPKKPDYDPEKALKDSLQPYEAAYLSKSLSRLFDPINLVFPPGG RNPPSSDELDGIIKTIASELNVAAVDTNLTLAVSKNVAKTIQLYSVKSEQLLS TQGDASQVIGPLTEGQRRNVAVVNSLYKLHQSVTKAIHALMENAVQPLLT	COG5-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SVGDAIEAIIITMHQEDFSGSLSSSGKPDVPCSLYMKELQGFIARVMSDYFK HFECLDFVFDNTEAIAQRAVELFIRHASLIRPLGEGGKMRLAADFAQMELA VGPFCRRVSDLGKSYRMLRSFRPLLFQASEHVASSPALGDVIPFSIIIQFLF TRAPAELKSPFQRAEWSHTRFSQWLDDHPSEKDRLLLIRGALEAYVQSVR SREGKEFAPVYPIMVQLLQKAMSALQ [SEQ ID NO:309]		
COG6	57511	0133103	A0A140VJG7, Q9Y2V7, A0A024RDW5	MAEGSGEVVAVSATGAANGLNNGAGGTSATTCNPLSRKLHKILETRLDND KEMLEALKALSTFFVENSLRTRRNLRGDIERKSLAINEEFVSIFKEVKEELE SISEDVQAMSNCCQDMTSRLQAAKEQTQDLIVKTTKLQSESQKLEIRAQV ADAFLSKFQLTSDEMSLLRGTREGPITEDFFKALGRVKQIHNDVKVLLRTN QQTAGLEIMEQMALLQETAYERLYRWAQSECRTLTQESCDVSPVLTQAM EALQDRPVLYKYTLDEFGTARRSTVVRGFIDALTRGGPGGTPRPIEMHSH DPLRYVGDMLAWLHQATASEKEHLEALLKHVTTQGVEENIQEVVGHITEG VCRPLKVRIEQVIVAEPGAVLLYKISNLLKFYHHTISGIVGNSATALLTTIEEM HLLSKKIFFNSLSLHASKLMDKVELPPPDLGPSSALNQTLMLLREVLASHD SSVVPLDARQADFVQVLSCVLDPLLQMCTVSASNLGTADMATFMVNSLY MMKTTLALFEFTDRRLEMLQFQIEAHLDTLINEQASYVLTRVGLSYIYNTVQ QHKPEQGSLANMPNLDSVTLKAAMVQFDRYLSAPDNLLIPQLNFLLSATVK EQIVKQSTELVCRAYGEVYAAVMNPINEYKDPENILHRSPQQVQTLLS [SEQ ID NO:310]	COG6-CDG	Distúrbio de glicosilação
COG7	91949	0168434	A0A0S2Z652, P83436	MDFSKFLADDFDVKEWINAAFRAGSKEAASGKADGHAATLVMKLQLFIQE VNHAVEETSHQALQNMPKVLRDVEALKQEASFLKEQMILVKEDIKKFEQDT SQSMQVLVEIDQVKSRMQLAAESLQEADKWSTLSADIEETFKTQDIAVISA KLTGMQNSLMMLVDTPDYSEKCVHLEALKNRLEALASPQIVAAFTSQAVD QSKVFVKVFTEIDRMPQLLAYYYKCHKVQLLAAWQELCQSDLSLDRQLTG LYDALLGAWHTQIQWATQVFQKPHEVVMVLLIQTLGALMPSLPSCLSNGV ERAGPEQELTRLLEFYDATAHFAKGLEMALLPHLHEHNLVKVTELVDAVYD PYKPYQLKYGDMEESNLLIQMSAVPLEHGEVIDCVQELSHSVNKLFGLASA AVDRCVRFTNGLGTCGLLSALKSLFAKYVSDFTSTLQSIRKKCKLDHIPPN SLFQEDWTAFQNSIRIIATCGELLRHCGDFEQQLANRILSTAGKYLSDSCS PRSLAGFQESILTDKKNSAKNPWQEYNYLQKDNPAEYASLMEILYTLKEKG SSNHNLLAAPRAALTRLNQQAHQLAFDSVFLRIKQQLLLISKMDSWNTAGI	COG7-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GETLTDELPAFSLTPLEYISNIGQYIMSLPLNLEPFVTQEDSALELALHAGKL PFPPEQGDELPELDNMADNWLGSIARATMQTYCDAILQIPELSPHSAKQLA TDIDYLINVMDALGLQPSRTLQHIVTLLKTRPEDYRQVSKGLPRRLATTVAT MRSVNY [SEQ ID NO:311]		
COG8	84342	0272617	A0A024R6Z6, Q96MW5	MATAATIPSVATATAAALGEVEDEGLLASLFRDRFPEAQWRERPDVGRYL RELSGSGLERLRREPERLAEERAQLLQQTRDLAFANYKTFIRGAECTERIH RLFGDVEASLGRLLDRLPSFQQSCRNFVKEAEEISSNRRMNSLTLNRHTEI LEILEIPQLMDTCVRNSYYEEALELAAYVRRLERKYSSIPVIQGIVNEVRQS MQLMLSQLIQQLRTNIQLPACLRVIGYLRRMDVFTEAELRVKFLQARDAWL RSILTAIPNDDPYFHITKTIEASRVHLFDIITQYRAIFSDEDPLLPPAMGEHTV NESAIFHGWVLQKVSQFLQVLETDLYRGIGGHLDSLLGQCMYFGLSFSRV GADFRGQLAPVFQRVAISTFQKAIQETVEKFQEEMNSYMLISAPAILGTSN MPAAVPATQPGTLQPPMVLLDFPPLACFLNNILVAFNDLRLCCPVALAQDV TGALEDALAKVTKIILAFHRAEEAAFSSGEQELFVQFCTVFLEDLVPYLNRC LQVLFPPAQIAQTLGIPPTQLSKYGNLGHVNIGAIQEPLAFILPKRETLFTLD DQALGPELTAPAPEPPAEEPRLEPAGPACPEGGRAETQAEPPSVGP [SEQ ID NO:312]	COG8-CDG	Distúrbio de glicosilação
DOLK	22845	0175283	A0A0S2Z597, Q9UPQ8	DRLLQQGSAVFQFRMSANSGLLPASMVMPLLGLVMKERCQTAGNPFFER FGIVVAATGMAVALFSSVLALGITRPVPTNTCVILGLAGGVIIYIMKHSLSVG EVIEVLEVLLIFVYLNMILLYLLPRCFTPGEALLVLGGISFVLNQLIKRSLTLVE SQGDPVDFFLLVVVVGMVLMGIFFSTLFVFMDSGTWASSIFFHLMTCVLSL GVVLPWLHRLIRRNPLLWLLQFLFQTDTRIYLLAYWSLLATLACLVVLYQNA KRSSSESKKHQAPTIARKYFHLIVVATYIPGIIFDRPLLYVAATVCLAVFIFLE YVRYFRIKPLGHTLRSFLSLFLDERDSGPLILTHIYLLLGMSLPIWLIPRPCT QKGSLGGARALVPYAGVLAVGVGDTVASIFGSTMGEIRWPGTKKTFEGT MTSIFAQIISVALILIFDSGVDLNYSYAWILGSISTVSLLEAYTTQIDNLLLPLY LLILLMA [SEQ ID NO:313]	DOLK-CDG	Distúrbio de glicosilação
DHDDS	79947	0117682	Q86SQ9	MSWIKEGELSLWERFCANIIKAGPMPKHIAFIMDGNRRYAKKCQVERQEG HSQGFNKLAETLRWCLNLGILEVTVYAFSIENFKRSKSEVDGLMDLARQKF SRLMEEKEKLQKHGVCIRVLGDLHLLPLDLQELIAQAVQATKNYNKCFLNV	DHDDS-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				CFAYTSRHEISNAVREMAWGVEQGLLDPSDISESLLDKCLYTNRSPHPDILI RTSGEVRLSDFLLWQTSHSCLVFQPVLWPEYTFWNLFEAILQFQMNHSVL QKARDMYAEERKRQQLERDQATVTEQLLREGLQASGDAQLRRTRLHKLS ARREERVQGFLQALELKRADWLARLGTASA [SEQ ID NO:314]		
DPAG T1	1798	0172269	A0A024R3H8, Q9H3H5	MWAFSELPMPLLINLIVSLLGFVATVTLIPAFRGHFIAARLCGQDLNKTSRQ QIPESQGVISGAVFLIILFCFIPFPFLNCFVKEQCKAFPHHEFVALIGALLAIC CMIFLGFADDVLNLRWRHKLLLPTAASLPLLMVYFTNFGNTTIVVPKPFRPI LGLHLDLGILYYVYMGLLAVFCTNAINILAGINGLEAGQSLVISASIIVFNLVEL EGDCRDDHVFSLYFMIPFFFTTLGLLYHNWYPSRVFVGDTFCYFAGMTFA VVGILGHFSKTMLLFFMPQVFNFLYSLPQLLHIIPCPRHRIPRLNIKTGKLEM SYSKFKTKSLSFLGTFILKVAESLQLVTVHQSETEDGEFTECNNMTLINLLL KVLGPIHERNLTLLLLLLQILGSAITFSIRYQLVRLFYDV [SEQ ID NO:315]	DPAGT1-CDG	Distúrbio de glicosilação
DPM1	8813	0000419	O60762, Q5QPK2, A0A0S2Z4Y5	MASLEVSRSPRRSRRELEVRSPRQNKYSVLLPTYNERENLPLIVWLLVKSF SESGINYEIIIIDDGSPDGTRDVAEQLEKIYGSDRILLRPREKKLGLGTAYIHG MKHATGNYIIIMDADLSHHPKFIPEFIRKQKEGNFDIVSGTRYKGNGGVYG WDLKRKIISRGANFLTQILLRPGASDLTGSFRLYRKEVLEKLIEKCVSKGYV FQMEMIVRARQLNYTIGEVPISFVDRVYGESK LGGNEIVSFLKGLLTLFATT [SEQ ID NO:316]	DPM1-CDG	Distúrbio de glicosilação
DPM2	8818	0136908	O94777	MATGTDQVVGLGLVAVSLIIFTYYTAWVILLPFIDSQHVIHKYFLPRAYAVAI PLAAGLLLLLFVGLFISYVMLKTKRVTKKAQ [SEQ ID NO:317]	DPM2-CDG	Distúrbio de glicosilação
DPM3	54344	0179085	A0A140VJI4, Q9P2X0, Q86TM7	MTKLAQWLWGLAILGSTWVALTTGALGLELPLSCQEVLWPLPAYLLVSAG CYALGTVGYRVATFHDCEDAARELQSQIQEARADLARRGLRF [SEQ ID NO:318]	DPM3-CDG	Distúrbio de glicosilação
G6PC 3	92579	0141349	Q9BUM1	MESTLGAGIVIAEALQNQLAWLENVWLWITFLGDPKILFLFYFPAAYYASRR VGIAVLWISLITEWLNLIFKWFLFGDRPFWWVHESGYYSQAPAQVHQFPS SCETGPGSPSGHCMITGAALWPIMTALSSQVATRARSRWVRVMPSLAYC TFLLAVGLSRIFILAHFPHQVLAGLITGAVLGWLMTPRVPMERELSFYGLTA LALMLGTSLIYWTLFTLGLDLSWSISLAFKWCERPEWIHVDSRPFASLSRD	Neutropenia congênita	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SGAALGLGIALHSPCYAQVRRAQLGNGQKIACLVLAMGLLGPLDWLGHPP QISLFYIFNFLKYTLWPCLVLALVPWAVHMFSAQEAPPIHSS [SEQ ID NO:319]		
GFPT 1	2673	0198380	Q06210	MCGIFAYLNYHVPRTRREILETLIKGLQRLEYRGYDSAGVGFDGGNDKDW EANACKIQLIKKKGKVKALDEEVHKQQDMDLDIEFDVHLGIAHTRWATHGE PSPVNSHPQRSDKNNEFIVIHNGIITNYKDLKKFLESKGYDFESETDTETIAK LVKYMYDNRESQDTSFTTLVERVIQQLEGAFALVFKSVHFPGQAVGTRRG SPLLIGVRSEHKLSTDHIPILYRTARTQIGSKFTRWGSQGERGKDKKGSCN LSRVDSTTCLFPVEEKAVEYYFASDASAVIEHTNRVIFLEDDDVAAVVDGR LSIHRIKRTAGDHPGRAVQTLQMELQQIMKGNFSSFMQKEIFEQPESVVNT MRGRVNFDDYTVNLGGLKDHIKEIQRCRRLILIACGTSYHAGVATRQVLEE LTELPVMVELASDFLDRNTPVFRDDVCFFLSQSGETADTLMGLRYCKERG ALTVGITNTVGSSISRETDCGVHINAGPEIGVASTKAYTSQFVSLVMFALM MCDDRISMQERRKEIMLGLKRLPDLIKEVLSMDDEIQKLATELYHQKSVLIM GRGYHYATCLEGALKIKEITYMHSEGILAGELKHGPLALVDKLMPVIMIIMR DHTYAKCQNALQQVVARQGRPVVICDKEDTETIKNTKRTIKVPHSVDCLQ GILSVIPLQLLAFHLAVLRGYDVDFPRNLAKSVTVE [SEQ ID NO:320]	Síndrome miastênica congênita	Distúrbio de glicosilação
GMPP A	29926	0144591	A0A024R482, Q96IJ6	MLKAVILIGGPQKGTRFRPLSFEVPKPLFPVAGVPMIQHHIEACAQVPGMQ EILLIGFYQPDEPLTQFLEAAQQEFNLPVRYLQEFAPLGTGGGLYHFRDQIL AGSPEAFFVLNADVCSDFPLSAMLEAHRRQRHPFLLLGTTANRTQSLNYG CIVENPQTHEVLHYVEKPSTFISDIINCGIYLFSPEALKPLRDVFQRNQQDG QLEDSPGLWPGAGTIRLEQDVFSALAGQGQIYVHL TDGIWSQIKSAGSALYASRLYLSRYQDTHPERLAKHTPGGPWIRGNVYIHP TAKVAPSAVLGPNVSIGKGVTVGEGVRLRESIVLHGATLQEHTCVLHSIVG WGSTVGRWARVEGTPSDPNPNDPRARMDSESLFKDGKLLPAITILGCRV RIPAEVLILNSIVLPHKELSRSFTNQIIL [SEQ ID NO:321]	GMPPA-CDG	Distúrbio de glicosilação
GMPP B	29925	0173540	Q9Y5P6	MKALILVGGYGTRLRPLTLSTPKPLVDFCNKPILLHQVEALAAAGVDHVILA VSYMSQVLEKEMKAQEQRLGIRISMSHEEEPLGTAGPLALARDLLSETAD PFFVLNSDVICDFPFQAMVQFHRHHGQEGSILVTKVEEPSKYGVVVCEAD TGRIHRFVEKPQVFVSNKINAGMYILSPAVLQRIQLQPTSIEKEVFPIMAKE	Distrofia muscular congênita, síndrome miastênica congênita e distro-	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GQLYAMELQGFWMDIGQPKDFLTGMCLFLQSLRQKQPERLCSGPGIVGN VLVDPSARIGQNCSIGPNVSLGPGVVVEDGVCIRRCTVLRDARIRSHSWLE SCIVGWRCRVGQWVRMENVTVLGEDVIVNDELYLNGASVLPHKSIGESVP EPRIIM [SEQ ID NO:322]	glicanopatia	
MAGT 1	84061	0102158	A0A087WU53, Q9H0U3	MAARWRFWCVSVTMVVALLIVCDVPSASAQRKKEMVLSEKVSQLMEWTN KRPVIRMNGDKFRRLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQIL ANSWRYSSAFTNRIFFAMVDFDEGSDVFQMLNMNSAPTFINFPAKGKPKR GDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNYAGPLMLGLLLAVIGGL VYLRRSNMEFLFNKTGWAFAALCFVLAMTSGQMWNHIRGPPYAHKNPHT GHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSDMDIGKRKIM CVAGIGLVVLFFSWMLSIFRSKYHGYPYSFLMS [SEQ ID NO:323]	MAGT1-CDG; Imunodeficiência ligada ao X com defeito de magné-sio, infecção pelo vírus Epstein-Barr e síndrome de neoplasia (XMEN)	Distúrbio de glicosilação
MAN1 B1	11253	0177239	Q9UKM7	MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVVMYPPPPPPPHRD FISVTLSFGENYDNSKSWRRRSCWRKWKQLSRLQRNMILFLLAFLLFCGL LFYINLADHWKALAFRLEEEQKMRPEIAGLKPANPPVLPAPQKADTDPENL PEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEATKRQEAPVDPRPEG DPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQGTPVHLNYR QKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTM WILGLRKEFEEARKWVSKKLHFEKDVDVNLFESTIRILGGLLSAYHLSGDSL FLRKAEDFGNRLMPAFRTPSKIPYSDVNIGTGVAHPPRWTSDSTVAEVTSI QLEFRELSRLTGDKKFQEAVEKVTQHIHGLSGKKDGLVPMFINTHSGLFTH LGVFTLGARADSYYEYLLKQWIQGGKQETQLLEDYVEAIEGVRTHLLRHSE PSKLTFVGELAHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQELM ETCYQMNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLLRPETVE SLFYLYRVTGDRKYQDWGWEILQSFSRFTRVPSGGYSSINNVQDPQKPE PRDKMESFFLGETLKYLFLLFSDDPNLLSLDAYVFNTEAHPLPIWTPA [SEQ ID NO:324]	MAN1B1-CDG	Distúrbio de glicosilação
MGAT 2	4247	0168282	Q10469	MRFRIYKRKVLILTLVVAACGFVLWSSNGRQRKNEALAPPLLDAEPARGA GGRGGDHPSVAVGIRRVSNVSAASLVPAVPQPEADNLTLRYRSLVYQLNF	MGAT2-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DQTLRNVDKAGTWAPRELVLVVQVHNRPEYLRLLLDSLRKAQGIDNVLVIF SHDFWSTEINQLIAGVNFCPVLQVFFPFSIQLYPNEFPGSDPRDCPRDLPK NAALKLGCINAEYPDSFGHYREAKFSQTKHHWWWKLHFVWERVKILRDY AGLILFLEEDHYLAPDFYHVFKKMWKLKQQECPECDVLSLGTYSASRSF YGMADKVDVKTWKSTEHNMGLALTRNAYQKLIECTDTFCTYDDYNWDWT LQYLTVSCLPKFWKVLVPQIPRIFHAGDCGMHHKKTCRPSTQSAQIESLLN NNKQYMFPETLTISEKFTVVAISPPRKNGGWGDIRDHELCKSYRRLQ [SEQ ID NO:325]		
MOGS	7841	0115275	Q13724, Q58F09	MARGERRRRAVPAEGVRTAERAARGGPGRRDGRGGGPRSTAGGVALA VVVLSLALGMSGRWVLAWYRARRAVTLHSAPPVLPADSSSPAVAPDLFW GTYRPHVYFGMKTRSPKPLLTGLMWAQQGTTPGTPKLRHTCEQGDGVG PYGWEFHDGLSFGRQHIQDGALRLTTEFVKRPGGQHGGDWSWRVTVEP QDSGTSALPLVSLFFYVVTDGKEVLLPEVGAKGQLKFISGHTSELGDFRFT LLPPTSPGDTAPKYGSYNVFWTSNPGLPLLTEMVKSRLNSWFQHRPPGA PPERYLGLPGSLKWEDRGPSGQGQGQFLIQQVTLKIPISIEFVFESGSAQA GGNQALPRLAGSLLTQALESHAEGFRERFEKTFQLKEKGLSSGEQVLGQA ALSGLLGGIGYFYGQGLVLPDIGVEGSEQKVDPALFPPVPLFTAVPSRSFF PRGFLWDEGFHQLVVQRWDPSLTREALGHWLGLLNADGWIGREQILGDE ARARVPPEFLVQRAVHANPPTLLLPVAHMLEVGDPDDLAFLRKALPRLHA WFSWLHQSQAGPLPLSYRWRGRDPALPTLLNPKTLPSGLDDYPRASHPS VTERHLDLRCWVALGARVLTRLAEHLGEAEVAAELGPLAASLEAAESLDEL HWAPELGVFADFGNHTKAVQLKPRPPQGLVRVVGRPQPQLQYVDALGYV SLFPLLLRLLDPTSSRLGPLLDILADSRHLWSPFGLRSLAASSSFYGQRNS EHDPPYWRGAVWLNVNYLALGALHHYGHLEGPHQARAAKLHGELRANV VGNVWRQYQATGFLWEQYSDRDGRGMGCRPFHGWTSLVLLAMAEDY [SEQ ID NO:326]	MOGS-CDG	Distúrbio de glicosilação
MPDU 1	9526	0129255	J3QW43, O75352, A0A0S2Z4W8, B4DLH7	MAAEADGPLKRLLVPILLPEKCYDQLFVQWDLLHVPCLKILLSKGLGLGIVA GSLLVKLPQVFKILGAKSAEGLSLQSVMLELVALTGTMVYSITNNFPFSSW GEALFLMLQTITICFLVMHYRGQTVKGVAFLACYGLVLLVLLSPLTPLTVVT LLQASNVPAVVVGRLLQAATNYHNGHTGQLSAITVFLLFGGSLARIFTSIQE TGDPLMAGTFVVSSLCNGLIAAQLLFYWNAKPPHKQKKAQ	MPDU1-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				[SEQ ID NO:327]		
MPI	4351	0178802	H3BPP3, Q8NHZ6, B4DW50, F5GX71, P34949, H3BPB8	MAAPRVFPLSCAVQQYAWGKMGSNSEVARLLASSDPLAQIAEDKPYAELWMGTHPRGDAKILDNRISQKTLSQWIAENQDSLGSKVKDTFNGNLPFLFKVLSVETPLSIQAHPNKELAEKLHLQAPQHYPDANHKPEMAIALTPFQGLCGFRPVEEIVTFLKKVPEFQFLIGDEAATHLKQTMSHDSQAVASSLQSCFSHLMKSEKKVVVEQLNLLVKRISQQAAAGNNMEDIFGELLLQLHQQYPGDIGCFAIYFLNLLTLKPGEAMFLEANVPHAYLKGDCVECMACSDNTVRAGLTPKFIDVPTLCEMLSYTPSSSKDRLFLPTRSQEDPYLSIYDPPVPDFTIMKTEVPGSVTEYKVLALDSASILLMVQGTVIASTPTTQTPIPLQRGGVLFIGANESVSLKLTEPKDLLIFRACCLL [SEQ ID NO:328]	MPI-CDG	Distúrbio de glicosilação
NGLY1	55768	0151092	Q96IV0	MAAAALGSSSGSASPAVAELCQNTPETFLEASKLLLTYADNILRNPNDEKYRSIRIGNTAFSTRLLPVRGAVECLFEMGFEEGETHLIFPKKASVEQLQKIRDLIAIERSSRLDGSNKSHKVKSSQQPAASTQLPTTPSSNPSGLNQHTRNRQGQSSDPPSASTVAADSAILEVLQSNIQHVLVYENPALQEKALACIPVQELKRKSQEKLSRARKLDKGINISDEDFLLLELLHWFKEEFFHWVNNVLCSKCGGQTRSRDRSLLPSDDELKWGAKEVEDHYCDACQFSNRFPRYNNPEKLLETRCGRCGEWANCFTLCCRAVGFEARYVWDYTDHVWTEVYSPSQQRWLHCDACEDVCDKPLLYEIGWGKKLSYVIAFSKDEVVDVTWRYSCKHEEVIARRTKVKEALLRDTINGLNKQRQLFLSENRRKELLQRIIVELVEFISPKTPKPGELGGRISGSVAWRVARGEMGLQRKETLFIPCENEKISKQLHLCYNIVKDRYVRVSNNNQTISGWENGVWKMESIFRKVETDWHMVYLARKEGSSFAYISWKFECGSVGLKVDSISIRTSSQTFQTGTVEWKLRSDTAQVELTGDNSLHSYADFSGATEVILEAELSRGDGDVAWQHTQLFRQSLNDHEENCLEIIIKFSDL [SEQ ID NO:329]	NGLY1-CDG	Distúrbio de glicosilação
PGM1	5236	0079739	B7Z6C2, P36871, B4DDQ8	MVKIVTVKTQAYQDQKPGTSGLRKRVKVFQSSANYAENFIQSIISTVEPAQRQEATLVVGGDGRFYMKEAIQLIARIAAANGIGRLVIGQNGILSTPAVSCIIRKIKAIGGIILTASHNPGGPNGDFGIKFNISNGGPAPEAITDKIFQISKTIEEYAVCPDLKVDLGVLGKQQFDLENKFKPFTVEIVDSVEAYATMLRSIFDFSALKELLSGPNRLKIRIDAMHGVVGPYVKKILCEELGAPANSAVNCVPLEDFGGHHP	PGM1-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DPNLTYAADLVETMKSGEHDFGAAFDGDGDRNMILGKHGFFVNPSDSVA VIAANIFSIPYFQQTGVRGFARSMPTSGALDRVASATKIALYETPTGWKFF GNLMDASKLSLCGEESFGTGSDHIREKDGLWAVLAWLSILATRKQSVEDIL KDHWQKYGRNFFTRYDYEEVEAEGANKMMKDLEALMFDRSFVGKQFSA NDKVYTVEKADNFEYSDPVDGSISRNQGLRLIFTDGSRIVFRLSGTGSAGA TIRLYIDSYEKDVAKINQDPQVMLAPLISIALKVSQLQERTGRTAPTVIT [SEQ ID NO:330]		
PGM3	5238	0013375	O95394, A0A087WT27	MDLGAITKYSALHAKPNGLILQYGTAGFRTKAEHLDHVMFRMGLLAVLRSK QTKSTIGVMVTASHNPEEDNGVKLVDPLGEMLAPSWEEHATCLANAEEQ DMQRVLIDISEKEAVNLQQDAFVVIGRDTRPSSEKLSQSVIDGVTVLGGQF HDYGLLTTPQLHYMVYCRNTGGRYGKATIEGYYQKLSKAFVELTKQASCS GDEYRSLKVDCANGIGALKLREMEHYFSQGLSVQLFNDGSKGKLNHLCG ADFVKSHQKPPQGMEIKSNERCCSFDGDADRIVYYYHDADGHFHLIDGDK IATLISSFLKELLVEIGESLNIGVVQTAYANGSSTRYLEEVMKVPVYCTKTGV KHLHHKAQEFDIGVYFEANGHGTALFSTAVEMKIKQSAEQLEDKKRKAAK MLENIIDLFNQAAGDAISDMLVIEAILALKGLTVQQWDALYTDLPNRQLKVQ VADRRVISTTDAERQAVTPPGLQEAINDLVKKYKLSRAFVRPSGTEDVVRV YAEADSQESADHLAHEVSLAVFQLAGGIGERPQPGF [SEQ ID NO:331]	PGM3-CDG	Distúrbio de glicosilação
RFT1	91869	0163933	Q96AA3	MGSQEVLGHAARLASSGLLLQVLFRLITFVLNAFILRFLSKEIVGVVNVRLTL LYSTTLFLAREAFRRACLSGGTQRDWSQTLNLLWLTVPLGVFWSLFLGWI WLQLLEVPDPNVVPHYATGVVLFGLSAVVELLGEPFWVLAQAHMFVKLKV IAESLSVILKSVLTAFLVLWLPHWGLYIFSLAQLFYTTVLVLCYVIYFTKLLGS PESTKLQTLPVSRITDLLPNITRNGAFINWKEAKLTWSFFKQSFLKQILTEG ERYVMTFLNVLNFGDQGVYDIVNNLGSLVARLIFQPIEESFYIFFAKVLERG KDATLQKQEDVAVAAAVLESLLKLALLAGLTITVFGFAYSQLALDIYGGTML SSGSGPVLLRSYCLYVLLLAINGVTECFTFAAMSKEEVDRYNFVMLALSSS FLVLSYLLTRWCGSVGFILANCFNMGIRITQSLCFIHRYYRRSPHRPLAGLH LSPVLLGTFALSGGVTAVSEVFLCCEQGWPARLAHIAVGAFCLGATLGTAF LTETKLIHFLRTQLGVPRRTDKMT [SEQ ID NO:332]	RFT1-CDG	Distúrbio de glicosilação
SEC23	10483	0101310	Q15437, B4DJW8	MATYLEFIQQNEERDGVRFSWNVWPSSRLEATRMVVPLACLLTPLKERPD	SEC23B-CDG	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
B				LPPVQYEPVLCSRPTCKAVLNPLCQVDYRAKLWACNFCFQRNQFPPAYG GISEVNQPAELMPQFSTIEYVIQRGAQSPLIFLYVVDTCLEEDDLQALKESL QMSLSLLPPDALVGLITFGRMVQVHELSCEGISKSYVFRGTKDLTAKQIQD MLGLTKPAMPMQQARPAQPQEHPFASSRFLQPVHKIDMNLTDLLGELQR DPWPVTQGKRPLRSTGVALSIAVGLLEGTFPNTGARIMLFTGGPPTQGPG MVVGDELKIPIRSWHDIEKDNARFMKKATKHYEMLANRTAANGHCIDIYAC ALDQTGLLEMKCCANLTGGYMVMGDSFNTSLFKQTFQRIFTKDFNGDFR MAFGATLDVKTSRELKIAGAIGPCVSLNVKGPCVSENELGVGGTSQWKIC GLDPTSTLGIYFEVVNQHNTPIPQGGRGAIQFVTHYQHSSTQRRIRVTTIAR NWADVQSQLRHIEAAFDQEAAAVLMARLGVFRAESEEGPDVLRWLDRQLI RLCQKFGQYNKEDPTSFRLSDSFSLYPQFMFHLRRSPFLQVFNNSPDESS YYRHHFARQDLTQSLIMIQPILYSYSFHGPPEPVLLDSSSILADRILLMDTFF QIVIYLGETIAQWRKAGYQDMPEYENFKHLLQAPLDDAQEILQARFPMPRY INTEHGGSQARFLLSKVNPSQTHNNLYAWGQETGAPILTDDVSLQVFMDH LKKLAVSSAC [SEQ ID NO:333]		
SLC35A1	10559	0164414	P78382	MAAPRDNVTLLFKLYCLAVMTLMAAVYTIALRYTRTSDKELYFSTTAVCITE VIKLLLSVGILAKETGSLGRFKASLRENVLGSPKELLKLSVPSLVYAVQNNM AFLALSNLDAAVYQVTYQLKIPCTALCTVLMLNRTLSKLQWVSVFMLCAGV TLVQWKPAQATKVVVEQNPLLGFGAIAIAVLCSGFAGVYFEKVLKSSDTSL WVRNIQMYLSGIIVTLAGVYLSDGAEIKEKGFFYGYTYYVWFVIFLASVGGL YTSVVVKYTDNIMKGFSAAAAIVLSTIASVMLFGLQITLTFALGTLLVCVSIYL YGLPRQDTTSIQQGETASKERVIGV [SEQ ID NO:334]	SLC35A1-CDG	Distúrbio de glicosilação
SLC35A2	7355	0102100	P78381, A6NFI1, A6NKM8, B4DE15	MAAVGAGGSTAAPGPGAVSAGALEPGTASAAHRRLKYISLAVLVVQNASLILS IRYARTLPGDRFFATTAVVMAEVLKGLTCLLLLFAQKRGNVKHLVLFLHEAVLV QYVDTLKLAVPSLIYTLQNNLQYVAISNLPAATFQVTYQLKILTTALFSVLMLNR SLSRLQWASLLLLFTGVAIVQAQQAGGGGPRPLDQNPGAGLAAVVASCLSS GFAGVYFEKILKGSSGSVWLRNLQLGLFGTALGLVGLWWAEGTAVATRGFFF GYTPAVWGVVLNQAFGGLLVAVVVKYADNILKGFATSLSIVLSTVASIRLFGFH VDPLFALGAGLVIGAVYLYSLPRGAAKAIASASASASGPCVHQQPPGQPPPP QLSSHRGDLITEPFLPKLLTKVKGS [SEQ ID NO:335]	SLC35A2-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
SLC35C1	55343	0181830	Q96A29, B3KQH0	MNRAPLKRSRILHMALTGASDPSAEAEANGEKPFLLRALQIALVVSLYWVTSISMVFLNKYLLDSPSLRLDTPIFVTFYQCLVTTLLCKGLSALAACCPGAVDFPSLRLDLRVARSVLPLSVVFIGMITFNNLCLKYVGVAFYNVGRSLTTVFNVLLSYLLLKQTTSFYALLTCGIIIGGFWLGVDQEGAEGTLSWLGTVFGVLASLCVSLNAIYTTKVLPAVDGSIWRLTFYNNVNACILFLPLLLLLGELQALRDFAQLGSAHFWGMMTLGGLFGFAIGYVTGLQIKFTSPLTHNVSG TAKACAQTVLAVLYYEETKSFLWWTSNMMVLGGSSAYTWVRGWEMKKTPEEPSPKDSEKSAMGV [SEQ ID NO:336]	SLC35C1-CDG	Distúrbio de glicosilação
SSR4	6748	0180879	P51571	MAAMASLGALALLLLSSLSRCSAEACLEPQITPSYYTTSDAVISTETVFIVEISLTCKNRVQNMALYADVGGKQFPVTRGQDVGRYQVSWSLDHKSAHAGTYEVRFFDEESYSLLRKAQRNNEDISIIPPLFTVSVDHRGTWNGPWVSTEVLAAAIGLVIYYLAFSAKSHIQA [SEQ ID NO:337]	SSR4-CDG	Distúrbio de glicosilação
SRD5A3	79644	0128039	Q9H8P0	MAPWAEAEHSALNPLRAVWLTLTAAFLLTLLLQLLPPGLLPGCAIFQDLIRYGKTKCGEPSRPAACRAFDVPKRYFSHFYIISVLWNGFLLWCLTQSLFLGAPFPSWLHGLLRILGAAQFQGGELALSAFLVLVFLWLHSLRRLFECLYVSVFSNVMIHVVQYCFGLVYYVLVGLTVLSQVPMDGRNAYITGKNLLMQARWFHILGMMMFIWSSAHQYKCHVILGNLRKNKAGVVIHCNHRIPFGDWFEYVSSPNYLAELMIYVSMAVTFGFHNLTWWLVVTNVFFNQALSAFLSHQFYKSKFVSYPKHRKAFLPFLF [SEQ ID NO:338]	SRD5A3-CDG	Distúrbio de glicosilação
TMEM165	55858	0134851	Q9HC07	MAAAAPGNGRASAPRLLLLFLVPLLWAPAAVRAGPDEDLSHRNKEPPAPAQQLQPQPVAVQGPEPARVEKIFTPAAPVHTNKEDPATQTNLGFIHAFVAAISVIIVSELGDKTFFIAAIMAMRYNRLTVLAGAMLALGLMTCLSVLFGYATTVIPRVYTYYVSTVLFAIFGIRMLREGLKMSPDEGQEELEEVQAELKKKDEEFQRTKLLNGPGDVETGTSITVPQKKWLHFISPIFVQALTLTFLAEWGDRSQLTTIVLAAREDPYGVAVGGTVGHCLCTGLAVIGGRMIAQKISVRTVTIIGGIVFLAFAFSALFISPDSGF [SEQ ID NO:339]	TMEM165-CDG	Distúrbio de glicosilação
TRIP11	9321	0100815	Q15643	MSSWLGGLGSGLGQSLGQVGGSLASLTGQISNFTKDMLMEGTEEVEAELPDSRTKEIEAIHAILRSENERLKKLCTDLEEKHEASEIQIKQQSTSYRNQLQ	TRIP11-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QKEVEISHLKARQIALQDQLLKLQSAAQSVPSGAGVPATTASSSFAYGISH HPSAFHDDDMDFGDIISSQQEINRLSNEVSRLESEVGHWRHIAQTSKAQG TDNSDQSEICKLQNIIKELKQNRSQEIDDHQHEMSVLQNAHQQKLTEISRR HREELSDYEERIEELENLLQQGGSGVIETDLSKIYEMQKTIQVLQIEKVEST KKMEQLEDKIKDINKKLSSAENDRDILRREQEQLNVEKRQIMEECENLKLE CSKLQPSAVKQSDTMTEKERILAQSASVEEVFRLQQALSDAENEIMRLSSL NQDNSLAEDNLKLKMRIEVLEKEKSLLSQEKEELQMSLLKLNNEYEVIKST ATRDISLDSELHDLRLNLEAKEQELNQSISEKETLIAEIEELDRQNQEATKH MILIKDQLSKQQNEGDSIISKLKQDLNDEKKRVHQLEDDKMDITKELDVQKE KLIQSEVALNDLHLTKQKLEDKVENLVDQLNKSQESNVSIQKENLELKEHIR QNEEELSRIRNELMQSLNQDSNSNFKDTLLKEREAEVRNLKQNLSELEQL NENLKKVAFDVKMENEKLVLACEDVRHQLEECLAGNNQLSLEKNTIVETLK MEKGEIEAELCWAKKRLLEEANKYEKTIEELSNARNLNTSALQLEHEHLIKL NQKKDMEIAELKKNIEQMDTDHKETKDVLSSSLEEQKQLTQLINKKEIFIEK LKERSSKLQEELDKYSQALRKNEILRQTIEEKDRSLGSMKEENNHLQEELE RLREEQSRTAPVADPKTLDSVTELASEVSQLNTIKEHLEEEIKHHQKIIEDQ NQSKMQLLQSLQEQKKEMDEFRYQHEQMNATHTQLFLEKDEEIKSLQKTI EQIKTQLHEERQDIQTDNSDIFQETKVQSLNIENGSEKHDLSKAETERLVK GIKERELEIKLLNEKNISLTKQIDQLSKDEVGKLTQIIQQKDLEIQALHARISS TSHTQDVVYLQQQLQAYAMEREKVFAVLNEKTRENSHLKTEYHKMMDIVA AKEAALIKLQDENKKLSTRFESSGQDMFRETIQNLSRIIREKDIEIDALSQKC QTLLAVLQTSSTGNEAGGVNSNQFEELLQERDKLKQQVKKMEEWKQQV MTTVQNMQHESAQLQEELHQLQAQVLVDSDNNSKLQVDYTGLIQSYEQN ETKLKNFGQELAQVQHSIGQLCNTKDLLLGKLDIISPQLSSASLLTPQSAEC LRASKSEVLSESSELLQQELEELRKSLQEKDATIRTLQENNHRLSDSIAATS ELERKEHEQTDSEIKQLKEKQDVLQKLLKEKDLLIKAKSDQLLSSNENFTNK VNENELLRQAVTNLKERILILEMDIGKLKGENEKIVETYRGKETEYQALQET NMKFSMMLREKEFECHSMKEKALAFEQLLKEKEQGKTGELNQLLNAVKS MQEKTVVFQQERDQVMLALKQKQMENTALQNEVQRLRDKEFRSNQELE RLRNHLLESEDSYTREALAAEDREAKLRKKVTVLEEKLVSSSNAMENASH QASVQVESLQEQLNVVSKQRDETALQLSVSQEQVKQYALSLANLQMVLE HFQQEEKAMYSAELEKQKQLIAEWKKNAENLEGKVISLQECLDEANAALD		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SASRLTEQLDVKEEQIEELKRQNELRQEMLDDVQKKLMSLANSSEGKVDK VLMRNLFIGHFHTPKNQRHEVLRLMGSILGVRREEMEQLFHDDQGGVTR WMTGWLGGGSKSVPNTPLRPNQQSVVNSSFSELFVKFLETESHPSIPPP KLSVHDMKPLDSPGRRKRDTNAPESFKDTAESRSGRRTDVNPFLAPRSA AVPLINPAGLGPGGPGHLLLKPISDVLPTFTPLPALPDNSAGVVLKDLLKQ [SEQ ID NO:340]		
TUSC 3	7991	0104723	Q13454	MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLA EKVEQLMEWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCS VCRQANEEYQILANSWRYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAP TFMHFPPKGRPKRADTFDLQRIGFAAEQLAKWIADRTDVHIRVFRPPNYS GTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMVSLCIVFAMTSGQMWNH IRGPPYAHKNPHNGQVSYIHGSSQAQFVAESHIILVLNAAITMGMVLLNE AATSKGDVGKRRIICLVGLGLVVFFFSFLLSIFRSKYHGYPYSDLDFE [SEQ ID NO:341]	TUSC3-CDG	Distúrbio de glicosilação
ALG14	199857	0172339	Q96F25	MVCVLVLAAAAGAVAVFLILRIWVVLRSMDVTPRESLSILVVAGSGGHTTEI LRLLGSLSNAYSPRHYVIADTDEMSANKINSFELDRADRDPSNMYTKYYIH RIPRSREVQQSWPSTVFTTLHSMWLSFPLIHRVKPDLVLCNGPGTCVPICV SALLLGILGIKKVIIVYVESICRVETLSMSGKILFHLSDYFIVQWPALKEKYPK SVYLGRIV [SEQ ID NO:342]	ALG14-CDG	Distúrbio de glicosilação
B4GA LT1	2683	0086062	P15291, W6MEN3	MRLREPLLSGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLSRL PQLVGVSTPLQGGSNSAAAIGQSSGELRTGGARPPPPLGASSQPRPGGD SSPVVDSGPGPASNLTSVPVPHTTALSLPACPEESPLLVGPMLIEFNMPVD LELVAKQNPNVKMGGRYAPRDCVSPHKVAIIIPFRNRQEHLKYWLYYLHPV LQRQQLDYGIYVINQAGDTIFNRAKLLNVGFQEALKDYDYTCFVFSDVDLIP MNDHNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSKQQFLTINGF PNNYWGWGGEDDDIFNRLVFRGMSISRPNAVVGRCRMIRHSRDKKNEPN PQRFDRIAHTKETMLSDGLNSLTYQVLDVQRYPLYTQITVDIGTPS [SEQ ID NO:343]	B4GALT1-CDG	Distúrbio de glicosilação
DDOS	1650	0244038	A0A024RAD5,	MGYFRCARAGSFGRRRKMEPSTAARAWALFWLLLPLLGAVCASGPRTLV	DDOST-CDG	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
T			P39656	LLDNLNVRETHSLFFRSLKDRGFELTFKTADDPSLSLIKYGEFLYDNLIIFSP SVEDFGGNINVETISAFIDGGGSVLVAASSDIGDPLRELGSECGIEFDEEKT AVIDHHNYDISDLGQHTLIVADTENLLKAPTIVGKSSLNPILFRGVGMVADP DNPLVLDILTGSSTSYSFFPDKPITQYPHAVGKNTLLIAGLQARNNARVIFS GSLDFFSDSFFNSAVQKAAPGSQRYSQTGNYELAVALSRWVFKEEGVLR VGPVSHHRVGETAPPNAYTVTDLVEYSIVIQQLSNGKWVPFDGDDIQLEFV RIDPFVRTFLKKKGGKYSVQFKLPDVYGVFQFKVDYNRLGYTHLYSSTQV SVRPLQHTQYERFIPSAYPYYASAFSMMLGLFIFSIVFLHMKEKEKSD [SEQ ID NO:344]		glicosilação
NUS1	116150	0153989	Q96E22	MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVL APLGFTLRKPPAVGRNRRHHRHPRGGSCLAAAHHRMRWRADGRSLEKL PVHMGLVITEVEQEPSFSDIASLVVWCMAVGISYISVYDHQGIFKRNNSRL MDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLNCHLAVKVLSPEDGKA DIVRAAQDFCQLVAQKQKRPTDLDVDTLASLLSSNGCPDPDLVLKFGPVD STLGFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK [SEQ ID NO:345]	NUS1-CDG	Distúrbio de glicosilação
RPN2	6185	0118705	P04844	MAPPGSSTVFLLALTIIASTWALTPTHYLTKHDVERLKASLDRPFTNLESAF YSIVGLSSLGAQVPDAKKACTYIRSNLDPSNVDSLFYAAQASQALSGCEISI SNETKDLLLAAVSEDSSVTQIYHAVAALSGFGLPLASQEALSALTARLSKEE TVLATVQALQTASHLSQQADLRSIVEEIEDLVARLDELGGVYLQFEEGLETT ALFVAATYKLMDHVGTEPSIKEDQVIQLMNAIFSKKNFESLSEAFSVASAAA VLSHNRYHVPVVVVPEGSASDTHEQAILRLQVTNVLSQPLTQATVKLEHAK SVASRATVLQKTSFTPVGDVFELNFMNVKFSSGYYDFLVEVEGDNRYIAN TVELRVKISTEVGITNVDLSTVDKDQSIAPKTTRVTYPAKAKGTFIADSHQN FALFFQLVDVNTGAELTPHQTFVRLHNQKTGQEVVFVAEPDNKNVYKFEL DTSERKIEFDSASGTYTLYLIIGDATLKNPILWNVADVVIKFPEEEAPSTVLS QNLFTPKQEIQHLFREPEKRPPTV VSNTFTALILSPLLLLFALWIRIGANVSNFTFAPSTIIFHLGHAAMLGLMYVY WTQLNMFQTLKYLAILGSVTFLAGNRMLAQQAVKRTAH [SEQ ID NO:346]	RPN2-CDG	Distúrbio de glicosilação
SEC23	10484	0100934	Q15436	MTTYLEFIQQNEERDGVRFSWNVWPSSRLEATRMVVPVAALFTPLKERP	SEC23A-CDG	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
A				DLPPIQYEPVLCSRTTCRAVLNPLCQVDYRAKLWACNFCYQRNQFPPSYAGISELNQPAELLPQFSSIEYVVLRGPQMPLIFLYVVDTCMEDEDLQALKESMQMSLSLLPPTALVGLITFGRMVQVHELGCEGISKSYVFRGTKDLSAKQLQEMLGLSKVPLTQATRGPQVQQPPPSNRFLQPVQKIDMNLTDLLGELQRDPWPVPQGKRPLRSSGVALSIAVGLLECTFPNTGARIMMFIGGPATQGPGM VVGDELKTPIRSWHDIDKDNAKYVKKGTKHFEALANRAATTGHVIDIYACALDQTGLLEMKCCPNLTGGYMVMGDSFNTSLFKQTFQRVFTKDMHGQFKMGFGGTLEIKTSREIKISGAIGPCVSLNSKGPCVSENEIGTGGTCQWKICGLSPTTTLAIYFEVVNQHNAPIPQGGRGAIQFVTQYQHSSGQRRIRVTTIARNWADAQTQIQNIAASFDQEAAAILMARLAIYRAETEEGPDVLRWLDRQLIRLCQKFGEYHKDDPSSFRFSETFSLYPQFMFHLRRSSFLQVFNNSPDESSYYRHHFMRQDLTQSLIMIQPILYAYSFSGPPEPVLLDSSSILADRILLMDTFFQILIYHGETIAQWRKSGYQDMPEYENFRHLLQAPVDDAQEILHSRFPMPRYIDTEHGGSQARFLLSKVNPSQTHNNMYAWGQESGAPILTDDVSLQVFMDHLKKLAVSSAA [SEQ ID NO:347]		glicosilação
SLC35A3	23443	0117620	Q9Y2D2, A0A1W2PRT7, A0A1W2PSD1, A0A1W2PQL8	MFANLKYVSLGILVFQTTSLVLTMRYSRTLKEEGPRYLSSTAVVVAELLKIMACILLVYKDSKCSLRALNRVLHDEILNKPMETLKLAIPSGIYTLQNNLLYVALSNLDAATYQVTYQLKILTTALFSVSMLSKKLGVYQWLSLVILMTGVAFVQWPSDSQLDSKELSAGSQFVGLMAVLTACFSSGFAGVYFEKILKETKQSVWIRNIQLGFFGSIFGLMGVYIYDGELVSKNGFFQGYNRLTWIVVVLQALGGLVIAAVIKYADNILKGFATSLSIILSTLISYFWLQDFVPTSVFFLGAILVITATFLYGYDPKPAGNPTKA [SEQ ID NO:348]	SLC35A3-CDG	Distúrbio de glicosilação
ST3GAL3	6487	0126091	Q11203	MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEEDSNSVVLSFDSAGQTLGSEYDRLGFLLNLDSKLPAELATKYANFSEGACKPGYASALMTAIFPRFSKPAPMFLDDSFRKWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPALDSLRCRRCIIVGNGGVLANKSLGSRIDDYDIVVRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSLFVLAGFKWQDFKWLKYIVYKERVSASDGFWKSVATRVPKEPPEIRILNPYFIQEAAFTLIGLPFNNGLMGRGNIPTLGSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYYETVRMAAIKESWTHNIQREKEFLRKLVKARVITDLSSGI [SEQ ID NO:349]	ST3GAL3-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
STT3A	3703	0134910	P46977	MTKFGFLRLSYEKQDTLLKLLILSMAAVLSFSTRLFAVLRFESVIHEFDPYF NYRTTRFLAEEGFYKFHNWFDDRAWYPLGRIIGGTIYPGLMITSAAIYHVLH FFHITIDIRNVCVFLAPLFSSFTTIVTYHLTKELKDAGAGLLAAAMIAVVPGYI SRSVAGSYDNEGIAIFCMLLTYYMWIKAVKTGSICWAAKCALAYFYMVSS WGGYVFLINLIPLHVLVLMLTGRFSHRIYVAYCTVYCLGTILSMQISFVGFQ PVLSSEHMAAFGVFGLCQIHAFVDYLRSKLNPQQFEVLFRSVISLVGFVLL TVGALLMLTGKISPWTGRFYSLLDPSYAKNNIPIIASVSEHQPTTWSSYYFD LQLLVFMFPVGLYYCFSNLSDARIFIIMYGVTSMYFSAVMVRLMLVLAPVM CILSGIGVSQVLSTYMKNLDISRPDKKSKKQQDSTYPIKNEVASGMILVMAF FLITYTFHSTWVTSEAYSSPSIVLSARGGDGSRIIFDDFREAYYWLRHNTPE DAKVMSWWDYGYQITAMANRTILVDNNTWNNTHISRVGQAMASTEEKAY EIMRELDVSYVLVIFGGLTGYSSDDINKFLWMVRIGGSTDTGKHIKENDYYT PTGEFRVDREGSPVLLNCLMYKMCYYRFGQVYTEAKRPPGFDRVRNAEI GNKDFELDVLEEAYTTEHWLVRIYKVKDLDNRGLSRT [SEQ ID NO:350]	STT3A-CDG	Distúrbio de glicosilação
STT3B	201595	0163527	Q8TCJ2	MAEPSAPESKHKSSLNSSPWSGLMALGNSRHGHHGPGAQCAHKAAGGA APPKPAPAGLSGGLSQPAGWQSLLSFTILFLAWLAGFSSRLFAVIRFESIIH EFDPWFNYRSTHHLASHGFYEFLNWFDERAWYPLGRIVGGTVYPGLMITA GLIHWILNTLNITVHIRDVCVFLAPTFSGLTSISTFLLTRELWNQGAGLLAAC FIAIVPGYISRSVAGSFDNEGIAIFALQFTYYLWVKSVKTGSVFWTMCCCLS YFYMVSAWGGYVFIINLIPLHVFVLLLMQRYSKRVYIAYSTFYIVGLILSMQIP FVGFQPIRTSEHMAAAGVFALLQAYAFLQYLRDRLTKQEFQTLFFLGVSLA AGAVFLSVIYLTYTGYIAPWSGRFYSLWDTGYAKIHIPIIASVSEHQPTTWV SFFFDLHILVCTFPAGLWFCIKNINDERVFVALYAISAVYFAGVMVRLMLTLT PVVCMLSAIAFSNVFEHYLGDDMKRENPPVEDSSDEDDKRNQGNLYDKA GKVRKHATEQEKTEEGLGPNIKSIVTMLMLMLLMMFAVHCTWVTSNAYSS PSVVLASYNHDGTRNILDDFREAYFWLRQNTDEHARVMSWWDYGYQIAG MANRTTLVDNNTWNNSHIALVGKAMSSNETAAYKIMRTLDVDYVLVIFGG VIGYSGDDINKFLWMVRIAEGEHPKDIRESDYFTPQGEFRVDKAGSPTLLN CLMYKMSYYRFGEMQLDFRTPPGFDRTRNAEIGNKDIKFKHLEEAFTSEH WLVRIYKVKAPDNRETLDHKPRVTNIFPKQKYLSKKTTKRKRGYIKNKLVF KKGKKISKKTV [SEQ ID NO:351]	STT3B-CDG	Distúrbio de glicosilação

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
AGA	175	0038002	P20933	MARKSNLPVLLVPFLLCQALVRCSSPLPLVVNTWPFKNATEAAWRALASG GSALDAVESGCAMCEREQCDGSVGFGGSPDELGETTLDAMIMDGTTMD VGAVGDLRRIKNAIGVARKVLEHTTHTLLVGESATTFAQSMGFINEDLSTTA SQALHSDWLARNCQPNYWRNVIPDPSKYCGPYKPPGILKQDIPIHKETED DRGHDTIGMVVIHKTGHIAAGTSTNGIKFKIHGRVGDSPIPGAGAYADDTA GAAAATGNGDILMRFLPSYQAVEYMRRGEDPTIACQKVISRIQKHFPEF	Aspartilglicosami-núria	Distúrbio de Armazenamento Lisossômico
				FGAVICANVTGSYGAACNKLSTFTQFSFMVYNSEKNQPTEEKVDCI [SEQ ID NO:352]		
ARSA	410	0100299	A0A0C4DFZ2, B4DVI5, P15289	MGAPRSLLLALAAGLAVARPPNIVLIFADDLGYGDLGCYGHPSSTTPNLDQ LAAGGLRFTDFYVPVSLCTPSRAALLTGRLPVRMGMYPGVLVPSSRGGLP LEEVTVAEVLAARGYLTGMAGKWHLGVGPEGAFLPPHQGFHRFLGIPYSH DQGPCQNLTCFPPATPCDGGCDQGLVPIPLLANLSVEAQPPWLPGLEAR YMAFAHDLMADAQRQDRPFFLYYASHHTHYPQFSGQSFAERSGRGPFG DSLMELDAAVGTLMTAIGDLGLLEETLVIFTADNGPETMRMSRGGCSGLL RC	Leucodistrofia metacromática	Distúrbio de Armazenamento Lisossômico
				GKGTTYEGGVREPALAFWPGHIAPGVTHELASSLDLLPTLAALAGAPLPNV TLDGFDLSPLLLGTGKSPRQSLFFYPSYPDEVRGVFAVRTGKYKAHFFTQ GSAHSDTTADPACHASSSLTAHEPPLLYDLSKDPGENYNLLGGVAGATPE VLQALKQLQLLKAQLDAAVTFGPSQVARGEDPALQICCHPGCTPRPACCH CPDPHA [SEQ ID NO:353]		
ARSB	411	0113273	A0A024RAJ9, P15848, A8K4A0	MGPRGAASLPRGPGPRRLLLPVVLPLLLLLLLAPPGSGAGASRPPHLVFLL ADDLGWNDVGFHGSRIRTPHLDALAAGGVLLDNYYTQPLCTPSRSQLLTG RYQIRTGLQHQIIWPCQPSCVPLDEKLLPQLLKEAGYTTHMVGKWHLGMY RKECLPTRRGFDTYFGYLLGSEDYYSHERCTLIDALNVTRCALDFRDGEE VATGYKNMYSTNIFTKRAIALITNHPPEKPLFLYLALQSVHEPLQVPEEYLK PYDFIQDKNRHHYAGMVSLMDEAVGNVTAALKSSGLWNNTVFIFSTDNG GQTLAGGNNWPLRGRKWSLWEGGVRGVGFVASPLLKQKGVKNRELIHIS DWLPTLVKLARGHTNGTKPLDGFDVWKTISEGSPSPRIELLHNIDPNFVDS SPCPRNSMAPAKDDSSLPEYSAFNTSVHAAIRHGNWKLLTGYPGCGYWF PPPSQYNVSEIPSSDPPTKTLWLFDIDRDPEERHDLSREYPHIVTKLLSRLQ	Mucopolissacari-dose tipo VI	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FYHKHSVPVYFPAQDPRCDPKATGVWGPWM [SEQ ID NO:354]		
ASAH1	427	0104763	A8K0B6, Q13510, Q53H01	MPGRSCVALVLLAAAVSCAVAQHAPPWTEDCRKSTYPPSGPTYRGAVPWYTINLDLPPYKRWHELMLDKAPVLKVIVNSLKNMINTFVPSGKIMQVVDEKLPGLLGNFPGPFEEEMKGIAAVTDIPLGEIISFNIFYELFTICTSIVAEDKKGHLIHGRNMDFGVFLGWNINNDTWVITEQLKPLTVNLDFQRNNKTVFKASSFAGYVGMLTGFKPGLFSLTLNERFSINGGYLGILEWILGKKDVMWIGFLTRTVLENSTSYEEAKNLLTKTKILAPAYFILGGNQSGEGCVITRDRKESLDVYELDAKQGRWYVVQTNYDRWKHPFFLDDRRTPAKMCLNRTSQENISFETMYDVLSTKPVLNKLTVYTTLIDVTKGQFETYLRDCPDPCIGW [SEQ ID NO:355]	Doença de Farber	Distúrbio de Armazenamento Lisossômico
ATP13A2	23400	0159363	Q8N4D4, Q9NQ11, Q8NBS1	MSADSSPLVGSTPTGYGTLTIGTSIDPLSSSVSSVRLSGYCGSPWRVIGYHVVVWMMAGIPLLLFRWKPLWGVRLRLRPCNLAHAETLVIEIRDKEDSSWQLFTVQVQTEAIGEGSLEPSPQSQAEDGRSQAAVGAVPEGAWKDTAQLHKSEEAVSVGQKRVLRYYLFQGQRYIWIETQQAFYQVSLLDHGRSCDDVHRSRHGLSLQDQMVRKAIYGPNVISIPVKSYPQLLVDEALNPYYGFQAFSIALWLADHYYWYALCIFLISSISICLSLYKTRKQSQTLRDMVKLSMRVCVCRPGGEEEWVDSSELVPGDCLVLPQEGGLMPCDAALVAGECMVNESSLTGESIPVLKTALPEGLGPYCAETHRRHTLFCGTLILQARAYVGPHVLAVVTRTGFCTAKGGLVSSILHPRPINFKFYKHSMKFVAALSVLALLGTIYSIFILYRNRVPLNEIVIRALDLVTVVVPPALPAAMTVCTLYAQSRLRRQGIFCIHPLRINLGGKLQLVCFDKTGTLTEDGLDVMGVVPLKGQAFLPLV PEPRRLPVGPLLRALATCHALSRLQDTPVGDPMDLKMVESTGWVLEEEPAADSAFGTQVLAVMRPPLWEPQLQAMEEPPVPVSVLHRFPFSSALQRMSVVVAWPGATQPEAYVKGSPELVAGLCNPETVPTDFAQMLQSYTAAGYRVVALASKPLPTVPSLEAAQQLTRDTVEGDLSLLGLLVMRNLLKPQTTPVIQALRRTRIRAVMVTGDNLQTAVTVARGCGMVAPQEHLIIVHATHPERGQPASLEFLPMESPTAVNGVKDPDQAASYTVEPDPRSRHLALSGPTFGIIVKHFPKL LPKVLVQGTVFARMAPEQKTELVCELQKLQYCVGMCGDGANDCGALKAADVGISLSQAEASVVSPFTSSMASIECVPMVIREGRCSLDTSFSVFKYMALYSLTQFISVLILYTINTNLGDLQFLAIDLVITTTVAVLMSRTGPALVLGRVRPPGALLSVPVLSSLLLQMVLVTGVQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY	Lipofuscinose ceroide neuronal 12 (CLN12), síndrome de Kufor-Rakeb (KRS)	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ENTVVFSLSSFQYLILAAAVSKGAPFRRPLYTNVPFLVALALLSSVLVGLVL VPGLLQGPLALRNITDTGFKLLLLGLVTLNFVGAFMLESVLDQCLPACLRRL RPKRASKKRFKQLERELAEQPWPPLPAGPLR [SEQ ID NO:356]		
CLN3	1201	0188603,0261832	A0A024QZB8, Q13286, B4DMY6, Q2TA70, B4DFF3	MGGCAGSRRRFSDSEGEETVPEPRLPLLDHQGAHWKNAVGFWLLGLCN NFSYVVMLSAAHDILSHKRTSGNQSHVDPGPTPIPHNSSSRFDCNSVSTA AVLLADILPTLVIKLLAPLGLHLLPYSPRVLVSGICAAGSFVLVAFSHSVGTS LCGVVFASISSGLGEVTFLSLTAFYPRAVISWWSSGTGGAGLLGALSYLGL TQAGLSPQQTLLSMLGIPALLLASYFLLLTSPEAQDPGGEEEAESAARQPLI RTEAPESKPGSSSSLSLRERWTVFKGLLWYIVPLVVVYFAEYFINQGLFEL LFFWNTSLSHAQQYRWYQMLYQAGVFASRSSLRCCRIRFTWALALLQCL NLVFLLADVWFGFLPSIYLVFLIILYEGLLGGAAYVNTFHNIALETSDEHREF AMAATCISDTLGISLSGLLALPLHDFLCQLS [SEQ ID NO:357]	Lipofuscinose ceroide neuronal 3 (CLN3)	Distúrbio de Armazenamento Lisossômico
CLN5	1203	0102805	A0A024R644, O75503	MAQEVDTAQGAEMRRGAGAARGRASWCWALALLWLAVVPGWSRVSGI PSRRHWPVPYKRFDFRPKPDPYCQAKYTFCPTGSPIPVMEGDDDIEVFRL QAPVWEFKYGDLLGHLKIMHDAIGFRSTLTGKNYTMEWYELFQLGNCTFP HLRPEMDAPFWCNQGAACFFEGIDDVHWKENGTLVQVATISGNMFNQMA KWVKQDNETGIYYETWNVKASPEKGAETWFDSYDCSKFVLRTFNKLAEF GAEFKNIETNYTRIFLYSGEPTYLGNETSVFGPTGNKTLGLAIKRFYYPFKP HLPTKEFLLSLLQIFDAVIVHKQFYLFYNFEYWFLPMKFPFIKITYEEIPLPIR NKTLSGL [SEQ ID NO:358]	Lipofuscinose ceroide neuronal 5 (CLN5)	Distúrbio de Armazenamento Lisossômico
CLN6	54982	0128973	A0A024R601, Q9NWW5	MEATRRRQHLGATGGPGAQLGASFLQARHGSVSADEAARTAPFHLDLWF YFTLQNWVLDFGRPIAMLVFPLEWFPLNKPSVGDYFHMAYNVITPFLLLKLI ERSPRTLPRSITYVSIIIFIMGASIHLVGDSVNHRLLFSGYQHHLSVRENPIIK NLKPETLIDSFELLYYYDEYLGHCMWYIPFFLILFMYFSGCFTASKAESLIPG PALLLVAPSGLYYWYLVTEGQIFILFIFTFFAMLALVLHQKRKRLFLDSNGLF LFSSFALTLLLVALWVAWLWNDPVLRKKYPGVIYVPEPWAFYTLHVSSRH [SEQ ID NO:359]	Lipofuscinose ceriide neuronal 6 (CLN6)	Distúrbio de Armazenamento Lisossômico
CLN8	2055	0182372,0278220	A0A024QZ57,	MNPASDGGTSESIFDLDYASWGIRSTLMVAGFVFYLGVFVVCHQLSSSLN	Lipofuscinose	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
			Q9UBY8	ATYRSLVAREKVFWDLAATRAVFGVQSTAAGLWALLGDPVLHADKARGQ QNWCWFHITTATGFFCFENVAVHLSNLIFRTFDLFLVIHHLFAFLGFLGCLV NLQAGHYLAMTTLLLEMSTPFTCVSWMLLKAGWSESLFWKLNQWLMIHM FHCRMVLTYHMWWVCFWHWDGLVSSLYLPHLTLFLVGLALLTLIINPYWT HKKTQQLLNPVDWNFAQPEAKSRPEGNGQLLRKKRP [SEQ ID NO:360]	ceroide neuronal 8 (CLN8)	Armazenamento Lisossômico
CTNS	1497	0040531	A0A0S2Z3I9, O60931, A0A0S2Z3K3	MIRNWLTIFILFPLKLVEKCESSVSLTVPPVVKLENGSSTNVSLTLRPPLNAT LVITFEITFRSKNITILELPDEVVVPPGVTNSSFQVTSQNVGQLTVYLHGNH SNQTGPRIRFLVIRSSAISIINQVIGWIYFVAWSISFYPQVIMNWRRKSVIGLS FDFVALNLTGFVAYSVFNIGLLWVPYIKEQFLLKYPNGVNPVNSNDVFFSL HAVVLTLIIIVQCCLYERGGQRVSWPAIGFLVLAWLFAFVTMIVAAVGVTTW LQFLFCFSYIKLAVTLVKYFPQAYMNFYYKSTEGWSIGNVLLDFTGGSFSLL QMFLQSYNNDQWTLIFGDPTKFGLGVFSIVFDVVFFIQHFCLYRKRPGYDQ LN [SEQ ID NO:361]	cistinose	Distúrbio de Armazenamento Lisossômico
CTSA	5476	0064601	P10619, X6R8A1, B4E324, X6R5C5	MIRAAPPPLFLLLLLLLLLVSWASRGEAAPDQDEIQRLPGLAKQPSFRQYS GYLKGSGSKHLHYWFVESQKDPENSPVVLWLNGGPGCSSLDGLLTEHGP FLVQPDGVTLEYNPYSWNLIANVLYLESPAGVGFSYSDDKFYATNDTEVA QSNFEALQDFFRLFPEYKNNKLFLTGESYAGIYIPTLAVLVMQDPSMNLQG LAVGNGLSSYEQNDNSLVYFAYYHGLLGNRLWSSLQTHCCSQNKCNFYD NKDLECVTNLQEVARIVGNSGLNIYNLYAPCAGGVPSHFRYEKDTVVVQD LGNIFTRLPLKRMWHQALLRSGDKVRMDPPCTNTTAASTYLNNPYVRKAL NIPEQLPQWDMCNFLVNLQYRRLYRSMNSQYLKLLSSQKYQILLYNGDVD MACNFMGDEWFVDSLNQKMEVQRRPWLVKYGDSGEQIAGFVKEFSHIAF LTIKGAGHMVPTDKPLAAFTMFSRFLNKQPY [SEQ ID NO:362]	Galactosialidose	Distúrbio de Armazenamento Lisossômico
CTSD	1509	0117984	P07339, V9HWI3	MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPV SKYSQAVPAVTEGPIPEVLKNYMDAQYYGEIGIGTPPQCFTVVFDTGSSNL WVPSIHCKLLDIACWIHHKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQD TVSVPCQSASSASALGGVKVERQVFG EATKQPGITFIAAKFDGILGMAYPRISVNNVLPVFDNLMQQKLVDQNIFSFY LSRDPDAQPGGELMLGGTDSKYYKGSLSYLNVTRKAYWQVHLDQVEVAS GLTLCKEGCEAIVDTGTSLMVGPVDEVRELQKAIGAVPLIQGEYMIPCEKV	Lipofuscinose ceroide neuronal 10 (CLN10)	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				STLPAITLKLGGKGYKLSPEDYTLKVSQAGKTLCLSGFMGMDIPPPSGPLW ILGDVFIGRYYTVFDRDNNRVGFAEAARL [SEQ ID NO:363]		
CTSF	8722	0174080	Q9UBX1	MAPWLQLLSLLGLLPGAVAAPAQPRAASFQAWGPPSPELLAPTRFALEMF NRGRAAGTRAVLGLVRGRVRRAGQGSLYSLEATLEEPPCNDPMVCRLPV SKKTLLCSFQVLDELGRHVLLRKDCGPVDTKVPGAGEPKSAFTQGSAMIS SLSQNHPDNRNETFSSVISLLNEDPLSQDLPVKMASIFKNFVITYNRTYESK EEARWRLSVFVNNMVRAQKIQALDRGTAQYGVTKFSDLTEEEFRTIYLNTL LRKEPGNKMKQAKSVGDLAPPEWDWRSKGAVTKVKDQGMCGSCWAFS VTGNVEGQWFLNQGTLLSLSEQELLDCDKMDKACMGGLPSNAYSAIKNL GGLETEDDYSYQGHMQSCNFSAEKAKVYINDSVELSQNEQKLAAWLAKR GPISVAINAFGMQFYRHGISRPLRPLCSPWLIDHAVLLVGYGNRSDVPFWA IKNSWGTDWGEKGYYYLHRGSGACGVNTMASSAVVD [SEQ ID NO:364]	Lipofuscinose ceroide neuronal 13 (CLN13)	Distúrbio de Armazenamento Lisossômico
CTSK	1513	0143387	P43235	MWGLKVLLLPVVSFALYPEEILDTHWELWKKTHRKQYNNKVDEISRRLIWE KNLKYISIHNLEASLGVHTYELAMNHLGDMTSEEVVQKMTGLKVPLSHSRS NDTLYIPEWEGRAPDSVDYRKKGYVTPVKNQGQCGSCWAFSSVGALEG QLKKKTGKLLNLSPQNLVDCVSENDGCGGGYMTNAFQYVQKNRGIDSED AYPYVGQEESCMYNPTGKAAKCRGYREIPEGNEKALKRAVARVGPVSVAI DASLTSFQFYSKGVYYDESCNSDNLNHAVLAVGYGIQKGNKHWIIKNSWG ENWGNKGYILMARNKNNACGIANLASFPKM [SEQ ID NO:365]	Picnodisostose	Distúrbio de Armazenamento Lisossômico
DNAJC5	80331	0101152	Q6AHX3, Q9H3Z4	MADQRQRSLSTSGESLYHVLGLDKNATSDDIKKSYRKLALKYHPDKNPDN PEAADKFKEINNAHAILTDATKRNIYDKYGSLGLYVAEQFGEENVNTYFVLS SWWAKALFVFCGLLTCCYCCCCLCCCFNCCCGKCKPKAPEGEETEFYVS PEDLEAQLQSDEREATDTPIVIQPASATETTQLTADSHPSYHTDGFN [SEQ ID NO:366]	Lipofuscinose ceroide neuronal 4 (CLN4)	Distúrbio de Armazenamento Lisossômico
FUCA1	2517	0179163	P04066, B5MDC5	MRAPGMRSRPAGPALLLLLLFLGAAESVRRAQPPRRYTPDWPSLDSRPL PAWFDEAKFGVFIHWGVFSVPAWGSEWFWWHWQGEGRPQYQRFMRD NYPPGFSYADFGPQFTARFFHPEEWADLFQAAGAKYVVLTTKHHEGFTN WPSPVSWNWNSKDVGPHRDLVGELGTALRKRNIRYGLYHSLLEWFHPLY	Fucosidose	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LLDKKNGFKTQHFVSAKTMPELYDLVNSYKPDLIWSDGEWECPDTYWNS TNFLSWLYNDSPVKDEVVVNDRWGQNCSCHHGGYYNCEDKFKPQSLPD HKWEMCTSIDKFSWGYRRDMALSDVTEESEIISELVQTVSLGGNYLLNIGP TKDGLIVPIFQERLLAVGK WLSINGEAIYASKPWRVQWEKNTTSVWYTSKGSAVYAIFLHWPENGVLNL ESPITTSTTKITMLGIQGDLKWSTDPDKGLFISLPQLPPSAVPAEFAWTIKLT GVK[SEQ ID NO:367]		
GAA	2548	0171298	P10253	MGVRHPPCSHRLLAVCALVSLATAALLGHILLHDFLLVPRELSGSSPVLEE THPAHQQGASRPGPRDAQAHPGRPRAVPTQCDVPPNSRFDCAPDKAITQ EQCEARGCCYIPAKQGLQGAQMGQPWCFFPPSYPSYKLENLSSSEMGY TATLTRTTPTFFPKDILTLRLDVMMETENRLHFTIKDPANRRYEVPLETPHV HSRAPSPLYSVEFSEEPFGVIVRRQLDGRVLLNTTVAPLFFADQFLQLSTS LPSQYITGLAEHLSPLMLSTSWTRITLWNRDLAPTPGANLYGSHPFYLALE DGGSAHGVFLLNSNAMDVVLQPSPALSWRSTGGILDVYIFLGPEPKSVVQ QYLDVVGYPFMPPYWGLGFHLCRWGYSSTAITRQVVENMTRAHFPLDVQ WNDLDYMDSRRDFTFNKDGFRDFPAMVQELHQGGRRYMMIVDPAISSSG PAGSYRPYDEGLRRGVFITNETGQPLIGKVWPGSTAFPDFTNPTALAWWE DMVAEFHDQVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVG GTLQAATICASSHQFLSTHYNLHNLYGLTEAIASHRALVKARGTRPFVISRS TFAGHGRYAGHWTGDVWSSWEQLASSVPEILQFNLLGVPLVGADVCGFL GNTSEELCVRWTQLGAFYPFMRNHNSLLSLPQEPYSFSEPAQQAMRKAL TLRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWTVDHQLLWGEA LLITPVLQAGKAEVTGYFPLGTWYDLQTVPVEALGSLPPPPAAPREPAIHS EGQWVTLPAPLDTINVHLRAGYIIPLQGPGLTTTESRQQPMALAVALTKGG EARGELFWDDGESLEVLERGAYTQVIFLARNNTIVNELVRVTSEGAGLQL QKVTVLGVATAPQQVLSNGVPVSNFTYSPDTKVLDICVSLLMGEQFLVSW C [SEQ ID NO:368]	Doença de Pompe	Distúrbio de Armazenamento Lisossômico
GALC	2581	0054983	A0A0A0MQV0, P54803	MAEWLLSASWQRRAKAMTAAAGSAGRAAVPLLLCALLAPGGAYVLDDSD GLGREFDGIGAVSGGGATSRLLVNYPEPYRSQILDYLFKPNFGASLHILKV EIGGDGQTTDGTEPSHMHYALDENYFRGYEWWLMKEAKKRNPNITLIGLP WSFPGWLGKGFDWPYVNLQLTAYYVVTWIVGAKRYHDLDIDYIGIWNERS	Doença de Krabbe	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				YNANYIKILRKMLNYQGLQRVKIIASDNLWESISASMLLDAELFKVVDVIGAH YPGTHSAKDAKLTGKKLWSSEDFSTLNSDMGAGCWGRILNQNYINGYMT STIAWNLVASYYEQLPYGRCGLMTAQEPWSGHYVVESPVWVSAHTTQFT QPGWYYLKTVGHLEKGGSYVALTDGLGNLTIIIETMSHKHSKCIRPFLPYFN VSQQFATFVLKGSFSEIPELQVWYTKLGKTSERFLFKQLDSLWLLDSDGSF TLSLHEDELFTLTTLTTGRKGSYPLPPKSQPFPSTYKDDFNVDYPFFSEAP NFADQTGVFEYFTNIEDPGEHHFTLRQVLNQRPITWAADASNTISIIGDYN WTNLTIKCDVYIETPDTGGVFIAGRVNKGGILIRSARGIFFWIFANGSYRVT GDLAGWIIYALGRVEVTAKKWYTLTLTIKGHFTSGMLNDKSLWTDIPVNFP KNGWAAIGTHSFEFAQFDNFLVEATR [SEQ ID NO:369]		
GALNS	2588	0141012	P34059, Q96I49, Q6YL38	MAAVVAATRWWQLLLVLSAAGMGASGAPQPPNILLLLMDDMGWGDLGV YGEPSRETPNLDRMAAEGLLFPNFYSANPLCSPSRAALLTGRLPIRNGFYT TNAHARNAYTPQEIVGGIPDSEQLLPELLKKAGYVSKIVGKWHLGHRPQFH PLKHGFDEWFGSPNCHFGPYDNKARPNIPVYRDWEMVGRYYEEFPINLK TGEANLTQIYLQEALDFIKRQARHHPFFLYWAVDATHAPVYASKPFLGTSQ RGRYGDAVREIDDSIGKILELLQDLHVADNTFVFFTSDNGAALISAPEQGG SNGPFLCGKQTTFEGGMREPALAWWPGHVTAGQVSHQLGSIMDLFTTSL ALAGLTPPSDRAIDGLNLLPTLLQGRLMDRPIFYYRGDTLMAATLGQHKAH FWTWTNSWENFRQGIDFCPGQNVSGVTTHNLEDHTKLPLIFHLGRDPGE RFPLSFASAEYQEALSRITSVVQQHQEALVPAQPQLNVCNWAVMNWAPP GCEKLGKCLTPPESIPKKCLWSH [SEQ ID NO:370]	Mucopolissacaridose tipo IVa	Distúrbio de Armazenamento Lisossômico
GLA	2717	0102393	P06280, Q53Y83	MQLRNPELHLGCALALRFLALVSWDIPGARALDNGLARTPTMGWLHWER FMCNLDCQEEPDSCISEKLFMEMAELMVSEGWKDAGYEYLCIDDCWMAP QRDSEGRLQADPQRFPHGIRQLANYVHSKGLKLGIYADVGNKTCAGFPG SFGYYDIDAQTFADWGVDLLKFDGCYCDSLENLADGYKHMSLALNRTGR SIVYSCEWPLYMWPFQKPNYTEIRQYCNHWRNFADIDDSWKSIKSILDWT SFNQERIVDVAGPGGWNDPDMLVIGNFGLSWNQQVTQMALWAIMAAPLF MSNDLRHISPQAKALLQDKDVIAINQDPLGKQGYQLRQGDNFEVWERPLS GLAWAVAMINRQEIG	Doença de Fabry	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GPRSYTIAVASLGKGVACNPACFITQLLPVKRKLGFYEWTSRLRSHINPTG TVLLQLENTMQMSLKDLL [SEQ ID NO:371]		
GLB1	2720	0170266	P16278, B7Z6Q5	MPGFLVRILPLLLVLLLLGPTRGLRNATQRMFEIDYSRDSFLKDGQPFRYIS GSIHYSRVPRFYWKDRLLKMKMAGLNAIQTYVPWNFHEPWPGQYQFSED HDVEYFLRLAHELGLLVILRPGPYICAEWEMGGLPAWLLEKESILLRSSDP DYLAAVDKWLGVLLPKMKPLLYQNGGPVITVQVENEYGSYFACDFDYLRF LQKRFRHHLGDDVVLFTTDGAHKTFLKCGALQGLYTTVDFGTGSNITDAFL SQRKCEPKGPLINSEFYTGWLDHWGQPHSTIKTEAVASSLYDILARG	Gangliosidose GM1, Mucopolis-sacaridose IVb	Distúrbio de Armazenamento Lisossômico
				ASVNLYMFIGGTNFAYWNGANSPYAAQPTSYDYDAPLSEAGDLTEKYFAL RNIIQKFEKVPEGPIPPSTPKFAYGKVTLEKLKTVGAALDILCPSGPIKSLYP LTFIQVKQHYGFVLYRTTLPQDCSNPAPLSSPLNGVHDRAYVAVDGIPQG VLERNNVITLNITGKAGATLDLLVENMGRVNYGAYINDFKGLVSNLTLSSNI LTDWTIFPLDTEDAVRSHLGGWGHRDSGHHDEAWAHNSSNYTLPAFYMG NFSIPSGIPDLPQDTFIQFPGWTKGQVWINGFNLGRYWPARGPQLTLFVP QHILMTSAPNTITVLELEWAPCSSDDPELCAVTFVDRPVIGSSVTYDHPSK PVEKRLMPPPPQKNKDSWLDHV [SEQ ID NO:372]		
GM2A	2760	0196743	P17900	MQSLMQAPLLIALGLLLAAPAQAHLKKPSQLSSFSWDNCDEGKDPAVIRSL TLEPDPIIVPGNVTLSVMGSTSVPLSSPLKVDLVLEKEVAGLWIKIPCTDYIG SCTFEHFCDVLDMLIPTGEPCPEPLRTYGLPCHCPFKEGTYSLPKSEFVVP DLELPSWLTTGNYRIESVLSSSGKRLGCIKIAASLKGI [SEQ ID NO:373]	GM2-gangliosidose, variante AB	Distúrbio de Armazenamento Lisossômico
GNP-TAB	79158	0111670	Q3T906	MLFKLLQRQTYTCLSHRYGLYVCFLGVVVTIVSAFQFGEVVLEWSRDQYH VLFDSYRDNIAGKSFQNRLCLPMPIDVVYTWVNGTDLELLKELQQVREQM EEEQKAMREILGKNTTEPTKKSEKQLECLLTHCIKVPMLVLDPALPANITLK DLPSLYPSFHSASDIFNVAKPKNPSTNVSVVVFDSTKDVEDAHSGLLKGNS RQTVWRGYLTTDKEVPGLVLMQDLAFLSGFPPTFKETNQLKTKLPENLSS KVKLLQLYSEASVALLKLNNPKDFQELNKQTKKNMTIDGKELTISPA	Mucolipidose tipo II alfa/beta, muco-lipidose III al-fa/beta	Distúrbio de Armazenamento Lisossômico
				YLLWDLSAISQSKQDEDISASRFEDNEELRYSLRSIERHAPWVRNIFIVTNG QIPSWLNLDNPRVTIVTHQDVFRNLSHLPTFSSPAIESHIHRIEGLSQKFIYL NDDVMFGKDVWPDDFYSHSKGQKVYLTWPVPNCAEGCPGSWIKDGYCD		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KACNNSACDWDGGDCSGNSGGSRYIAGGGGTGSIGVGQPWQFGGGINS VSYCNQGCANSWLADKFCDQACNVLSCGFDAGDCGQDHFHELYKVILLP NQTHYIIPKGECLPYFSFAEVAKRGVEGAYSDNPIIRHASIANKWKTIHLIMH SGMNATTIHFNLTFQNTNDEEFKMQITVEVDTREGPKLNSTAQKGYENLV SPITLLPEAEILFEDIPKEKRFPKFKRHDVNSTRRAQEEVKIPLVNISLLPKDA QLSLNTLDLQLEHGDITLKGYNLSKSALLRSFLMNSQHAKIKNQAIITDETN DSLVAPQEKQVHKSILPNSLGVSERLQRLTFPAVSVKVNGHDQGQNPPLD LETTARFRVETHTQKTIGGNVTKEKPPSLIVPLESQMTKEKKITGKEKENSR MEENAENHIGVTEVLLGRKLQHYTDSYLGFLPWEKKKYFQDLLDEEESLK TQLAYFTDSKNTGRQLKDTFADSLRYVNKILNSKFGFTSRKVPAHMPHMID RIVMQELQDMFPEEFDKTSFHKVRHSEDMQFAFSYFYYLMSAVQPLNISQ VFDEVDTDQSGVLSDREIRTLATRIHELPLSLQDLTGLEHMLINCSKMLPAD ITQLNNIPPTQESYYDPNLPPVTKSLVTNCKPVTDKIHKAYKDKNKYRFEIM GEEEIAFKMIRTNVSHVVGQLDDIRKNPRKFVCLNDNIDHNHKDAQTVKAV LRDFYESMFPIPSQFELPREYRNRFLHMHELQEWRAYRDKLKFWTHCVLA TLIMFTIFSFFAEQLIALKRKIFPRRRIHKEASPNRIRV[SEQ ID NO:374]		
GNPTG	84572	0090581	Q9UJJ9	MAAGLARLLLLLGLSAGGPAPAGAAKMKVVEEPNAFGVNNPFLPQASRLQ AKRDPSPVSGPVHLFRLSGKCFSLVESTYKYEFCPFHNVTQHEQTFRWN AYSGILGIWHEWEIANNTFTGMWMRDGDACRSRSRQSKVELACGKSNRL AHVSEPSTCVYALTFETPLVCHPHALLVYPTLPEALQRQWDQVEQDLADE LITPQGHEKLLRTLFEDAGYLKTPEENEPTQLEGGPDSLGFETLENCRKAH KELSKEIKRLKGLLTQHGIPYTRPTETSNLEHLGHETPRAKSPEQLRGDPG LRGSL [SEQ ID NO:375]	Mucolipidose III gama	Distúrbio de Armazenamento Lisossômico
GNS	2799	0135677	A0A024RBC5, P15586, Q7Z3X3	MRLLPLAPGRLRRGSPRHLPSCSPALLLLVLGGCLGVFGVAAGTRRPNVV LLLTDDQDEVLGGMTPLKKTKALIGEMGMTFSSAYVPSALCCPSRASILTG KYPHNHHVVNNTLEGNCSSKSWQKIQEPNTFPAILRSMCGYQTFFAGKYL NEYGAPDAGGLEHVPLGWSYWYALEKNSKYYNYTLSINGKARKHGENYS VDYLTDVLANVSLDFLDYKSNFEPFFMMIATPAPHSPWTAAPQYQKAFQN VFAPRNKNFNIHGTNKHWLIRQAKTPMTNSSIQFLDNAFRKRWQTLLSVD DLVEKLVKRLEFTGELNNTYIFYTSDNGYHTGQFSLPIDKRQLYEFDIKVPL	Mucopolissacaridose tipo IIID	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LVRGPGIKPNQTSKMLVANIDLGPTILDIAGYDLNKTQMDGMSLLPILRGAS NLTWRSDVLVEYQGEGRNVTDPTCPSLSPGVSQCFPDCVCEDAYNNTYA CVRTMSALWNLQYCEFDDQEVFVEVYNLTADPDQITNIAKTIDPELLGKMN YRLMMLQSCSGPTCRTPGVFDPGYRFDPRLMFSNRGSVRTRRFSKHLL [SEQ ID NO:376]		
GRN	2896	0030582	P28799	MWTLVSWVALTAGLVAGTRCPDGQFCPVACCLDPGGASYSCCRPLLDK WPTTLSRHLGGPCQVDAHCSAGHSCIFTVSGTSSCCPFPEAVACGDGHH CCPRGFHCSADGRSCFQRSGNNSVGAIQCPDSQFECPDFSTCCVMVDG SWGCCPMPQASCCEDRVHCCPHGAFCDLVHTRCITPTGTHPLAKKLPAQ RTNRAVALSSSVMCPDARSRCPDGSTCCELPSGKYGCCPMPNATCCSD HLHCCPQDTVCDLIQSKCLSKENATTDLLTKLPAHTVGDVKCDMEVSCPD GYTCCRLQSGAWGCCPFTQAVCCEDHIHCCPAGFTCDTQKGTCEQGPH QVPWMEKAPAHLSLPDPQALKRDVPCDNVSSCPSSDTCCQLTSGEWGC CPIPEAVCCSDHQHCCPQGYTCVAEGQCQRGSEIVAGLEKMPARRASLS HPRDIGCDQHTSCPVGQTCCPSLGGSWACCQLPHAVCCEDRQHCCPAG YTCNVKARSCEKEVVSAQPATFLARSPHVGVKDVECGEGHFCHDNQTCC RDNRQGWACCPYRQGVCCADRRHCCPAGFRCAARGTKCLRREAPRWD APLRDPALRQLL [SEQ ID NO:377]	Lipofuscinose ceroide neuronal 11 (CLN11), demência frontotemporal	Distúrbio de Armazenamento Lisossômico
GUSB	2990	0169919	P08236	MARGSAVAWAALGPLLWGCALGLQGGMLYPQESPSRECKELDGLWSFR ADFSDNRRRGFEEQWYRRPLWESGPTVDMPVPSSFNDISQDWRLRHFV GWVWYEREVILPERWTQDLRTRVVLRIGSAHSYAIVWVNGVDTLEHEGG YLPFEADISNLVQVGPLPSRLRITIAINNTLTPTTLPPGTIQYLTDTSKYPKGY FVQNTYFDFFNYAGLQRSVLLYTTPTTYIDDITVTTSVEQDSGLVNYQISVK GSNLFKLEVRLLDAENKVVANGTGTQGQLKVPGVSLWWPYLMHERPAYL YSLEVQLTAQTSLGPVSDFYTLPVGIRTVAVTKSQFLINGKPFYFHGVNKH EDADIRGKGFDWPLLVKDFNLLRWLGANAFRTSHYPYAEEVMQMCDRYG IVVIDECPGVGLALPQFFNNVSLHHHMQVMEEVVRRDKNHPAVVMWSVA NEPASHLESAGYYLKMVIAHTKSLDPSRPVTFVSNSNYAADKGAPYVDVIC LNSYYSWYHDYGHLELIQLQLATQFENWYKKYQKPIIQSEYGAETIAGFHQ DPPLMFTEEYQKSLLEQYHLGLDQKRRKYVVGELIWNFADFMTEQSPTRV LGNKKGIFTRQRQPKSAAFLLRERYWKIANETRYPHSVAKSQCLENSLFT	Mucopolissacaridose tipo VII	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				[SEQ ID NO:378]		
HEXA	3073	0213614	A0A0S2Z3W3, P06865, B4DVA7, H3BP20	MTSSRLWFSLLLAAAFAGRATALWPWPQNFQTSDQRYVLYPNNFQFQYD VSSAAQPGCSVLDEAFQRYRDLLFGSGSWPRPYLTGKRHTLEKNVLVVS VVTPGCNQLPTLESVENYTLTINDDQCLLLSETVWGALRGLETFSQLVWK SAEGTFFINKTEIEDFPRFPHRGLLLDTSRHYLPLSSILDTLDVMAYNKLNV FHWHLVDDPSFPYESFTFPELMRKGSYNPVTHIYTAQDVKEVIEYARLRGI RVLAEFDTPGHTLSWGPGIPGLLTPCYSGSEPSGTFGPVNPSLNNTYEFM STFFLEVSSVFPDFYLHLGGDEVDFTCWKSNPEIQDFMRKKGFGEDFKQL ESFYIQTLLDIVSSYGKGYVVWQEVFDNKVKIQPDTIIQVWREDIPVNYMKE LELVTKAGFRALLSAPWYLNRISYGPDWKDFYIVEPLAFEGTPEQKALVIG GEACMWGEYVDNTNLVPRLWPRAGAVAERLWSNKLTSDLTFAYERLSHF RCELLRRGVQAQPLNVGFCEQEFEQT [SEQ ID NO:379]	Doença de Tay-Sachs	Distúrbio de Armazenamento Lisossômico
HEXB	3074	0049860	A0A024RAJ6, P07686, Q5URX0	MELCGLGLPRPPMLLALLLATLLAAMLALLTQVALVVQVAEAARAPSVSAK PGPALWPLPLSVKMTPNLLHLAPENFYISHSPNSTAGPSCTLLEEAFRRYH GYIFGFYKWHHEPAEFQAKTQVQQLLVSITLQSECDAFPNISSDESYTLLV KEPVAVLKANRVWGALRGLETFSQLVYQDSYGTFTINESTIIDSPRFSHRGI LIDTSRHYLPVKIILKTLDAMAFNKFNVLHWHIVDDQSFPYQSITFPELSNKG SYSLSHVYTPNDVRMVIEYARLRGIRVLPEFDTPGHTLSWGKGQKDLLTP CYSRQNKLDSFGPINPTLNTTYSFLTTFFKEISEVFPDQFIHLGGDEVEFKC WESNPKIQDFMRQKGFGTDFKKLESFYIQKVLDIIATINKGSIVWQEVFDDK AKLAPGTIVEVWKDSAYPEELSRVTASGFPVILSAPWYLDLISYGQDWRKY YKVEPLDFGGTQKQKQLFIGGEACLWGEYVDATNLTPRLWPRASAVGER LWSSKDVRDMDDAYDRLTRHRCRMVERG IAAQPLYAGYCNHENM [SEQ ID NO:380]	Doença de Sandhoff	Distúrbio de Armazenamento Lisossômico
HGS-NAT	138050	0165102	Q68CP4, Q8IVU6	MTGARASAAEQRRAGRSGQARAAERAAGMSGAGRALAALLLAASVLSAA LLAPGGSSGRDAQAAPPRDLDKKRHAELKMDQALLLIHNELLWTNLTVYW KSECCYHCLFQVLVNVPQSPKAGKPSAAAASVSTQHGSILQLNDTLEEKE VCRLEYRFGEFGNYSLLVKNIHNGVSEIACDLAVNEDPVDSNLPVSIAFLIG LAVIIVISFLRLLLSLDDFNNWISKAISSRETDRLINSELGSPSRTDPLDGDVQ PATWRLSALPPRLRSVDTFRGIALILMVFVNYGGGKYWYFKHASWNGLTV	Mucopolissacari-dose tipo IIIC	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ADLVFPWFVFIMGSSIFLSMTSILQRGCSKFRLLGKIAWRSFLLICIGIIIVNP NYCLGPLSWDKVRIPGVLQRLGVTYFVVAVLELLFAKPVPEHCASERSCLS LRDITSSWPQWLLILVLEGLWLGLTFLLPVPGCPTGYLGPGGIGDFGKYPN CTGGAAGYIDRLLLGDDHLYQHPSSAVLYHTEVAYDPEGILGTINSIVMAFL GVQAGKILLYYKARTKDILIRFTAWCC		
				ILGLISVALTKVSENEGFIPVNKNLWSLSYVTTLSSFAFFILLVLYPVVDVKGL WTGTPFFYPGMNSILVYVGHEVFENYFPFQWKLKDNQSHKEHLTQNIVAT ALWVLIAYILYRKKIFWKI [SEQ ID NO:381]		
HYAL1	3373	0114378	A0A024R2X3, Q12794, B3KUI5, A0A0S2Z3Q0	MAAHLLPICALFLTLLDMAQGFRGPLLPNRPFTTVWNANTQWCLERHGVD VDVSVFDVVANPGQTFRGPDMTIFYSSQLGTYPYYTPTGEPVFGGLPQNA SLIAHLARTFQDILAAIPAPDFSGLAVIDWEAWRPRWAFNWDTKDIYRQRS RALVQAQHPDWPAPQVEAVAQDQFQGAARAWMAGTLQLGRALRPRGL WGFYGFPDCYNYDFLSPNYTGQCPSGIRAQNDQLGWLWGQSRALYPSIY MPAVLEGTGKSQMYVQHRVAEAFRVAVAAGDPNLPVLPYVQIFYDTTNHF LPLDELEHSLGESAAQGAAGVVLWVSWENTRTKESCQAIKEYMDTTLGPF ILNVTSGALLCSQ	Mucopolissacaridose tipo IX	Distúrbio de Armazenamento Lisossômico
				ALCSGHGRCVRRTSHPKALLLLNPASFSIQLTPGGGPLSLRGALSLEDQA QMAVEFKCRCYPGWQAPWCERKSMW [SEQ ID NO:382]		
IDS	3423	0010404	P22304, B4DGD7	MPPPRTGRGLLWLGLVLSSVCVALGSETQANSTTDALNVLLIIVDDLRPSL GCYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTGRRPDTTR LYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHPGISSNHTDDSPY SWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTLPDKQ STEQAIQLLEKMKTSASPFFLAVGYHKPHIPFRYPKEFQKLYPLENITLAPD PEVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPIPVDFQRKIRQSYFAS VSYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDV ATHVPLIFYVPGRTASLPEAGEKLFPYLDPFDSASQLMEPGRQSMDLVELV SLFPTLAGLAGLQVPPRCPVPSFHVELCREGKNLLKHFRFRDLEEDPYLP GNPRELIAYSQYPRPSDIPQWNSDKPSLKDIKIMGYSIRTIDYRYTVWVGF NPDEFLANFSDIHAGELYFVDSDPLQDHNMYNDSQGGDLFQLLMP [SEQ ID NO:383]	Mucopolissacaridose tipo II	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
IDUA	3425	0127415	P35475	MRPLRPRAALLALLASLLAAPPVAPAEAPHLVHVDAARALWPLRRFWRSTGFCPPLPHSQADQYVLSWDQQLNLAYVGAVPHRGIKQVRTHWLLELVTTRGSTGRGLSYNFTHLDGYLDLLRENQLLPGFELMGSASGHFTDFEDKQQVFEWKDLVSSLARRYIGRYGLAHVSKWNFETWNEPDHHDFDNVSMTMQGFLNYYDACSEGLRAASPALRLGGPGDSFHTPPRSPLSWGLLRHCHDGTNFFTGEAGVRLDYISLHRKGARSSISILEQEKVVAQQIRQLFPKFADTPIYNDEADPLVGWSLPQPWRADVTYAAMVVKVIAQHQNLLLANTTSAFPYALLSNDNAFLSYHPHPFAQRTLTARFQVNNTRPPHVQLLRKPVLTAMGLLALLDEEQLWAEVSQAGTVLDSNHTVGVLASAHRPQGPADAWRAAVLIYASDDTRAHPNRSVAVTLRLRGVPPGPGLVYVTRYLDNGLCSPDGEWRRLGRPVFPTAEQFRRMRAAEDPVAAAPRPLPAGGRLTLRPALRLPSLLLVHVCARPEKPPGQVTRLRALPLTQGQLVLVWSDEHVGSKCLWTYEIQFSQDGKAYTPVSRKPSTFNLFVFSPDTGAVSGSYRVRALDYWARPGPFSDPVPYLEVPVPRGPPSPGNP [SEQ ID NO:384]	Mucopolissacaridose tipo I	Distúrbio de Armazenamento Lisossômico
KCTD7	154881	0243335	Q96MP8, A0A024RDN7	MVVVTGREPDSRRQDGAMSSSDAEDDFLEPATPTATQAGHALPLLPQEFPEVVPLNIGGAHFTTRLSTLRCYEDTMLAAMFSGRHYIPTDSEGRYFIDRDGTHFGDVLNFLRSGDLPPRERVRAVYKEAQYYAIGPLLEQLENMQPLKGEKVRQAFLGLMPYYKDHLERIVEIARLRAVQRKARFAKLKVCVFKEEMPITPYECPLLNSLRFERSESDGQLFEHHCEVDVSFGPWEAVADVYDLLHCLVTDLSAQGLTVDHQCIGVCDKHLVNHYYCKRPIYEFKITWW [SEQ ID NO:385]	Lipofuscinose ceroide neuronal 14 (CLN14)	Distúrbio de Armazenamento Lisossômico
LAMP2	3920	0005893	P13473	MVCFRLFPVPGSGLVLVCLVLGAVRSYALELNLTDSENATCLYAKWQMNFTVRYETTNKTYKTVTISDHGTVTYNGSICGDDQNGPKIAVQFGPGFSWIANFTKAASTYSIDSVSFSYNTGDNTTFPDAEDKGILTVDELLAIRIPLNDLFRCNSLSTLEKNDVVQHYWDVLVQAFVQNGTVSTNEFLCDKDKTSTVAPTIHTTVPSPTTTPTPKEKPEAGTYSVNNGNDTCLLATMGLQLNITQDKVASVININPNTTHSTGSCRSHTALLRLNSSTIKYLDFVFAVKNENRFYLKEVNISMYLVNGSVFSIANNNLSYWDAPLGSSYMCNKEQTVSVSGAFQINTFDLRVQPFNVTQGKYSTAQDCSADDDNFLVPIAVGAALAGVLILVLLAYFIGLKHHHAGYEQF [SEQ ID NO:386]	Doença de Danon	Distúrbio de Armazenamento Lisossômico
MAN2	4125	0104774	O00754, A8K6A7	MGAYARASGVCARGCLDSAGPWTMSRALRPPLPPLCFFLLLLAAAGARA	alfa-manosidose	Distúrbio de

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
B1				GGYETCPTVQPNMLNVHLLPHTHDDVGWLKTVDQYFYGIKNDIQHAGVQ YILDSVISALLADPTRRFIYVEIAFFSRWWHQQTNATQEVVRDLVRQGRLE FANGGWVMNDEAATHYGAIVDQMTLGLRFLEDTFGNDGRPRVAWHIDPF GHSREQASLFAQMGFDGFFFGRLDYQDKWVRMQKLEMEQVWRASTSLK PPTADLFTGVLPNGYNPPRNLCWDVLCVDQPLVEDPRSPEYNAKELVDYF LNVATAQGRYYRTNHTVMTMGSDFQYENANMWFKNLDKLIRLVNAQQAK GSSVHVLYSTPACYLWELNKANLTWSVKHDDFFPYADGPHQFWTGYFSS RPALKRYERLSYNFLQVCNQLEALVGLAANVGPYGSGDSAPLNEAMAVL QHHDAVSGTSRQHVANDYARQLAAGWGPCEVLLSNALARLRGFKDHFTF CQQLNISICPLSQTAARFQVIVYNPLGRKVNWMVRLPVSEGVFVVKDPNG RTVPSDVVIFPSSDSQAHPPELLFSASLPALGFSTYSVAQVPRWKPQARA PQPIPRRSWSPALTIENEHIRATFDPDTGLLMEIMNMNQQLLLPVRQTFFW YNASIGDNESDQASGAYIFRPNQQKPLPVSRWAQIHLVKTPLVQEVHQNF SAWCSQVVRLYPGQRHLELEWSVGPIPVGDTWGKEVISRFDTPLETKGR FYTDSNGREILERRRDYRPTWKLNQTEPVAGNYYPVNTRIYITDGNMQLT VLTDRSQGGSSLRDGSLELMVHRRLLKDDGRGVSEPLMENGSGAWVRG RHLVLLDTAQAAAAGHRLLAEQEVLAPQVVLAPGGGAAYNLGAPPRTQFS GLRRDLPPSVHLLTLASWGPEMVLLRLEHQFAVGEDSGRNLSAPVTLNLR DLFSTFTITRLQETTLVANQLREAASRLKWTTNTGPTPHQTPYQLDPANITL EPMEIRTFLASVQWKEVDG [SEQ ID NO:387]		
MAN-BA	4126	0109323	O00462	MRLHLLLLLALCGAGTTAAELSYSLRGNWSICNGNGSLELPGAVPGCVHSAL FQQGLIQDSYYRFNDLNYRWVSLDNWTYSKEFKIPFEISKWQKVNLILEGVDT VSKILFNEVTIGETDNMFNRYSFDITNVVRDVNSIELRFQSAVLYAAQQSKAHT RYQVPPDCPPLVQKGECHVNFVRKEQCSFSWDWGPSFPTQGIWKDVRIEA YNICHLNYFTFSPIYDKSAQEWNLEIESTFDVVSSKPVGGQVIVAIPKLQTQQT YSIELQPGKRIVELFVNISKNITVETWWPHGHGNQTGYNMTVLFELDGGLNIE KSAKVYFRTVELIEEPIKGSPGLSFYFKINGFPIFLKGSNWIPADSFQDRVTSEL LRLLLQSVVDANMNTLRVWGGGIYEQDEFYELCDELGIMVWQDFMFACALY PTDQGFLDSVTAEVAYQIKRLKSHPSIIIWSGNNENEEALMMNWYHISFTDRPI YIKDYVTLYVKNIRELVLAGDKSRPFITSSPTNGAETVAEAWVSQNPNSNYFG DVHFYDYISDC	beta-manosidose	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				WNWKVFPKARFASEYGYQSWPSFSTLEKVSSTEDWSFNSKFSLHRQHHEG GNKQMLYQAGLHFKLPQSTDPLRTFKDTIYLTQVMQAQCVKTETEFYRRSRS EIVDQQGHTMGALYWQLNDIWQAPSWASLEYGGKWKMLHYFAQNFFAPLL PVGFENENTFYIYGVSDLHSDYSMTLSVRVHTWSSLEPVCSRVTERFVMKG GEAVCLYEEPVSELLRRCGNCTRESCVVSFYLSADHELLSPTNYHFLSSPKE AVGLCKAQITAIISQQGDIFVFDLETSAVAPFVWLDVGSIPGRFSDNGFLMTEK TRTILFYPWEPTSKNELEQSFHVTSLTDIY [SEQ ID NO:388]		
MCOLN1	57192	0090674	Q9GZU1	MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFF MSPCDKFRAKGRKPCKLMLQVVKILVVTVQLILFGLSNQLAVTFREENTIAF RHLFLLGYSDGADDTFAAYTREQLYQAIFHAVDQYLALPDVSLGRYAYVR GGGDPWTNGSGLALCQRYYHRGHVDPANDTFDIDPMVVTDCIQVDPPER PPPPPSDDLTLLESSSSYKNLTLKFHKLVNVTIHFRLKTINLQSLINNEIPDC YTFSVLITFDNKAHSGRIPISLETQAHIQECKHPSVFQHGDNSFRLLFDVVVI LTCSLSFLLCARSLLRGFLLQNEFVGFMWRQRGRVISLWERLEFVNGWYI LLVTSDVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIRYLTFFHN YNILIATLRVALPSVMRFCCCVAVIYLGYCFCGWIVLGPYHVKFRSLSMVSE CLFSLINGDDMFVTFAAMQAQQGRSSLVWLFSQLYLYSFISLFIYMVLSLFI ALITGAYDTIKHPGGAGAEESELQAYIAQCQDSPTSGKFRRGSGSACSLLC CCGRDPSEEHSLLVN [SEQ ID NO:389]	Mucolipidose tipo IV	Distúrbio de Armazenamento Lisossômico
MFSD8	256471	0164073	Q8NHS3	MAGLRNESEQEPLLGDTPGSREWDILETEEHYKSRWRSIRILYLTMFLSSV GFSVVMMSIWPYLQKIDPTADTSFLGWVIASYSLGQMVASPIFGLWSNYR PRKEPLIVSILISVAANCLYAYLHIPASHNKYYMLVARGLLGIGAGNVAVVRS YTAGATSLQERTSSMANISMCQALGFILGPVFQTCFTFLGEKGVTWDVIKL QINMYTTPVLLSAFLGILNIILILAILREHRVDDSGRQCKSINFEEASTDEAQV PQGNIDQVAVVAINVLFFVTLFIFALFETIITPLTMDMYAWTQEQAVLYNGIIL AALGVEAVVIFLGVKLLSKKIGERAILLGGLIVVWVGFFILLPWGNQFPKIQW EDLHNNSIPNTTFGEIIIGLWKSPMEDDNERPTGCSIEQAWCLYTPVIHLAQ FLTSAVLIGLGYPVCNLMSYTLYSKILGPKPQGVYMGWLTASGSGARILGP MFISQVYAHWGPRWAFSLVCGIIVLTITLLGVVYKRLIALSVRYGRIQE[SEQ ID NO:390]	Lipofuscinose ceroide neuronal 7 (CLN7)	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
NAGA	4668	0198951	A0A024R1Q5, P17050	MLLKTVLLLGHVAQVLMLDNGLLQTPPMGWLAWERFRCNINCDEDPKNCI SEQLFMEMADRMAQDGWRDMGYTYLNIDDCWIGGRDASGRLMPDPKRF PHGIPFLADYVHSLGLKLGIYADMGNFTCMGYPGTTLDKVVQDAQTFAEW KVDMLKLDGCFSTPEERAQGYPKMAAALNATGRPIAFSCSWPAYEGGLP PRVNYSLLADICNLWRNYDDIQDSWWSVLSILNWFVEHQDILQPVAGPGH WNDPDMLLIGNFGLSLEQSRAQMALWTVLAAPLLMSTDLRTISAQNMDIL QNPLMIKINQDPLGIQGRRIHKEKSLIEVYMRPLSNKASALVFFSCRTDMPY RYHSSLGQLNFTGSVIYEAQDVYSGDIISGLRDETNFTVIINPSGVVMWYLY PIKNLEMSQQ[SEQ ID NO:391]	Doença de Schindler	Distúrbio de Armazenamento Lisossômico
NAGLU	4669	0108784	A0A140VJE4, P54802	MEAVAVAAAVGVLLLAGAGGAAGDEAREAAAVRALVARLLGPGPAADFS VSVERALAAKPGLDTYSLGGGGAARVRVRGSTGVAAAAGLHRYLRDFCG CHVAWSGSQLRLPRPLPAVPGELTEATPNRYRYYQNVCTQSYSFVWWD WARWEREIDWMALNGINLALAWSGQEAIWQRVYLALGLTQAEINEFFTGP AFLAWGRMGNLHTWDGPLPPSWHIKQLYLQHRVLDQMRSFGMTPVLPA FAGHVPEAVTRVFPQVNVTKMGSWGHFNCSYSCSFLLAPEDPIFPIIGSLF LRELIKEFGTDHIYGADTFNEMQPPSSEPSYLAAATTAVYEAMTAVDTEAV WLLQGWLFQHQPQFWGPAQIRAVLGAVPRGRLLVLDLFAESQPVYTRTA SFQGQPFIWCMLHNFGGNHGLFGALEAVNGGPEAARLFPNSTMVGTGM APEGISQNEVVYSLMAELGWRKDPVPDLAAWVTSFAARRYGVSHPDAGA AWRLLLRSVYNCSGEACRGHNRSPLVRRPSLQMNTSIWYNRSDVFEAW RLLLTSAPSLATSPAFRYDLLDLTRQAVQELVSLYYEEARSAYLSKELASLL RAGGVLAYELLPALDEVLASDSRFLLGSWLEQARAAAVSEAEADFYEQNS RYQLTLWGPEGNILDYANKQLAGLVANYYTPRWRLFLEALVDSVAQGIPF QQHQFDKNVFQLEQAFVLSKQRYPSQPRGDTVDLAKKIFLKYYPRWVAG SW [SEQ ID NO:392]	Mucopolissacaridose IIIB	Distúrbio de Armazenamento Lisossômico
NEU1	4758	0204386,0227315,0227129,0223957,0234846,0184494,0228691,0234343	Q5JQI0, Q99519	MTGERPSTALPDRRWGPRILGFWGGCRVWVFAAIFLLLSLAASWSKAEN DFGLVQPLVTMEQLLWVSGRQIGSVDTFRIPLITATPRGTLLAFAEARKMS SSDEGAKFIALRRSMDQGSTWSPTAFIVNDGDVPDGLNLGAVVSDVETGV VFLFYSLCAHKAGCQVASTMLVWSKDDGVSWSTPRNLSLDIGTEVFAPGP GSGIQKQREPRKGRLIVCGHGTLERDGVFCLLSDDHGASWRYGSGVSGI PYGQPKQENDFNPDECQPYELPDGSVVINARNQNNYHCHCRIVLRSYDA	Mucolipidose tipo I, Sialidose I	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				CDTLRPRDVTFDPELVDPVVAAGAVVTSSGIVFFSNPAHPEFRVNLTLRW SFSNGTSWRKETVQLWPGPSGYSSLATLEGSMDGEEQAPQLYVLYEKG RNHYTESISVAKISVYGTL [SEQ ID NO:393]		
NPC1	4864	0141458	O15118	MTARGLALGLLLLLLCPAQVFSQSCVWYGECGIAYGDKRYNCEYSGPPKP LPKDGYDLVQELCPGFFFGNVSLCCDVRQLQTLKDNLQLPLQFLSRCPSC FYNLLNLFCELTCSPRQSQFLNVTATEDYVDPVTNQTKTNVKELQYYVGQ SFANAMYNACRDVEAPSSNDKALGLLCGKDADACNATNWIEYMFNKDNG QAPFTITPVFSDFPVHGMEPMNNATKGCDESVDEVTAPCSCQDCSIVCGP KPQPPPPPAPWTILGLDAMYVIMWITYMAFLLVFFGAFFAVWCYRKRYFV SEYTPIDSNIAFSVNASDKGEASCCDPVSAAFEGCLRRLFTRWGSFCVRN PGCVIFFSLVFITACSSGLVFVRVTTNPVDLWSAPSSQARLEKEYFDQHFG PFFRTEQLIIRAPLTDKHIYQPYPSGADVPFGPPLDIQILHQVLDLQIAIENIT ASYDNETVTLQDICLAPLSPYNTNCTILSVLNYFQNSHSVLDHKKGDDFFV YADYHTHFLYCVRAPASLNDTSLLHDPCLGTFGGPVFPWLVLGGYDDQN YNNATALVITFPVNNYYNDTEKLQRAQAWEKEFINFVKNYKNPNLTISFTAE RSIEDELNRESDSDVFTVVISYAIMFLYISLALGHMKSCRRLLVDSKVSLGIA GILIVLSSVACSLGVFSYIGLPLTLIVIEVIPFLVLAVGVDNIFILVQAYQRDER LQGETLDQQLGRVLGEVAPSMFLSSFSETVAFFLGALSVMPAVHTFSLFA GLAVFIDFLLQITCFVSLLGLDIKRQEKNRLDIFCCVRGAEDGTSVQASESC LFRFFKNSYSPLLLKDWMRPIVIAIFVGVLSFSIAVLNKVDIGLDQSLSMPDD SYMVDYFKSISQYLHAGPPVYFVLEEGHDYTSSKGQNMVCGGMGCNND SLVQQIFNAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRVDNITDQFC NASVVDPACVRCRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGG HAAYSSAVNILLGHGTRVGATYFMTYHTVLQTSADFIDALKKARLIASNVTE TMGINGSAYRVFPYSVFYVFYEQYLTIIDDTIFNLGVSLGAIFLVTMVLLGCE LWSAVIMCATIAMVLVNMFGVMWLWGISLNAVSLVNLVMSCGISVEFCSHI TRAFTVSMKGSRVERAEEALAHMGSSVFSGITLTKFGGIVVLAFAKSQIFQI FYFRMYLAMVLLGATHGLIFLPVLLSYIGPSVNKAKSCATEERYKGTERER LLNF [SEQ ID NO:394]		
NPC2	10577	0119655	A0A024R6C0, P61916, G3V3E8	MRFLAATFLLLALSTAAQAEPVQFKDCGSVDGVIKEVNVSPCPTQPCQLS KGQSYSVNVTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPDGCKSGINCPI	Niemann-Pick tipo C	Distúrbio de Armazenamento

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QKDKTYSYLNKLPVKSEYPSIKLVVEWQLQDDKNQSLFCWEIPVQIVSHL [SEQ ID NO:395]		Lisossômico
SGSH	6448	0181523	P51688	MSCPVPACCALLLVLGLCRARPRNALLLLADDGGFESGAYNNSAIATPHLD ALARRSLLFRNAFTSVSSCSPSRASLLTGLPQHQNGMYGLHQDVHHFNSF DKVRSLPLLLSQAGVRTGIIGKKHVGPETVYPFDFAYTEENGSVLQVGRNI TRIKLLVRKFLQTQDDRPFFLYVAFHDPHRCGHSQPQYGTFCEKFGNGES GMGRIPDWTPQAYDPLDVLVPYFVPNTPAARADLAAQYTTVGRMDQGVG LVLQELRDAGVLNDTLVIFTSDNGIPFPSGRTNLYWPGTAEPLLVSSPE HPKRWGQVSEAYVSLLDLTPTILDWFSIPYPSYAIFGSKTIHLTGRSLLPAL EAEPLWATVFGSQSHHEVTMSYPMRSVQHRHFRLVHNLNFKMPFPIDQD FYVSPTFQDLLNRTTAGQPTGWYKDLRHYYYRARWELYDRSRDPHETQN LATDPRFAQLLEMLRDQLAKWQWETHDPWVCAPDGVLEEKLSPQCQPL HNEL [SEQ ID NO:396]	Mucopolissacaridose IIIA	Distúrbio de Armazenamento Lisossômico
PPT1	5538	0131238	P50897	MASPGCLWLLAVALLPWTCASRALQHLDPPAPLPLVIWHGMGDSCCNPL SMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVNSQVTTVCQALAK DPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHQGVFGLP RCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPIKEDVYRNHSIF LADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQ AKETIPLQETSLYTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHII PFLG [SEQ ID NO:397]	Lipofuscinose ceroide neuronal 1 (CLN1)	Distúrbio de Armazenamento Lisossômico
PSAP	5660	0197746	P07602, A0A024QZQ2	MYALFLLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCL QTVWNKPTVKSLPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPK PNMSASCKEIVDSYLPVILDIIKGEMSRPGEVCSALNLCESLQKHLAELNHQ KQLESNKIPELDMTEVVAPFMANIPLLLYPQDGPRSKPQPKDNGDVCQDCI QMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQYSEIAI QMMMHMQPKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPALELVEPIK KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCSKLPKSLS EECQEVVDTYGSSILSILLEEVSPELVCSMLHLCSGTRLPALTVHVTQPKD GGFCEVCKKLVGYLDRNLEKNSTKQEILAALEKGCSFLPDPYQKQCDQFV AEYEPVLIEILVEVMDPSFVCLKIGACPSAHKPLLGTEKCIWGPSYWCQNT	Deficiência de prosaposina, deficiência de SapA (variante de Krabbe), deficiência de SapB (variante de MLD), deficiência de SapC (variante de Gaucher)	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ETAAQCNAVEHCKRHVWN [SEQ ID NO:398]		
SLC17A5	26503	0119899	Q9NRA2	MRSPVRDLARNDGEESTDRTPLLPGAPRAEAAPVCCSARYNLAILAFFGF FIVYALRVNLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHHNQTGKKY QWDAETQGWILGSFFYGYIITQIPGGYVASKIGGKMLLGFGILGTAVLTLFT PIAADLGVGPLIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLLSISYA GAQLGTVISLPLSGIICYYMNWTYVFYFFGTIGIFWFLLWIWLVSDTPQKHK RISHYEKEYILSSLRNQLSSQKSVPWVPILKSLPLWAIVVAHFSYNWTFYTL LTLLPTYMKEILRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKWNFS TLCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFLTISTTLGGFCSSGFSIN HLDIAPSYAGILLGITNTFATIPGMVGPVIAKSLTPDNTVGEWQTVFYIAAAI NVFGAIFFTLFAKGEVQNWALNDHHGHRH [SEQ ID NO:399]	Doença do armazenamento de ácido siálico infantil, doença de Salla	Distúrbio de Armazenamento Lisossômico
SMPD1	6609	0166311	P17405, Q59EN6, E9LUE8, Q8IUN0, E9LUE9	MPRYGASLRQSCPRSGREQGQDGTAGAPGLLWMGLVLALALALALALAL SDSRVLWAPAEAHPLSPQGHPARLHRIVPRLRDVFGWGNLTCPICKGLFT AINLGLKKEPNVARVGSVAIKLCNLLKIAPPAVCQSIVHLFEDDMVEVWRRS VLSPSEACGLLLGSTCGHWDIFSSWNISLPTVPKPPPKPPSPPAPGAPVS RILFLTDLHWDHDYLEGTDPDCADPLCCRRGSGLPPASRPGAGYWGEYS KCDLPLRTLESLLSGLGPAGPFDMVYWTGDIPAHDVWHQTRQDQLRALT TVTALVRKFLGPVPVYPAVGNHESTPVNSFPPPFIEGNHSSRWLYEAMAK AWEPWLPAEALRTLRIGGFYALSPYPGLRLISLNMNFCSRENFWLLINSTD PAGQLQWLVGELQAAEDRGDKVHIIGHIPPGHCLKSWSWNYYRIVARYEN TLAAQFFGHTHVDEFEVFYDEETLSRPLAVAFLAPSATTYIGLNPGYRVYQI DGNYSGSSHVVLDHETYILNLTQANIPGAIPHWQLLYRARETYGLPNTLPT AWHNLVYRMRGDMQLFQTFWFLYHKGHPPSEPCGTPCRLATLCAQLSA RADSPALCRHLMPDGSLPEAQSLWPRPLFC [SEQ ID NO:400]	Niemann Pick tipos A e B	Distúrbio de Armazenamento Lisossômico
SUMF1	285362	0144455	Q8NBK3	MAAPALGLVCGRCPELGLVLLLLLLSLLCGAAGSQEAGTGAGAGSLAGSC GCGTPQRPGAHGSSAAAHRYSREANAPGPVPGERQLAHSKMVPIPAGVF TMGTDDPQIKQDGEAPARRVTIDAFYMDAYEVSNTEFEKFVNSTGYLTEA EKFGDSFVFEGMLSEQVKTNIQQAVAAAPWWLPVKGANWRHPEGPDSTI LHRPDHPVLHVSWNDAVAYCTWAGKRLPTEAEWEYSCRGGLHNRLFPW GNKLQPKGQHYANIWQGEFPVTNTGEDGFQGTAPVDAFPPNGYGLYNIV	Deficiência múltipla de sulfatase	Distúrbio de Armazenamento Lisossômico

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GNAWEWTSDWWTVHHSVEETLNPKGPPSGKDRVKKGGSYMCHRSYCY RYRCAARSQNTPDSSASNLGFRCAADRLPTMD [SEQ ID NO:401]		
TPP1	1200	0166340	O14773	MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLT FALRQQNVERLSELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQ KWLLAAGAQKCHSVITQDFLTCWLSIRQAELLLPGAEFHHYVGGPTETHV VRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRPEPQVTGTVGLHLG VTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGR HEGQEPFLQWLMLLSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKA AARGLTLLFASGDSGAGCWSVSGRHQFRPTFPASSPYVTTVGGTSFQEP FLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSPHLPPSSYFNASGRA YPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSGRPP LGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTG WGTPNFPALLKTLLNP [SEQ ID NO:402]	Lipofuscinose ceroide neuronal 2 (CLN2)	Distúrbio de Armazenamento Lisossômico
AHCY	191	0101444	P23526, Q1RMG2	MSDKLPYKVADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARI AGCLHMTVETAVLIETLVTLGAEVQWSSCNIFSTQDHAAAAIAKAGIPVYA WKGETDEEYLWCIEQTLYFKDGPLNMILDDGGDLTNLIHTKYPQLLPGIRGI SEETTTGVHNLYKMMANGILKVPAINVNDSVTKSKFDNLYGCRESLIDGIK RATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPINALQAAME GYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIG HFDVEIDVKWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAM GHPSFVMSNSFTNQVMAQIELWTHPDKYPVGVHFLPKKLDEAVAEAHLG KLNVKLTKLTEKQAQYLGMSCDGPFKPDHYRY [SEQ ID NO:403]	Hipermetioninemia	Aminoacidopatia
GNMT	27232	0124713	A0A0S2Z5F2, Q14749, V9HW60	MVDSVYRTRSLGVAAEGLPDQYADGEAARVWQLYIGDTRSRTAEYKAWL LGLLRQHGCQRVLDVACGTGVDSIMLVEEGFSVTSVDASDKMLKYALKER WNRRHEPAFDKWVIEEANWMTLDKDVPQSAEGGFDAVICLGNSFAHLPD CKGDQSEHRLALKNIASMVRAGGLLVIDHRNYDHILSTGCAPPGKNIYYKS DLTKDVTTSVLIVNNKAHMVTLDYTVQVPGAGQDGSPGLSKFRLSYYPHC LASFTELLQAAFGGKCQHSVLGDFKPYKPGQTYIPCYFIHVLKRTD [SEQ ID NO:404]	Hipermetioninemia	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
MAT1A	4143	0151224	Q00266	MNGPVDGLCDHSLSEGVFMFTSESVGEGHPDKICDQISDAVLDAHLKQDPNAKVACETVCKTGMVLLCGEITSMAMVDYQRVVRDTIKHIGYDDSAKGFDFKTCNVLVALEQQSPDIAQCVHLDRNEEDVGAGDQGLMFGYATDETEECMPLTIILAHKLNARMADLRRSGLLPWLRPDSKTQVTVQYMQDNGAVIPVRIHTIVISVQHNEDITLEEMRRALKEQVIRAVVPAKYLDEDTVYHLQPSGRFVIGGPQGDAGVTGRKIIVDTYGGWGAHGGGAFSGKDYTKVDRSAAYAARWVAKSLVKAGLCRRVLVQVSYAIGVAEPLSISIFTYGTSQKTERELLDVVHKNFDLRPGVIVRDLDLKKPIYQKTACYGHFGRSEFPWEVPRKLVF [SEQ ID NO:405]	Hipermetioninemia	Aminoacidopatia
GCH1	2643	0131979	A0A024R642, P30793, Q8IZH9	MEKGPVRAPAEKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSAQPADGWKGERPRSEEDNELNLPNLAAAYSSILSSLGENPQRQGLLKTPWRAASAMQFFTKGYQETISDVLNDAIFDEDHDEMVIVKDIDMFSMCEHHLVPFVGKVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAGVGVVVEATHMCMVMRGVQKMNSKTVTSTMLGVFREDPKTREEFLTLIRS [SEQ ID NO:406]	Deficiência de cofator BH4	Aminoacidopatia
PCBD1	5092	0166228	P61457	MAGKAHRLSAEERDQLLPNLRAVGWNELEGRDAIFKQFHFKDFNRAFGFMTRVALQAEKLDHHPEWFNVYNKVHITLSTHECAGLSERDINLASFIEQVAVSMT [SEQ ID NO:407]	Deficiência de cofator BH4	Aminoacidopatia
PTS	5805	0150787	Q03393	MSTEGGGRRCQAQVSRRISFSASHRLYSKFLSDEENLKLFGKCNNPNGHGHNYKVVVTVHGEIDPATGMVMNLADLKKYMEEAIMQPLDHKNLDMDVPYFADVVSTTENVAVYIWDNLQKVLPVGVLYKVKVYETDNNIVVYKGE [SEQ ID NO:408]	Deficiência de cofator BH4	Aminoacidopatia
QDPR	5860	0151552	A0A140VKA9, P09417	MAAAAAAGEARRVLVYGGRGALGSRCVQAFRARNWWVASVDVVENEEAASIIVKMTDSFTEQADQVTAEVGKLLGEEKVDAILCVAGGWAGGNAKSKSLFKNCDLMWKQSIWTSTISSHLATKHLKEGGLLTLAGAKAALDGTPGMIGYGMAKGAVHQLCQSLAGKNSGMPPGAAAIAVLPVTLDTPMNRKSMPEADFSSWTPLEFLVETFHDWITGKNRPSSGSLIQVVTTEGRTELTPAYF [SEQ ID NO:409]	Deficiência de cofator BH4	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
SPR	6697	0116096	P35270	MEGGLGRAVCLLTGASRGFGRTLAPLLASLLSPGSVLVLSARNDEALRQL EAELGAERSGLRVVRVPADLGAEAGLQQLLGALRELPRPKGLQRLLLINNA GSLGDVSKGFVDLSDSTQVNNYWALNLTSMLCLTSSVLKAFPDSPGLNRT VVNISSLCALQPFKGWALYCAGKAARDMLFQVLALEEPNVRVLNYAPGPL DTDMQQLARETSVDPDMRKGLQELKAKGKLVDCKVSAQKLLSLLEKDEFK SGAHVDFYDK [SEQ ID NO:410]	Deficiência de cofator BH4	Aminoacidopatia
DNAJ C12	56521	0108176	Q6IAH1, Q9UKB3	MDAILNYRSEDTEDYYTLLGCDELSSVEQILAEFKVRALECHPDKHPENPK AVETFQKLQKAKEILTNEESRARYDHWRRSQMSMPFQQWEALNDSVKTS MHWVVRGKKDLMLEESDKTHTTKMENEECNEQRERKKEELASTAEKTEQ KEPKPLEKSVSPQNSDSSGFADVNGWHLRFRWSKDAPSELLRKFRNYEI [SEQ ID NO:411]	Deficiência de co-chaperona de choque térmico de fenilalanina, tirosi-na e triptofano hidroxilases	Aminoacidopatia
ALDH 4A1	8659	0159423	P30038, A0A024RAD8	MLLPAPALRRALLSRPWTGAGLRWKHTSSLKVANEPVLAFTQGSPERDAL QKALKDLKGRMEAIPCVVGDEEVWTSDVQYQVSPFNHGHKVAKFCYADK SLLNKAIEAALAARKEWDLKPIADRAQIFLKAADMLSGPRRAEILAKTMVGQ GKTVIQAEIDAAAELIDFFRFNAKYAVELEGQQPISVPPSTNSTVYRGLEGF VAAISPFNFTAIGGNLAGAPALMGNVVLWKPSDTAMLASYAVYRILREAGL PPNIIQFVPADGPLFGDTVTSSEHLCGINFTGSVPTFKHLWKQVAQ NLDRFHTFPRLAGECGGKNFHFVHRSADVESVVSGTLRSAFEYGGQKCS ACSRLYVPHSLWPQIKGRLLEEHSRIKVGDPAEDFGTFFSAVIDAKSFARIK KWLEHARSSPSLTILAGGKCDDSVGYFVEPCIVESKDPQEPIMKEEIFGPV LSVYVYPDDKYKETLQLVDSTTSYGLTGAVFSQDKDVVQEATKVLRNAAG NFYINDKSTGSIVGQQPFGGARASGTNDKPGGPHYILRWTSPQVIKETHK PLGDWSYAYMQ [SEQ ID NO:412]	Hiperprolinemia	Aminoacidopatia
PROD H	5625	0100033	O43272	MALRRALPALRPCIPRFVQLSTAPASREQPAAGPAAVPGGGSATAVRPPV PAVDFGNAQEAYRSRRTWELARSLLVLRLCAWPALLARHEQLLYVSRKLL GQRLFNKLMKMTFYGHFVAGEDQESIQPLLRHYRAFGVSAILDYGVEEDL SPEEAEHKEMESCTSAAERDGSGTNKRDKQYQAHRAFGDRRNGVISART YFYANEAKCDSHMETFLRCIEASGRVSDDGFIAIKLTALGRPQFLLQFSEVL AKWRCFFHQMAVEQGQAGLAAMDTKLEVAVLQESVAKLGIASRAEIEDW	Hiperprolinemia	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FTAETLGVSGTMDLLDWSSLIDSRTKLSKHLVVPNAQTGQLEPLLSRFTEEEELQMTRMLQRMDVLAKKATEMGVRLMVDAEQTYFQPAISRLTLEMQRKFNVEKPLIFNTYQCYLKDAYDNVTLDVELARREGWCFGAKLVRGAYLAQERARAAEIGYEDPINPTYEATNAMYHRCLDYVLEELKHNAKAKVMVASHNEDTVRFALRRMEELGLHPADHQVYFGQLLGMCDQISFPLGQAGYPVYKYVPYGPVMEVLPYLSRRALENSSLMKGTHRERQLLWLELLRRLRTGNLFHRPA [SEQ ID NO:413]		
HPD	3242	0158104	P32754	MTTYSDKGAKPERGRFLHFHSVTFWVGNAKQAASFYCSKMGFEPLAYRGLETGSREVVSHVIKQGKIVFVLSSALNPWNKEMGDHLVKHGDGVKDIAFEVEDCDYIVQKARERGAKIMREPWVVEQDKFGKVKFAVLQTYGDTTHTLVEKMNYIGQFLPGYEAPAFMDPLLPKLPKCSLEMIDHIVGNQPDQEMVSASEWYLKNLQFHRFWSVDDTQVHTEYSSLRSIVVANYEESIKMPINEPAPGKKKSQIQEYVDYNGGAGVQHIALKTEDIITAIRHLRERGLEFLSVPSTYYKQLREKLKTAKIKVKENIDALEELKILVDYDEKGYLLQIFTKPVQDRPTLFLEVIQRHNHQGFGAGNFNSLFKAFEEEQNLRGNLTNMETNGVVPGM [SEQ ID NO:414]	Tirosinemia tipo II	Aminoacidopatia
GBA	2629	0177628,0262446	A0A068F658, P04062, B7Z6S9	MEFSSPSREECPKPLSRVSIMAGSLTGLLLLQAVSWASGARPCIPKSFGYSSVVCVCNATYCDSFDPPTFPALGTFSRYESTRSGRRMELSMGPIQANHTGTGLLLTLQPEQKFQKVKGFGGAMTDAAALNILALSPPAQNLLLKSYFSEEGIGYNIIRVPMASCDFSIRTYTYADTPDDFQLHNFSLPEEDTKLKIPLIHRALQLAQRPVSLLASPWTSPTWLKTNGAVNGKGSLKGQP [SEQ ID NO:415] GDIYHQTWARYFVKFLDAYAEHKLQFWAVTAENEPSAGLLSGYPFQCLGFTPEHQRDFIARDLGPTLANSTHHNVRLLMLDDQRLLLPHWAKVVLTDPEAAKYVHGIAVHWYLDFLAPAKATLGETHRLFPNTMLFASEACVGSKFWEQSVRLGSWDRGMQYSHSIITNLLYHVVGWTDWNLALNPEGGPNWVRNFVDSPIIVDITKDTFYKQPMFYHLGHFSKFIPEGSQRVGLVASQKNDLDAVALMHPDGSAVVVVLNRSSKDVPLTIKDPAVGFLETISPGYSIHTYLWRRQ [SEQ ID NO:415]	Doença de Gaucher	
HGD	3081	0113924	Q93099, B3KW64	MAELKYISGFGNECSSEDPRCPGSLPEGQNNPQVCPYNLYAEQLSGSAF	Alcaptonúria	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				TCPRSTNKRSWLYRILPSVSHKPFESIDEGQVTHNWDEVDPDPNQLRWK PFEIPKASQKKVDFVSGLHTLCGAGDIKSNNGLAIHIFLCNTSMENRCFYNS DGDFLIVPQKGNLLIYTEFGKMLVQPNEICVIQRGMRFSIDVFEETRGYILE VYGVHFELPDLGPIGANGLANPRDFLIPIAWYEDRQVPGGYTVINKYQGKL FAAKQDVSPFNVVAWHGNYTPYKYNLKNFMVINSVAFDHADPSIFTVLTAK SVRPGVAIADFVIFPPRWGVADKTFRPPYYHRNCMSEFMGLIRGHYEAKQ GGFLPGGGSLHSTMTPHGPDADCFEKASKVKLAPERIADGTMAFMFESSL SLAVTKWGLKASRCLDENYHKCWEPLKSHFTPNSRNPAEPN [SEQ ID NO:416]		
AMN	81693	0166126	Q9BXJ7, B3KP64	MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAV EFPADKMVSVLVQEGHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCG AGEPAVFRDSDRFSWHDPHLWRSGDEAPGLFFVDAERVPCRHDDVFFP PSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASRAGRLRFHGP GALSVGPEDCADPSGCVCGNAEAQPWICAALLQPLGGRCPQAACHSALR PQGQCCDLCGAVVLLTHGPAFDLERYRARILDTFLGLPQYHGLQVAVSKV PRSSRLREADTEIQVVLVENGPETGGAGRLARALLADVAENGEALGVLEA TMRESGAHVWGSSAAGLAGGVAAAVLLALLVLLVAPPLLRRAGRLRWRR HEAAAPAGAPLGFRNPVFDVTASEELPLPRRLSLVPKAAADSTSHSYFVN PLFAGAEAEA [SEQ ID NO:417]	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica
CD320	51293	0167775	Q9NPF0	MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGP SSGSCPPTKFQCRTSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKG QCPPPPGLPCPCTGVSDCSGGTDKKLRNCSRLACLAGELRCTLSDDCIPL TWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVTLESVTSLRNATT MGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLLL LSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP [SEQ ID NO:418]	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica
CUBN	8029	0107611	O60494	MMNMSLPFLWSLLTLLIFAEVNGEAGELELQRQKRSINLQQPRMATERGN LVFLTGSAQNIEFRTGSLGKIKLNDEDLSECLHQIQKNKEDIIELKGSAIGLP QNISSQIYQLNSKLVDLERKFQGLQQTVDKKVCSSNPCQNGGTCLNLHDS FFCICPPQWKGPLCSADVNECEIYSGTPLSCQNGGTCVNTMGSYSCHCP PETYGPQCASKYDDCEGGSVARCVHGICEDLMREQAGEPKYSCVCDAG	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				WMFSPNSPACTLDRDECSFQPGPCSTLVQCFNTQGSFYCGACPTGWQG NGYICEDINECEINNGGCSVAPPVECVNTPGSSHCQACPPGYQGDGRVC TLTDICSVSNGGCHPDASCSSTLGSLPLCTCLPGYTGNGYGPNGCVQLSN ICLSHPCLNGQCIDTVSGYFCKCDSGWTGVNCTENINECLSNPCLNGGTC VDGVDSFSCECTRLWTGALCQVPQQVCGESLSGINGSFSYRSPDVGYVH DVNCFWVIKTEMGKVLRITFTFFRLESMDNCPHEFLQVYDGDSSSAFQLG RFCGSSLPHELLSSDNALYFHLYSEHLRNGRGFTVRWETQQPECGGILTG PYGSIKSPGYPGNYPPGRDCVWIVVTSPDLLVTFTFGTLSLEHHDDCNKD YLEIRDGPLYQDPLLGKFCTTFSVPPLQTTGPFARIHFHSDSQISDQGFHIT YLTSPSDLRCGGNYTDPEGELFLPELSGPFTHTRQCVYMMKQPQGEQIQI NFTHVELQCQSDSSQNYIEVRDGETLLGKVCGNGTISHIKSITNSVWIRFKI DASVEKASFRAVYQVACGDELTGEGVIRSPFFPNVYPGERTCRWTIHQPQ SQVILLNFTVFEIGSSAHCETDYVEIGSSSILGSPENKKYCGTDIPSFITSVY NFLYVTFVKSSSTENHGFMAKFSAEDLACGEILTESTGTIQSPGHPNVYPH GINCTWHILVQPNHLIHLMFETFHLEFHYNCTNDYLEVYDTDSETSLGRYC GKSIPPSLTSSGNSLMLVFVTDSDLAYEGFLINYEAISAATACLQDYTDDLG TFTSPNFPNNYPNNWECIYRITVRTGQLIAVHFTNFSLEEAIGNYYTDFLEIR DGGYEKSPLLGIFYGSNLPPTIISHSNKLWLKFKSDQIDTRSGFSAYWDGS STGCGGNLTTSSGTFISPNYPMPYYHSSECYWWLKSSHGSAFELEFKDF HLEHHPNCTLDYLAVYDGPSSNSHLLTQLCGDEKPPLIRSSGDSMFIKLRT DEGQQGRGFKAEYRQTCENVVIVNQTYGILESIGYPNPYSENQHCNWTIR ATTGNTVNYTFLAFDLEHHINCSTDYLELYDGPRQMGRYCGVDLPPPGST TSSKLQVLLLTDGVGRREKGFQMQWFVYGCGGELSGATGSFSSPGFPN RYPPNKECIWYIRTDPGSSIQLTIHDFDVEYHSRCNFDVLEIYGGPDFHSP RIAQLCTQRSPENPMQVSSTGNELAIRFKTDLSINGRGFNASWQAVTGGC GGIFQAPSGEIHSPNYPSPYRSNTDCSWVIRVDRNHRVLLNFTDFDLEPQ DSCIMAYDGLSSTMSRLARTCGREQLANPIVSSGNSLFLRFQSGPSRQNR GFRAQFRQACGGHILTSSFDTVSSPRFPANYPNNQNCSWIIQAQPPLNHIT LSFTHFELERSTTCARDFVEILDGGHEDAPLRGRYCGTDMPHPITSFSSAL TLRFVSDSSISAGGFHTTVTASVSACGGTFYMAEGIFNSPGYPDIYPPNVE CVWNIVSSPGNRLQLSFISFQLEDSQDCSRDFVEIREGNATGHLVGRYCG NSFPLNYSSIVGHTLWVRFISDGSGSGTGFQATFMKIFGNDNIVGTHGKVA		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SPFWPENYPHNSNYQWTVNVNASHVVHGRILEMDIEEIQNCYYDKLRIYD GPSIHARLIGAYCGTQTESFSSTGNSLTFHFYSDSSISGKGFLLEWFAVDA PDGVLPTIAPGACGGFLRTGDAPVFLFSPGWPDSYSNRVDCTWLIQAPDS TVELNILSLDIESHRTCAYDSLVIRDGDNNLAQQLAVLCGREIPGPIRSTGE YMFIRFTSDSSVTRAGFNASFHKSCGGYLHADRGIITSPKYPETYPSNLNC SWHVLVQSGLTIAVHFEQPFQIPNGDSSCNQGDYLVLRNGPDICSPPLGP PGGNGHFCGSHASSTLFTSDNQMFVQFISDHSNEGQGFKIKYEAKSLACG GNVYIHDADSAGYVTSPNHPHNYPPHADCIWILAAPPETRIQLQFEDRFDI EVTPNCTSNYLELRDGVDSDAPILSKFCGTSLPSSQWSSGEVMYLRFRSD NSPTHVGFKAKYSIAQCGGRVPGQSGVVESIGHPTLPYRDNLFCEWHLQ GLSGHYLTISFEDFNLQNSSGCEKDFVEIWDNHTSGNILGRYCGNTIPDSI DTSSNTAVVRFVTDGSVTASGFRLRFESSMEECGGDLQGSIGTFTSPNYP NPNPHGRICEWRITAPEGRRITLMFNNLRLATHPSCNNEHVIVFNGIRSNS PQLEKLCSSVNVSNEIKSSGNTMKVIFFTDGSRPYGGFTASYTSSEDAVC GGSLPNTPEGNFTSPGYDGVRNYSRNLNCEWTLSNPNQGNSSISIHFEDF YLESHQDCQFDVLEFRVGDADGPLMWRLCGPSKPTLPLVIPYSQVWIHFV TNERVEHIGFHAKYSFTDCGGIQIGDSGVITSPNYPNAYDSLTHCSSLLEAP QGHTITLTFSDFDIEPHTTCAWDSVTVRNGGSPESPIIGQYCGNSNPRTIQ SGSNQLVVTFNSDHSLQGGGFYATWNTQTLGCGGIFHSDNGTIRSPHWP QNFPENSRCSWTAITHKSKHLEISFDNNFLIPSGDGQCQNSFVKVWAGTE EVDKALLATGCGNVAPGPVITPSNTFTAVFQSQEAPAQGFSASFVSRCGS NFTGPSGYIISPNYPKQYDNNMNCTYVIEANPLSVVLLTFVSFHLEARSAVT GSCVNDGVHIIRGYSVMSTPFATVCGDEMPAPLTIAGPVLLNFYSNEQITD FGFKFSYRIISCGGVFNFSSGIITSPAYSYADYPNDMHCLYTITVSDDKVIEL KFSDFDVVPSTSCSHDYLAIYDGANTSDPLLGKFCGSKRPPNVKSSNNSM LLVFKTDSFQTAKGWKMSFRQTLGPQQGCGGYLTGSNNTFASPDSDSNG MYDKNLNCVWIIIAPVNKVIHLTFNTFALEAASTRQRCLYDYVKLYDGDSEN ANLAGTFCGSTVPAPFISSGNFLTVQFISDLTLEREGFNATYTIMDMPCGG TYNATWTPQNISSPNSSDPDVPFSICTWVIDSPPHQQVKITVWALQLTSQD CTQNYLQLQDSPQGHGNSRFQFCGRNASAVPVFYSSMSTAMVIFKSGVV NRNSRMSFTYQIADCNRDYHKAFGNLRSPGWPDNYDNDKDCTVTLTAPQ NHTISLFFHSLGIENSVECRNDFLEVRNGSNSNSPLLGKYCGTLLPNPVFS		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QNNELYLRFKSDSVTSDRGYEIIWTSSPSGCGGTLYGDRGSFTSPGYPGT YPNNTYCEWVLVAPAGRLVTINFYFISIDDPGDCVQNYLTLYDGPNASSPS SGPYCGGDTSIAPFVASSNQVFIKFHADYARRPSAFRLTWDS[SEQ ID NO:419]		
GIF	2694	0134812	P27352	MAWFALYLLSLLWATAGTSTQTQSSCSVPSAQEPLVNGIQVLMENSVTSS AYPNPSILIAMNLAGAYNLKAQKLLTYQLMSSDNNDLTIGQLGLTIMALTSS CRDPGDKVSILQRQMENWAPSSPNAEASAFYGPSLAILALCQKNSEATLPI AVRFAKTLLANSSPFNVDTGAMATLALTCMYNKIPVGSEEGYRSLFGQVL KDIVEKISMKIKDNGIIGDIYSTGLAMQALSVTPEPSKKEWNCKKTTDMILNE IKQGKFHNPMSIAQILPSLKGKTYLDVPQVTCSPDHEVQPTLPSNPGPGPT SASNITVIYTINNQLRGVELLFNETINVSVKSGSVLLVVLEEAQRKNPMFKF ETTMTSWGLVVSSINNIAENVNHKTYWQFLSGVTPLNEGVADYIPFNHEHI TANFTQY [SEQ ID NO:420]	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica
TCN1	6947	0134827	P20061	MRQSHQLPLVGLLLFSFIPSQLCEICEVSEENYIRLKPLLNTMIQSNYNRGT SAVNVVLSLKLVGIQIQTLMQKMIQQIKYNVKSRLSDVSSGELALIILALGVC RNAEENLIYDYHLIDKLENKFQAEIENMEAHNGTPLTNYYQLSLDVLALCLF NGNYSTAEVVNHFTPENKNYYFGSQFSVDTGAMAVLALTCVKKSLINGQI KADEGSLKNISIYTKSLVEKILSEKKENGLIGNTFSTGEAMQALFVSSDYYN ENDWNCQQTLNTVLTEISQGAFSNPNAAAQVLPALMGKTFLDINKDSSCV SASGNFNISADEPITVTPPDSQSYISVNYSVRINETYFTNVTVLNGSVFLSV MEKAQKMNDTIFGFTMEERSWGPYITCIQGLCANNNDRTYWELLSGGEPL SQGAGSYVVRNGENLEVRWSKY [SEQ ID NO:421]	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica
TCN2	6948	0185339	P20062	MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHL NPSIYVGLRLSSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKP SMGQLALYLLALRANCEFVRGHKGDRLVSQLKWFLEDEKRAIGHDHKGH PHTSYYQYGLGILALCLHQKRVHDSVVDKLLYAVEPFHQGHHSVDTAAMA GLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGNVYSTPLAL QFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGST VEDVLKKAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNT	Acidemia metilmalônica combinada e homocistinúria	Acidemia orgânica

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				PLLQGIADYRPKDGETIELRLVSW [SEQ ID NO:422]		
PREPL	9581	0138078	Q4J6C6	MQQKTKLFLQALKYSIPHLGKCMQKQHLNHYNFADHCYNRIKLKKYHLTK CLQNKPKISELARNIPSRSFSCKDLQPVKQENEKPLPENMDAFEKVRTKLE TQPQEEYEIINVEVKHGGFVYYQEGCCLVRSKDEEADNDNYEVLFNLEEL KLDQPFIDCIRVAPDEKYVAAKIRTEDSEASTCVIIKLSDQPVMEASFPNVS SFEWVKDEEDEDVLFYTFQRNLRCHDVYRATFGDNKRNERFYTEKDPSY FVFLYLTKDSRFLTINIMNKTTSEVWLIDGLSPWDPPVLIQKRIHGVLYYVEH RDDELYILTNVGEPTEFKLMRTAADTPAIMNWDLFFTMKRNTKVIDLDMFK DHCVLFLKHSNLLYVNVIGLADDSVRSLKLPPWACGFIMDTNSDPKNCPF QLCSPIRPPKYYTYKFAEGKLFEETGHEDPITKTSRVLRLEAKSKDGKLVP MTVFHKTDSEDLQKKPLLVHVYGAYGMDLKMNFRPERRVLVDDGWILAY CHVRGGGELGLQWHADGRLTKKLNGLADLEACIKTLHGQGFSQPSLTTLT AFSAGGVLAGALCNSNPELVRAVTLEAPFLDVLNTMMDTTLPLTLEELEE WGNPSSDEKHKNYIKRYCPYQNIKPQHYPSIHITAYENDERVPLKGIVSYT EKLKEAIAEHAKDTGEGYQTPNIILDIQPGGNHVIEDSHKKITAQIKFLYEEL GLDSTSVFEDLKKYLKF[SEQ ID NO:423]	Cistinúria	Aminoacidopatia
PHGDH	26227	0092621	O43175	MAFANLRKVLISDSLDPCCRKILQDGGLQVVEKQNLSKEELIAELQDCEGLI VRSATKVTADVINAAEKLQVVGRAGTGVDNVDLEAATRKGILVMNTPNGN SLSAAELTCGMIMCLARQIPQATASMKDGKWERKKFMGTELNGKTLGILG LGRIGREVATRMQSFGMKTIGYDPIISPEVSASFGVQQLPLEEIWPLCDFIT VHTPLLPSTTGLLNDNTFAQCKKGVRVVNCARGGIVDEGALLRALQSGQC AGAALDVFTEEPPRDRALVDHENVISCPHLGASTKEAQSRCGEEIAVQFV DMVKGKSLTGVVNAQALTSAFSPHTKPWIGLAEALGTLMRAWAGSPKGTI QVITQGTSLKNAGNCLSPAVIVGLLKEASKQADVNLVNAKLLVKEAGLNVT TSHSPAAPGEQGFGECLLAVALAGAPYQAVGLVQGTTPVLQGLNGAVFR PEVPLRRDLPLLLFRTQTSDPAMLPTMIGLLAEAGVRLLSYQTSLVSDGET WHVMGISSLLPSLEAWKQHVTEAFQFHF [SEQ ID NO:424]	Distúrbios da Biossíntese de Serina	Aminoacidopatia
PSAT1	29968	0135069	A0A024R280, Q9Y617, A0A024R222	MDAPRQVVNFGPGPAKLPHSVLLEIQKELLDYKGVGISVLEMSHRSSDFA KIINNTENLVRELLAVPDNYKVIFLQGGGCGQFSAVPLNLIGLKAGRCADYV VTGAWSAKAAEEAKKFGTINIVHPKLGSYTKIPDPSTWNLNPDASYVYYCA	Distúrbios da Biossíntese de Serina	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				NETVHGVEFDFIPDVKGAVLVCDMSSNFLSKPVDVSKFGVIFAGAQKNVG SAGVTVVIVRDDLLGFALRECPSVLEYKVQAGNSSLYNTPPCFSIYVMGLV LEWIKNNGGAAAMEKLSSIKSQTIYEIIDNSQGFYVCPVEPQNRSKMNIPFR IGNAKGDDALEKRFLDKALELNMLSLKGHRSVGGIRASLYNAVTIEDVQKL AAFMKKFLEMHQL [SEQ ID NO:425]		
PSPH	5723	0146733	A0A024RDL3, P78330	MVSHSELRKLFYSADAVCFDVDSTVIREEGIDELAKICGVEDAVSEMTRRA MGGAVPFKAALTERLALIQPSREQVQRLIAEQPPHLTPGIRELVSRLQERN VQVFLISGGFRSIVEHVASKLNIPATNVFANRLKFYFNGEYAGFDETQPTAE SGGKGKVIKLLKEKFHFKKIIMIGDGATDMEACPPADAFIGFGGNVIRQQVK DNAKWYITDFVELLGELEE [SEQ ID NO:426]	Distúrbios da Biossíntese de Serina	Aminoacidopatia
AMT	275	0145020	A0A024R2U7, P48728	MQRAVSVVARLGFRLQAFPPALCRPLSCAQEVLRRTPLYDFHLAHGGKM VAFAGWSLPVQYRDSHTDSHLHTRQHCSLFDVSHMLQTKILGSDRVKLM ESLVVGDIAELRPNQGTLSLFTNEAGGILDDLIVTNTSEGHLYVVSNAGCW EKDLALMQDKVRELQNQGRDVGLEVLDNALLALQGPTAAQVLQAGVADD LRKLPFMTSAVMEVFGVSGCRVTRCGYTGEDGVEISVPVAGAVHLATAIL KNPEVKLAGLAARDSLRLEAGLCLYGNDIDEHTTPVEGSLSWTLGKRRRA AMDFPGAKVIVPQLKGRVQRRRVGLMCEGAPMRAHSPILNMEGTKIGTVT SGCPSPSLKKNVAMGYVPCEYSRPGTMLLVEVRRKQQMAVVSKMPFVPT NYYTLK [SEQ ID NO:427]	Encefalopatia Glicina	Aminoacidopatia
GCSH	2653	0140905	P23434	MALRVVRSVRALLCTLRAVPSPAAPCPPRPWQLGVGAVRTLRTGPALLSV RKFTEKHEWVTTENGIGTVGISNFAQEALGDVVYCSLPEVGTKLNKQDEF GALESVKAASELYSPLSGEVTEINEALAENPGLVNKSCYEDGWLIKMTLSN PSELDELMSEEAYEKYIKSIEE [SEQ ID NO:428]	Encefalopatia Glicina	Aminoacidopatia
GLDC	2731	0178445	P23378	MQSCARAWGLRLGRGVGGGRRLAGGSGPCWAPRSRDSSSGGGDSAAA GASRLLERLLPRHDDFARRHIGPGDKDQREMLQTLGLASIDELIEKTVPANI RLKRPLKMEDPVCENEILATLHAISSKNQIWRSYIGMGYYNCSVPQTILRNL LENSGWITQYTPYQPEVSQGRLESLLNYQTMVCDITGLDMANASLLDEGT AAAEALQLCYRHNKRRKFLVDPRCHPQTIAVVQTRAKYTGVLTELKLPCE MDFSGKDVSGVLFQYPDTEGKVEDFTELVERAHQSGSLACCATDLLALC	Encefalopatia Glicina	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ILRPPGEFGVDIALGSSQRFGVPLGYGGPHAAFFAVRESLVRMMPGRMV GVTRDATGKEVYRLALQTREQHIRRDKATSNICTAQALLANMAAMFAIYHG SHGLEHIARRVHNATLILSEGLKRAGHQLQHDLFFDTLKIQCGCSVKEVLG RAAQRQINFRLFEDGTLGISLDETVNEKDLDDLLWIFGCESSAELVAESMG EECRGIPGSVFKRTSPFLTHQVFNSYHSETNIVRYMKKLENKDISLVHSMIP LGSCTMKLNSSSELAPITWKEFANIHPFVPLDQAQGYQQLFRELEKDLCEL TGYDQVCFQPNSGAQGEYAGLATIRAYLNQKGEGHRTVCLIPKSAHGTNP ASAHMAGMKIQPVEVDKYGNIDAVHLKAMVDKHKENLAAIMITYPSTNGVF EENISDVCDLIHQHGGQVYLDGANMNAQVGICRPGDFGSDVSHLNLHKTF CIPHGGGGPGMGPIGVKKHLAPFLPNHPVISLKRNEDACPVGTVSAAPWG SSSILPISWAYIKMMGGKGLKQATETAILNANYMAKRLETHYRILFRGARGY VGHEFILDTRPFKKSANIEAVDVAKRLQDYGFHAPTMSWPVAGTLMVEPT ESEDKAELDRFCDAMISIRQEIADIEEGRIDPRVNPLKMSPHSLTCVTSSH WDRPYSREVAAFPLPFVKPENKFWPTIARIDDIYGDQHLVCTCPPMEVYE SPFSEQKRASS [SEQ ID NO:429]		
LIAS	11019	0121897	O43766, Q6P5Q6, B4E0L7, A0A024R9W0, A0A1W2PQE9, A0A1X7SBR7	MSLRCGDAARTLGPRVFGRYFCSPVRPLSSLPDKKKELLQNGPDLQDFV SGDLADRSTWDEYKGNLKRQKGERLRLPPWLKTEIPMGKNYNKLKNTLR NLNLHTVCEEARCPNIGECWGGGEYATATATIMLMGDTCTRGCRFCSVKT ARNPPPLDASEPYNTAKAIAEWGLDYVVLTSVDRDDMPDGGAEHIAKTVS YLKERNPKILVECLTPDFRGDLKAIEKVALSGLDVYAHNVETVPELQSKVR DPRANFDQSLRVLKHAKKVQPDVISKTSIMLGLGENDEQVYATMKALREA DVDCLTLGQYMQPTRRHLKVEEYITPEKFKYWEKVGNELGFHYTASGPLV RSSYKAGEFFLKNLVAKRKTKDL [SEQ ID NO:430]	Encefalopatia Glicina	Aminoacidopatia
NFU1	27247	0169599	Q9UMS0	MAATARRGWGAAAVAAGLRRRFCHMLKNPYTIKKQPLHQFVQRPLFPLP AAFYHPVRYMFIQTQDTPNPNSLKFIPGKPVLETRTMDFPTPAAAFRSPLA RQLFRIEGVKSVFFGPDFITVTKENEELDWNLLKPDIYATIMDFFASGLPLV TEETPSGEAGSEEDDEVVAMIKELLDTRIRPTVQEDGGDVIYKGFEDGIVQ LKLQGSCTSCPSSIITLKNGIQNMLQFYIPEVEGVEQVMDDESDEKEANSP [SEQ ID NO:431]	Encefalopatia Glicina	Aminoacidopatia
SLC6A	6536	0196517	P48067, B7Z3W8,	MSGGDTRAAIARPRMAAAHGPVAPSSPEQVTLLPVQRSFFLPPFSGATPS	Encefalopatia	Aminoacidopatia

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
9			B7Z589	TSLAESVLKVWHGAYNSGLLPQLMAQHSLAMAQNGAVPSEATKRDQNLK RGNWGNQIEFVLTSVGYAVGLGNVWRFPYLCYRNGGGAFMFPYFIMLIFC GIPLFFMELSFGQFASQGCLGVWRISPMFKGVGYGMMVVSTYIGIYYNVVI CIAFYYFFSSMTHVLPWAYCNNPWNTHDCAGVLDASNLTNGSRPAALPS NLSHLLNHSLQRTSPSEEYWRLYVLKLSDDIGNFGEVRLPLLGCLGVSWL VVFLCLIRGVKSSGKVVYFTATFPYVVLTILFVRGVTLEGAFDGIMYYLTPQ WDKILEAKVWGDAASQIFYSLGCAWGGLITMASYNKFHNNCYRDSVIISIT NCATSVYAGFVIFSILGFMANHLGVDVSRVADHGPGLAFVAYPEALTLLPIS PLWSLLFFFMLILLGLGTQFCLLETLVTAIVDEVGNEWILQKKTYVTLGVAV AGFLLGIPLTSQAGIYWLLLMDNYAASFSLVVISCIMCVAIMYIYGHRNYFQ DIQMMLGFPPPLFFQICWRFVSPAIIFFILVFTVIQYQPITYNHYQYPGWAVA IGFLMALSSVLCIPLYAMFRLCRTDGDTLLQRLKNATKPSRDWGPALLEHR TGRYAPTIAPSPEDGFEVQPLHPDKAQIPIVGSNGSSRLQDSRI [SEQ ID NO:432]	Glicina	
SLC2A 1	6513	0117394	P11166, Q59GX2	MEPSSKKLTGRLMLAVGGAVLGSLQFGYNTGVINAPQKVIEEFYNQTWVH RYGESILPTTLTTLWSLSVAIFSVGGMIGSFSVGLFVNRFGRRNSMLMMNL LAFVSAVLMGFSKLGKSFEMLILGRFIIGVYCGLTTGFVPMYVGEVSPTALR GALGTLHQLGIVVGILIAQVFGLDSIMGNKDLWPLLLSIIFIPALLQCIVLPFCP ESPRFLLINRNEENRAKSVLKKLRGTADVTHDLQEMKEESRQMMREKKVT ILELFRSPAYRQPILIAVVLQLSQQLSGINAVFYYSTSIFEKAGVQQPVYATI GSGIVNTAFTVVSLFVVERAGRRTLHLIGLAGMAGCAILMTIALALLEQLPW MSYLSIVAIFGFVAFFEVGPGPIPWFIVAELFSQGPRPAAIAVAGFSNWTSN FIVGMCFQYVEQLCGPYVFIIFTVLLVLFFIFTYFKVPETKGRTFDEIASGFR QGGASQSDKTPEELFHPLGADSQV [SEQ ID NO:433]	Deficiência do Transportador de Glicose Tipo 1	Distúrbio de carbo-hidratos
ATP7A	538	0165240	B4DRW0, Q04656, Q762B6	MDPSMGVNSVTISVEGMTCNSCVWTIEQQIGKVNGVHHIKVSLEEKNATII YDPKLQTPKTLQEAIDDMGFDAVIHNPDPLPVLTDTLFLTVTASLTLPWDHI QSTLLKTKGVTDIKIYPQKRTVAVTIIPSIVNANQIKELVPELSLDTGTLEKKS GACEDHSMAQAGEVVLKMKVEGMTCHSCTSTIEGKIGKLQGVQRIKVSLD NQEATIVYQPHLISVEEMKKQIEAMGFPAFVKKQPKYLKLGAIDVERLKNTP VKSSEGSQQRSPSYTNDSTATFIIDGMHCKSCVSNIESTLSALQYVSSIVVS LENRSAIVKYNASSVTPESLRKAIEAVSPGLYRVSITSEVESTSNSPSSSSL	Doenças relacionadas a ATP7A Distúrbio do Metabolismo do Cobre	Distúrbio de Transporte de Metal

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QKIPLNVVSQPLTQETVINIDGMTCNSCVQSIEGVISKKPGVKSIRVSLANS NGTVEYDPLLTSPETLRGAIEDMGFDATLSDTNEPLVVIAQPSSEMPLLTS TNEFYTKGMTPVQDKEEGKNSSKCYIQVTGMTCASCVANIERNLRREEGI YSILVALMAGKAEVRYNPAVIQPPMIAEFIRELGFGATVIENADEGDGVLEL VVRGMTCASCVHKIESSLTKHRGILYCSVALATNKAHIKYDPEIIGPRDIIHTI ESLGFEASLVKKDRSASHLDHKREIRQWRRSFLVSLFFCIPVMGLMIYMMV MDHHFATLHHNQNMSKEEMINLHSSMFLERQILPGLSVMNLLSFLLCVPV QFFGGWYFYIQAYKALKHKTANMDVLIVLATTIAFAYSLIILLVAMYERAKVN PITFFDTPPMLFVFIALGRWLEHIAKGKTSEALAKLISLQATEATIVTLDSDNI LLSEEQVDVELVQRGDIIKVVPGGKFPVDGRVIEGHSMVDESLITGEAMPV AKKPGSTVIAGSINQNGSLLICATHVGADTTLSQIVKLVEEAQTSKAPIQQF ADKLSGYFVPFIVFVSIATLLVWIVIGFLNFEIVETYFPGYNRSISRTETIIRFA FQASITVLCIACPCSLGLATPTAVMVGTGVGAQNGILIKGGEPLEMAHKVK VVVFDKTGTITHGTPVVNQVKVLTESNRISHHKILAIVGTAESNSEHPLGTAI TKYCKQELDTETLGTCIDFQVVPGCGISCKVTNIEGLLHKNNWNIEDNNIKN ASLVQIDASNEQSSTSSSMIIDAQISNALNAQQYKVLIGNREWMIRNGLVIN NDVNDFMTEHERKGRTAVLVAVDDELCGLIAIADTVKPEAELAIHILKSMGL EVVLMTGDNSKTARSIASQVGITKVFAEVLPSHKVAKVKQLQEEGKRVAM VGDGINDSPALAMANVGIAIGTGTDVAIEAADVVLIRNDLLDVVASIDLSRKT VKRIRINFVFALIYNLVGIPIAAGVFMPIGLVLQPWMGSAAMAASSVSVVLS SLFLKLYRKPTYESYELPARSQIGQKSPSEISVHVGIDDTSRNSPKLGLLDR IVNYSRASINSLLSDKRSLNSVVTSEPDKHSLLVGDFREDDDTAL [SEQ ID NO:434]		
AP1S1	1174	0106367	A0A024QYT6, P61966	MMRFMLLFSRQGKLRLQKWYLATSDKERKKMVRELMQVVLARKPKMCS FLEWRDLKVVYKRYASLYFCCAIEGQDNELITLELIHRYVELLDKYFGSVCE LDIIFNFEKAYFILDEFLMGGDVQDTSKKSVLKAIEQADLLQEEDESPRSVL EEMGLA [SEQ ID NO:435]	Distúrbio do Metabolismo do Cobre	Distúrbio de Transporte de Metal
CP	1356	0047457	A5PL27, P00450	MKILILGIFLFLCSTPAWAKEKHYYIGIIETTWDYASDHGEKKLISVDTEHSNI YLQNGPDRIGRLYKKALYLQYTDETFRTTIEKPVWLGFLGPIIKAETGDKVY VHLKNLASRPYTFHSHGITYYKEHEGAIYPDNTTDFQRADDKVYPGEQYT YMLLATEEQSPGEGDGNCVTRIYHSHIDAPKDIASGLIGPLIICKKDSLDKEK	Distúrbio do Metabolismo do Cobre	Distúrbio de Transporte de Metal

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				EKHIDREFVVMFSVVDENFSWYLEDNIKTYCSEPEKVDKDNEDFQESNRM YSVNGYTFGSLPGLSMCAEDRVKWYLFGMGNEVDVHAAFFHGQALTNK NYRIDTINLFPATLFDAYMVAQNPGEWMLSCQNLNHLKAGLQAFFQVQEC NKSSSKDNIRGKHVRHYYIAAEEIIWNYAPSGIDIFTKENLTAPGSDSAVFF EQGTTRIGGSYKKLVYREYTDASFTNRKERGPEEEHLGILGPVIWAEVGDT IRVTFHNKGAYPLSIEPIGVRFNKNNEGTYYSPNYNPQSRSVPPSASHVAP TETFTYEWTVPKEVGPTNADPVCLAKMYYSAVDPTKDIFTGLIGPMKICKK GSLHANGRQKDVDKEFYLFPTVFDENESLLLEDNIRMFTTAPDQVDKEDE DFQESNKMHSMNGFMYGNQPGLTMCKGDSVVWYLFSAGNEADVHGIYF SGNTYLWRGERRDTANLFPQTSLTLHMWPDTEGTFNVECLTTDHYTGGM KQKYTVNQCRRQSEDSTFYLGERTYYIAAVEVEWDYSPQREWEKELHHL QEQNVSNAFLDKGEFYIGSKYKKVVYRQYTDSTFRVPVERKAEEEHLGIL GPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVTPTLPGETLTYVW KIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRPYLKVFNP RRKLEFALLFLVFDENESWYLDDNIKTYSDHPEKVNKDDEEFIESNKMHAI NGRMFGNLQGLTMHVGDEVNWYLMGMGNEIDLHTVHFHGHSFQYKHRG VYSSDVFDIFPGTYQTLEMFPRTPGIWLLHCHVTDHIHAGMETTYTVLQNE DTKSG [SEQ ID NO:436]		
SLC33A1	9197	0169359	O00400	MSPTISHKDSSRQRRPGNFSHSLDMKSGPLPPGGWDDSHLDSAGREGD REALLGDTGTGDFLKAPQSFRAELSSILLLLFLYVLQGIPLGLAGSIPLILQS KNVSYTDQAFFSFVFWPFSLKLLWAPLVDAVYVKNFGRRKSWLVPTQYIL GLFMIYLSTQVDRLLGNTDDRTPDVIALTVAFFLFEFLAATQDIAVDGWALT MLSRENVGYASTCNSVGQTAGYFLGNVLFLALESADFCNKYLRFQPQPR GIVTLSDFLFFWGTVFLITTTLVALLKKENEVSVVKEETQGITDTYKLLFAIIK MPAVLTFCLLILTAKIGFSAADAVTGLKLVEEGVPKEHLALLAVPMVPLQIIL PLIISKYTAGPQPLNTFYKAMPYRLLLGLEYALLVWWTPKVEHQGGFPIYY YIVVLLSYALHQVTVYSMYVSIMAFNAKVSDPLIGGTYMTLLNTVSNLGGN WPSTVALWLVDPLTVKECVGASNQNCRTPDAVELCKKLGGSCVTALDGY YVESIICVFIGFGWWFFLGPKFKKLQDEGSSSWKCKRNN[SEQ ID NO:437]	Distúrbio do Metabolismo do Cobre	Distúrbio de Transporte de Metal
PEX7	5191	01112357	O00628, Q6FGN1	MSAVCGGAARMLRTPGRHGYAAEFSPYLPGRLACATAQHYGIAGCGTLLI LDPDEAGLRLFRSFDWNDGLFDVTWSENNEHVLITCSGDGSLQLWDTAK	Doença de Refsum em Adultos	Distúrbios Peroxissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				AAGPLQVYKEHAQEVYSVDWSQTRGEQLVVSGSWDQTVKLWDPTVGKS LCTFRGHESIIYSTIWSPHIPGCFASASGDQTLRIWDVKAAGVRIVIPAHQA EILSCDWCKYNENLLVTGAVDCSLRGWDLRNVRQPVFELLGHTYAIRRVK FSPFHASVLASCSYDFTVRFWNFSKPDSLLETVEHHTEFTCGLDFSLQSP TQVADCSWDETIKIYDPACLTIPA [SEQ ID NO:438]	Espectro de Condrodisplasia punctata rizomélica	
PHYH	5264	0107537	O14832	MEQLRAAARLQIVLGHLGRPSAGAVVAHPTSGTISSASFHPQQFQYTLDN NVLTLEQRKFYEENGFLVIKNLVPDADIQRFRNEFEKICRKEVKPLGLTVMR DVTISKSEYAPSEKMITKVQDFQEDKELFRYCTLPEILKYVECFTGPNIMAM HTMLINKPPDSGKKTSRHPLHQDLHYFPFRPSDLIVCAWTAMEHISRNNG CLVVLPGTHKGSLKPHDYPKWEGGVNKMFHGIQDYEENKARVHLVMEKG DTVFFHPLLIHGSGQNKTQGFRKAISCHFASADCHYIDVKGTSQENIEKEV VGIAHKFFGAENSVNLKDIWMFRARLVKGERTNL [SEQ ID NO:439]	Doença de Refsum em Adultos	Distúrbios Peroxissomais
AGPS	8540	0018510	O00116, B7Z3Q4	MAEAAAAAGGTGLGAGASYGSAADRDRDPDPDRAGRRLRVLSGHLLGR PREALSTNECKARRAASAATAAPTATPAAQESGTIPKKRQEVMKWNGWG YNDSKFIFNKKGQIELTGKRYPLSGMGLPTFKEWIQNTLGVNVEHKTTSKA SLNPSDTPPSVVNEDFLHDLKETNISYSQEADDRVFRAHGHCLHEIFLLRE GMFERIPDIVLWPTCHDDVVKIVNLACKYNLCIIPIGGGTSVSYGLMCPADE TRTIISLDTSQMNRILWVDENNLTAHVEAGITGQELERQLKESGYCTGHEP DSLEFSTVGGWVSTRASGMKKNIYGNIEDLVVHIKMVTPRGIIEKSCQGPR MSTGPDIHHFIMGSEGTLGVITEATIKIRPVPEYQKYGSVAFPNFEQGVACL REIAKQRCAPASIRLMDNKQFQFGHALKPQVSSIFTSFLDGLKKFYITKFKG FDPNQLSVATLLFEGDREKVLQHEKQVYDIAAKFGGLAAGEDNGQRGYLL TYVIAYIRDLALEYYVLGESFETSAPWDRVVDLCRNVKERITRECKEKGVQ FAPFSTCRVTQTYDAGACIYFYFAFNYRGISDPLTVFEQTEAAAREEILANG GSLSHHHGVGKLRKQWLKESISDVGFGMLKSVKEYVDPNNIFGNRNLL [SEQ ID NO:440]	Espectro de Condrodisplasia punctata rizomélica	Distúrbios Peroxissomais
GNPAT	8443	0116906	O15228	MESSSSSNSYFSVGPTSPSAVVLLYSKELKKWDEFEDILEERRHVSDLKFA MKCYTPLVYKGITPCKPIDIKCSVLNSEEIHYVIKQLSKESLQSVDVLREEVS EILDEMSHKLRLGAIRFCAFTLSKVFKQIFSKVCVNEEGIQKLQRAIQEHPV VLLPSHRSYIDFLMLSFLLYNYDLPVPVIAAGMDFLGMKMVGELLRMSGAF	Espectro de Condrodisplasia punctata rizomélica	Distúrbios Peroxissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FMRRTFGGNKLYWAVFSEYVKTMLRNGYAPVEFFLEGTRSRSAKTLTPKF GLLNIVMEPFFKREVFDTYLVPISISYDKILEETLYVYELLGVPKPKESTTGLL KARKILSENFGSIHVYFGDPVSLRSLAAGRMSRSSYNLVPRYIPQKQSEDM HAFVTEVAYKMELLQIENMVLSPWTLIVAVLLQNRPSMDFDALVEKTLWLK GLTQAFGGFLIWPDNKPAEEVVPASILLHSNIASLVKDQVILKVDSGDSEVV DGLMLQHITLLMCSAYRNQLLNIFVRPSLVAVALQMTPGFRKEDVYSCFRF LRDVFADEFIFLPGNTLKDFEEGCYLLCKSEAIQVTTKDILVTEKGNTVLEFL VGLFKPFVESYQIICKYLLSEEEDHFSEEQYLAAVRKFTSQLLDQGTSQCY DVLSSDVQKNALAACVRLGVVEKKKINNNCIFNVNEPATTKLEEMLGCKTP IGKPATAKL [SEQ ID NO:441]		
ABCD 1	215	0101986	P33897	MPVLSRPRPWRGNTLKRTAVLLALAAYGAHKVYPLVRQCLAPARGLQAP AGEPTQEASGVAAAKAGMNRVFLQRLLWLLRLLFPRVLCRETGLLALHSA ALVSRTFLSVYVARLDGRLARCIVRKDPRAFGWQLLQWLLIALPATFVNSAI RYLEGQLALSFRSRLVAHAYRLYFSQQTYYRVSNMDGRLRNPDQSLTED VVAFAASVAHLYSNLTKPLLDVAVTSYTLLRAARSRGAGTAWPSAIAGLVV FLTANVLRAFSPKFGELVAEEARRKGELRYMHSRVVANSEEIAFYGGHEV ELALLQRSYQDLASQINLILLERLWYVMLEQFLMKYVWSASGLLMVAVPIIT ATGYSESDAEAVKKAALEKKEEELVSERTEAFTIARNLLTAAADAIERIMSS YKEVTELAGYTARVHEMFQVFEDVQRCHFKRPRELEDAQAGSGTIGRSG VRVEGPLKIRGQVVDVEQGIICENIPIVTPSGEVVVASLNIRVEEGMHLLITG PNGCGKSSLFRILGGLWPTYGGVLYKPPPQRMFYIPQRPYMSVGSLRDQ VIYPDSVEDMQRKGYSEQDLEAILDVVHLHHILQREGGWEAMCDWKDVL SGGEKQRIGMARMFYHRPKYALLDECTSAVSIDVEGKIFQAAKDAGIALLSI THRPSLWKYHTHLLQFDGEGGWKFEKLDSAARLSLTEEKQRLEQQLAGIP KMQRRLQELCQILGEAVAPAHVPAPSPQGPGGLQGAST [SEQ ID NO:442]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
ACOX 1	51	0161533	Q15067	MNPDLRRERDSASFNPELLTHILDGSPEKTRRRREIENMILNDPDFQHEDL NFLTRSQRYEVAVRKSAIMVKKMREFGIADPDEIMWFKKLHLVNFVEPVGL NYSMFIPTLLNQGTTAQKEKWLLSSKGLQIIGTYAQTEMGHGTHLRGLETT ATYDPETQEFILNSPTVTSIKWWPGGLGKTSNHAIVLAQLITKGKCYGLHAF IVPIREIGTHKPLPGITVGDIGPKFGYDEIDNGYLKMDNHRIPRENMLMKYA	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QVKPDGTYVKPLSNKLTYGTMVFVRSFLVGEAARALSKACTIAIRYSAVRH QSEIKPGEPEPQILDFQTQQYKLFPLLATAYAFQFVGAYMKETYHRINEGIG QGDLSELPELHALTAGLKAFTSWTANTGIEACRMACGGHGYSHCSGLPNI YVNFTPSCTFEGENTVMMLQTARFLMKSYDQVHSGKLVCGMVSYLNDLP SQRIQPQQVAVWPTMVDINSPESLTEAYKLRAARLVEIAAKNLQKEVIHRK SKEVAWNLTSVDLVRASEAHCHYVVVKLFSEKLLKIQDKAIQAVLRSLCLLY SLYGISQNAGDFLQGSIMTEPQITQVNQRVKELLTLIRSDAVALVDAFDFQD VTLGSVLGRYDGNVYENLFEWAKNSPLNKAEVHESYKHLKSLQSKL [SEQ ID NO:443]		
PEX1	5189	0127980	O43933, A0A0C4DG33, B4DER6	MWGSDRLAGAGGGGAAVTVAFTNARDCFLHLPRRLVAQLHLLQNQAIEV VWSHQPAFLSWVEGRHFSDQGENVAEINRQVGQKLGLSNGGQVFLKPC SHVVSCQQVEVEPLSADDWEILELHAVSLEQHLLDQIRIVFPKAIFPVWVD QQTYIFIQIVALIPAASYGRLETDTKLLIQPKTRRAKENTFSKADAEYKKLHS YGRDQKGMMKELQTKQLQSNTVGITESNENESEIPVDSSSVASLWTMIGS IFSFQSEKKQETSWGLTEINAFKNMQSKVVPLDNIFRVCKSQPPSIYNASA TSVFHKHCAIHVFPWDQEYFDVEPSFTVTYGKLVKLLSPKQQQSKTKQNV LSPEKEKQMSEPLDQKKIRSDHNEEDEKACVLQVVWNGLEELNNAIKYTK NVEVLHLGKVWIPDDLRKRLNIEMHAVVRITPVEVTPKIPRSLKLQPRENLP KDISEEDIKTVFYSWLQQSTTTMLPLVISEEEFIKLETKDGLKEFSLSIVHSW EKEKDKNIFLLSPNLLQKTTIQVLLDPMVKEENSEEIDFILPFLKLSSLGGVN SLGVSSLEHITHSLLGRPLSRQLMSLVAGLRNGALLLTGGKGSGKSTLAKA ICKEAFDKLDAHVERVDCKALRGKRLENIQKTLEVAFSEAVWMQPSVVLLD DLDLIAGLPAVPEHEHSPDAVQSQRLAHALNDMIKEFISMGSLVALIATSQS QQSLHPLLVSAQGVHIFQCVQHIQPPNQEQRCEILCNVIKNKLDCDINKFT DLDLQHVAKETGGFVARDFTVLVDRAIHSRLSRQSISTREKLVLTTLDFQK ALRGFLPASLRSVNLHKPRDLGWDKIGGLHEVRQILMDTIQLPAKYPELFA NLPIRQRTGILLYGPPGTGKTLLAGVIARESRMNFISVKGPELLSKYIGASE QAVRDIFIRAQAAKPCILFFDEFESIAPRRGHDNTGVTDRVVNQLLTQLDG VEGLQGVYVLAATSRPDLIDPALLRPGRLDKCVYCPPPDQVSRLEILNVLS DSLPLADDVDLQHVASVTDSFTGADLKALLYNAQLEALHGMLLSSGLQDG SSSSDSDLSLSSMVFLNHSSGSDDSAGDGECGLDQSLVSLEMSEILPDES KFNMYRLYFGSSYESELGNGTSSDLSSQCLSAPSSMTQDLPGVPGKDQL	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FSQPPVLRTASQEGCQELTQEQRDQLRADISIIKGRYRSQSGEDESMNQP GPIKTRLAISQSHLMTALGHTRPSISEDDWKNFAELYESFQNPKRRKNQS GTMFRPGQKVTLA [SEQ ID NO:444]		
PEX2	5828	0164751	P28328	MASRKENAKSANRVLRISQLDALELNKALEQLVWSQFTQCFHGFKPGLLA RFEPEVKACLWVFLWRFTIYSKNATVGQSVLNIKYKNDFSPNLRYQPPSK NQKIWYAVCTIGGRWLEERCYDLFRNHHLASFGKVKQCVNFVIGLLKLGG LINFLIFLQRGKFATLTERLLGIHSVFCKPQNICEVGFEYMNRELLWHGFAE FLIFLLPLINVQKLKAKLSSWCIPLTGAPNSDNTLATSGKECALCGEWPTMP HTIGCEHIFCYFCAKSSFLFDVYFTCPKCGTEVHSLQPLKSGIEMSEVNAL [SEQ ID NO:445]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
PEX3	8504	0034693	P56589	MLRSVWNFLKRHKKKCIFLGTVLGGVYILGKYGQKKIREIQEREAAEYIAQA RRQYHFESNQRTCNMTVLSMLPTLREALMQQLNSESLTALLKNRPSNKLE IWEDLKIISFTRSTVAVYSTCMLVVLLRVQLNIIGGYIYLDNAAVGKNGTTILA PPDVQQQYLSSIQHLLGDGLTELITVIKQAVQKVLGSVSLKHSLSLLDLEQK LKEIRNLVEQHKSSSWINKDGSKPLLCHYMMPDEETPLAVQACGLSPRDIT TIKLLNETRDMLESPDFSTVLNTCLNRGFSRLLDNMAEFFRPTEQDLQHG NSMNSLSSVSLPLAKIIPIVNGQIHSVCSETPSHFVQDLLTMEQVKDFAANV YEAFSTPQQLEK [SEQ ID NO:446]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
PEX5	5830	0139197	A0A0S2Z480, P50542, B4DR50, A0A0S2Z4F3, A0A0S2Z4H1, B4E0T2	MAMRELVEAECGGANPLMKLAGHFTQDKALRQEGLRPGPWPPGAPASE AASKPLGVASEDELVAEFLQDQNAPLVSRAPQTFKMDDLLAEMQQIEQSN FRQAPQRAPGVADLALSENWAQEFLAAGDAVDVTQDYNETDWSQEFISE VTDPLSVSPARWAEEYLEQSEEKLWLGEPEGTATDRWYDEYHPEEDLQH TASDFVAKVDDPKLANSEFLKFVRQIGEGQVSLESGAGSGRAQAEQWAA EFIQQQGTSDAWVDQFTRPVNTSALDMEFERAKSAIESDVDFWDKLQAEL EEMAKRDAEAHPWLSDYDDLTSATYDKGYQFEEENPLRDHPQPFEEGLR RLQEGDLPNAVLLFEAAVQQDPKHMEAWQYLGTTQAENEQELLAISALRR CLELKPDNQTALMALAVSFTNESLQRQACETLRDWLRYTPAYAHLVTPAE EGAGGAGLGPSKRILGSLLSDSLFLEVKELFLAAVRLDPTSIDPDVQCGLG VLFNLSGEYDKAVDCFTAALSVRPNDYLLWNKLGATLANGNQSEEAVAAY RRALELQPGYIRSRYNLGISCINLGAHREAVEHFLEALNMQRKSRGPRGE	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GGAMSENIWSTLRLALSMLGQSDAYGAADARDLSTLLTMFGLPQ [SEQ ID NO:447]		
PEX6	5190	0124587	A0A024RD09, Q13608	MALAVLRVLEPFPTETPPLAVLLPPGGPWPAAELGLVLALRPAGESPAGP ALLVAALEGPDAGTEEQGPGPPQLLVSRALLRLLALGSGAWVRARAVRRP PALGWALLGTSLGPGLGPRVGPLLVRRGETLPVPGPRVLETRPALQGLLG PGTRLAVTELRGRARLCPESGDSSRPPPPPVVSSFAVSGTVRRLQGVLG GTGDSLGVSRSCLRGLGLFQGEWVWVAQARESSNTSQPHLARVQVLEP RWDLSDRLGPGSGPLGEPLADGLALVPATLAFNLGCDPLEMGELRIQRYL EGSIAPEDKGSCSLLPGPPFARELHIEIVSSPHYSTNGNYDGVLYRHFQIPR VVQEGDVLCVPTIGQVEILEGSPEKLPRWREMFFKVKKTVGEAPDGPASA YLADTTHTSLYMVGSTLSPVPWLPSEESTLWSSLSPPGLEALVSELCAVLK PRLQPGGALLTGTSSVLLRGPPGCGKTTVVAAACSHLGLHLLKVPCSSLC AESSGAVETKLQAIFSRARRCRPAVLLLTAVDLLGRDRDGLGEDARVMAV LRHLLLNEDPLNSCPPLMVVATTSRAQDLPADVQTAFPHELEVPALSEGQ RLSILRALTAHLPLGQEVNLAQLARRCAGFVVGDLYALLTHSSRAACTRIK NSGLAGGLTEEDEGELCAAGFPLLAEDFGQALEQLQTAHSQAVGAPKIPS VSWHDVGGLQEVKKEILETIQLPLEHPELLSLGLRRSGLLLHGPPGTGKTL LAKAVATECSLTFLSVKGPELINMYVGQSEENVREVFARARAAAPCIIFFDE LDSLAPSRGRSGDSGGVMDRVVSQLLAELDGLHSTQDVFVIGATNRPDLL DPALLRPGRFDKLVFVGANEDRASQLRVLSAITRKFKLEPSVSLVNVLDCC PPQLTGADLYSLCSDAMTAALKRRVHDLEEGLEPGSSALMLTMEDLLQAA ARLQPSVSEQELLRYKRIQRKFAAC [SEQ ID NO:448]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
PEX10	5192	0157911	A0A024R068, O60683, A0A024R0A4	MAPAAASPPEVIRAAQKDEYYRGGLRSAAGGALHSLAGARKWLEWRKEV ELLSDVAYFGLTTLAGYQTLGEEYVSIIQVDPSRIHVPSSLRRGVLVTLHAV LPYLLDKALLPLEQELQADPDSGRPLQGSLGPGGRGCSGARRWMRHHTA TLTEQQRRALLRAVFVLRQGLACLQRLHVAWFYIHGVFYHLAKRLTGITYL RVRSLPGEDLRARVSYRLLGVISLLHLVLSMGLQLYGFRQRQRARKEWRL HRGLSHRRASLEERAVSRNPLCTLCLEERRHPTATPCGHLFCWECITAWC SSKAECPLCREKFPPQKLIYLRHYR[SEQ ID NO:449]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
PEX12	5193	0108733	O00623	MAEHGAHFTAASVADDQPSIFEVVAQDSLMTAVRPALQHVVKVLAESNPT	Adrenoleucodistro-	Distúrbios Pero-

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				HYGFLWRWFDEIFTLLDLLLQQHYLSRTSASFSENFYGLKRIVMGDTHKS QRLASAGLPKQQLWKSIMFLVLLPYLKVKLEKLVSSLREEDEYSIHPPSSR WKRFYRAFLAAYPFVNMAWEGWFLVQQLRYILGKAQHHSPLLRLAGVQL GRLTVQDIQALEHKPAKASMMQQPARSVSEKINSALKKAVGGVALSLSTG LSVGVFFLQFLDWWYSSENQETIKSLTALPTPPPPVHLDYNSDSPLLPKMK TVCPLCRKTRVNDTVLATSGYVFCYRCVFHYVRSHQACPITGYPTEVQHLI KLYSPEN [SEQ ID NO:450]	fia Ligada ao X	xissomais
PEX13	5194	0162928	Q92968	MASQPPPPPKPWETRRIPGAGPGPGPGPTFQSADLGPTLMTRPGQPALT RVPPPILPRPSQQTGSSSVNTFRPAYSSFSSGYGAYGNSFYGGYSPYSY GYNGLGYNRLRVDDLPPSRFVQQAEESSRGAFQSIESIVHAFASVSMMM DATFSAVYNSFRAVLDVANHFSRLKIHFTKVFSAFALVRTIRYLYRRLQRML GLRRGSENEDLWAESEGTVACLGAEDRAATSAKSWPIFLFFAVILGGPYLI WKLLSTHSDEVTDSINWASGEDDHVVARAEYDFAAVSEEEISFRAGDMLN LALKEQQPKVRGWLLASLDGQTTGLIPANYVKILGKRKGRKTVESSKVSK QQQSFTNPTLTKGATVADSLDEQEAAFESVFVETNKVPVAPDSIGKDGEK QDL[SEQ ID NO:451]	Adrenoleucodistro-fia Ligada ao X	Distúrbios Pero-xissomais
PEX14	5195	0142655	O75381	MASSEQAEQPSQPSSTPGSENVLPREPLIATAVKFLQNSRVRQSPLATRR AFLKKKGLTDEEIDMAFQQSGTAADEPSSLGPATQVVPVQPPHLISQPYSP AGSRWRDYGALAIIMAGIAFGFHQLYKKYLLPLILGGREDRKQLERMEAGL SELSGSVAQTVTQLQTTLASVQELLIQQQQKIQELAHELAAAKATTSTNWIL ESQNINELKSEINSLKGLLLNRRQFPPSPSAPKIPSWQIPVKSPSPSSPAAV NHHSSSDISPVSNESTSSSPGKEGHSPEGSTVTYHLLGPQEEGEGVVDVK GQVRMEVQGEEEKREDKEDEEDEEDDDVSHVDEEDCLGVQREDRRGG DGQINEQVEKLRRPEGASNESERD[SEQ ID NO:452]	Adrenoleucodistro-fia Ligada ao X	Distúrbios Pero-xissomais
PEX16	9409	0121680	Q9Y5Y5	MEKLRLLGLRYQEYVTRHPAATAQLETAVRGFSYLLAGRFADSHELSELV YSASNLLVLLNDGILRKELRKKLPVSLSQQKLLTWLSVLECVEVFMEMGAA KVWGEVGRWLVIALVQLAKAVLRMLLLLWFKAGLQTSPPIVPLDRETQAQ PPDGDHSPGNHEQSYVGKRSNRVVRTLQNTPSLHSRHWGAPQQREGR QQQHHEELSATPTPLGLQETIAEFLYIARPLLHLLSLGLWGQRSWKPWLLA GVVDVTSLSLLSDRKGLTRRERRELRRRTILLLYYLLRSPFYDRFSEARIL	Adrenoleucodistro-fia Ligada ao X	Distúrbios Pero-xissomais

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FLLQLLADHVPGVGLVTRPLMDYLPTWQKIYFYSWG [SEQ ID NO:453]		
PEX19	5824	0162735	P40855, A0A0S2Z497	MAAAEEGCSVGAEADRELEELLESALDDFDKAKPSPAPPSTTTAPDASGPQKRSPGDTAKDALFASQEKFFQELFDSELASQATAEFEKAMKELAEEEPHLVEQFQKLSEAAGRVGSDMTSQQEFTSCLKETLSGLAKNATDLQNSSMSEEELTKAMEGLGMDEGDGEGNILPIMQSIMQNLLSKDVLYPSLKEITEKYPEWLQSHRESLPPEQFEKYQEQHSVMCKICEQFEAETPTDSETTQKARFEMVLDLMQQLQDLGHPPKELAGEMPPGLNFDLDALNLSGPPGASGEQCLIM [SEQ ID NO:454]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
PEX26	55670	0215193	A0A024R100, Q7Z412, A0A0S2Z5M7, Q7Z2D7	MKSDSSTSAAPLRGLGGPLRSSEPVRAVPARAPAVDLLEEAADLLVVHLDFRAALETCERAWQSLANHAVAEEPAGTSLEVKCSLCVVGIQALAEMDRWQEVLSWVLQYYQVPEKLPPKVLELCILLYSKMQEPGAVLDVVGAWLQDPANQNLPEYGALAEFHVQRVLLPLGCLSEAEELVVGSAAFGEERRLDVLQAIHTARQQQKQEHSGSEEAQKPNLEGSVSHKFLSLPMLVRQLWDSAVSHFFSLPFKKSLLAALILCLLVVRFDPASPSSLHFLYKLAQLFRWIRKAAFSRLYQLRIRD [SEQ ID NO:455]	Adrenoleucodistrofia Ligada ao X	Distúrbios Peroxissomais
AMACR	23600	0242110	Q9UHK6	MALQGISVVELSGLAPGPFCAMVLADFGARVVRVDRPGSRYDVSRLGRGKRSLVLDLKQPRGAAVLRRLCKRSDVLLEPFRRGVMEKLQLGPEILQRENPRLIYARLSGFGQSGSFCRLAGHDINYLALSGVLSKIGRSGENPYAPLNLLADFAGGGLMCALGIIMALFDRTRTGKGQVIDANMVEGTAYLSSFLWKTQKLSLWEAPRGQNMLDGGAPFYTTYRTADGEFMAVGAIEPQFYELLIKGLGLKSDELPNQMSMDDWPEMKKKFADVFAEKTKAEWCQIFDGTDACVTPVLTFEEVVHHDHNKERGSFITSEEQDVSPRPAPLLLNTPAIPSFKRDPFIGEHTEEILEEFGFSREEIYQLNSDKIIESNKVVKASL [SEQ ID NO:456]	Transtorno do Espectro de Zellweger	Distúrbios Peroxissomais
ADA	100	0196839	A0A0S2Z381, P00813, F5GWI4	MAQTPAFDKPKVELHVHLDGSIKPETILYYGRRRGIALPANTAEGLLNVIGMDKPLTLPDFLAKFDYYMPAIAGCREAIKRIAYEFVEMKAKEGVVYVEVRYSPHLLANSKVEPIPWNQAEGDLTPDEVVALVGQGLQEGERDFGVKARSILCCMRHQPNWSPKVVELCKKYQQQTVVAIDLAGDETIPGSSLLPGHVQAYQEAVKSGIHRTVHAGEVGSAEVVKEAVDILKTERLGHGYHTLEDQALYNRLRQENMHFEICPWSSYLTGAWKPDTEHAVIRLKNDQANYSLNTDDPLIF	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				KSTLDTDYQMTKRDMGFTEEEFKRLNINAAKSSFLPEDEKRELLDLLYKAY GMPPSASAGQNL [SEQ ID NO:457]		
ADSL	158	0239900	P30566, X5D8S6, X5D7W4, A0A1B0GWJ0	MAAGGDHGSPDSYRSPLASRYASPEMCFVFSDRYKFRTWRQLWLWLAE AEQTLGLPITDEQIQEMKSNLENIDFKMAAEEEKRLRHDVMAHVHTFGHC CPKAAGIIHLGATSCYVGDNTDLIILRNALDLLLPKLARVISRLADFAKERAS LPTLGFTHFQPAQLTTVGKRCCLWIQDLCMDLQNLKRVRDDLRFRGVKGT TGTQASFLQLFEGDDHKVEQLDKMVTEKAGFKRAFIITGQTYTRKVDIEVL SVLASLGASVHKICTDIRLLANLKEMEEPFEKQQIGSSAMPYKRNPMRSER CCSLARHLMTLVMDPLQTASVQWFERTLDDSANRRICLAEAFLTADTILNT LQNISEGLVVYPKVIERRIRQELPFMATENIIMAMVKAGGSRQDCHEKIRVL SQQAASVVKQEGGDNDLIERIQVDAYFSPIHSQLDHLLDPSSFTGRASQQ VQRFLEEEVYPLLKPYESVMKVKAELCL [SEQ ID NO:458]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
AMPD 1	270	0116748	P23109	MNVRIFYSVSQSPHSLLSLLFYCAILESRISATMPLFKLPAEEKQIDDAMRN FAEKVFASEVKDEGGRQEISPFDVDEICPISHHEMQAHIFHLETLSTSTEAR RKKRFQGRKTVNLSIPLSETSSTKLSHIDEYISSSPTYQTVPDFQRVQITGD YASGVTVEDFEIVCKGLYRALCIREKYMQKSFQRFPKTPSKYLRNIDGEAW VANESFYPVFTPPVKKGEDPFRTDNLPENLGYHLKMKDGVVYVYPNEAAV SKDEPKPLPYPNLDTFLDDMNFLLALIAQGPVKTYTHRRLKFLSSKFQVHQ MLNEMDELKELKNNPHRDFYNCRKVDTHIHAAACMNQKHLLRFIKKSYQI DADRVVYSTKEKNLTLKELFAKLKMHPYDLTVDSLDVHAGRQTFQRFDKF NDKYNPVGASELRDLYLKTDNYINGEYFATIIKEVGADLVEAKYQHAEPRL SIYGRSPDEWSKLSSWFVCNRIHCPNMTWMIQVPRIYDVFRSKNFLPHFG KMLENIFMPVFEATINPQADPELSVFLKHITGFDSVDDESKHSGHMFSSKS PKPQEWTLEKNPSYTYYAYYMYANIMVLNSLRKERGMNTFLFRPHCGEA GALTHLMTAFMIADDISHGLNLKKSPVLQYLFFLAQIPIAMSPLSNNSLFLEY AKNPFLDFLQKGLMISLSTDDPMQFHFTKEPLMEEYAIAAQVFKLSTCDMC EVARNSVLQCGISHEEKVKFLGDNYLEEGPAGNDIRRTNVAQIRMAYRYE TWCYELNLIAEGLKSTE [SEQ ID NO:459]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
GPHN	10243	0171723	Q9NQX3	MATEGMILTNHDHQIRVGVLTVSDSCFRNLAEDRSGINLKDLVQDPSLLGG TISAYKIVPDEIEEIKETLIDWCDEKELNLILTTGGTGFAPRDVTPEATKEVIE	Transtorno do Metabolismo da	Transtorno do Metabolismo da

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				REAPGMALAMLMGSLNVTPLGMLSRPVCGIRGKTLIINLPGSKKGSQECF QFILPALPHAIDLLRDAIVKVKEVHDELEDLPSPPPPLSPPPTTSPHKQTED KGVQCEEEEEEKKDSGVASTEDSSSSHITAAAIAAKIPDSIISRGVQVLPRD TASLSTTPSESPRAQATSRLSTASCPTPKVQSRCSSKENILRASHSAVDIT KVARRHRMSPFPLTSMDKAFITVLEMTPVLGTEIINYRDGMGRVLAQDVYA KDNLPPFPASVKDGYAVRAADGPGDRFIIGESQAGEQPTQTVMPGQVMR VTTGAPIPCGADAVVQVEDTELIRESDDGTEELEVRILVQARPGQDIRPIGH DIKRGECVLAKGTHMGPSEIGLLATVGVTEVEVNKFPVVAVMSTGNELLNP EDDLLPGKIRDSNRSTLLATIQEHGYPTINLGIVGDNPDDLLNALNEGISRA DVIITSGGVSMGEKDYLKQVLDIDLHAQIHFGRVFMKPGLPTTFATLDIDGV RKIIFALPGNPVSAVVTCNLFVVPALRKMQGILDPRPTIIKARLSCDVKLDPR PEYHRCILTWHHQEPLPWAQSTGNQMSSRLMSMRSANGLLMLPPKTEQ YVELHKGEVVDVMVIGRL [SEQ ID NO:460]	Purina	Purina
MO-COS	55034	0075643	Q96EN8	MAGAAAESGRELWTFAGSRDPSAPRLAYGYGPGSLRELRAREFSRLAGT VYLDHAGATLFSQSQLESFTSDLMENTYGNPHSQNISSKLTHDTVEQVRY RILAHFHTTAEDYTVIFTAGSTAALKLVAEAFPWVSQGPESSGSRFCYLTD SHTSVVGMRNVTMAINVISTPVRPEDLWSAEERSASASNPDCQLPHLFCY PAQSNFSGVRYPLSWIEEVKSGRLHPVSTPGKWFVLLDAASYVSTSPLDL SAHQADFVPISFYKIFGFPTGLGALLVHNRAAPLLRKTYFGGGTASAYLAG EDFYIPRQSVAQRFEDGTISFLDVIALKHGFDTLERLTGGMENIKQHTFTLA QYTYVALSSLQYPNGAPVVRIYSDSEFSSPEVQGPIINFNVLDDKGNIIGYS QVDKMASLYNIHLRTGCFCNTGACQRHLGISNEMVRKHFQAGHVCGDNM DLIDGQPTGSVRISFGYMSTLDDVQAFLRFIIDTRLHSSGDWPVPQAHADT GETGAPSADSQADVIPAVMGRRSLSPQEDALTGSRVWNNSSTVNAVPVA PPVCDVARTQPTPSEKAAGVLEGALGPHVVTNLYLYPIKSCAAFEVTRWP VGNQGLLYDRSWMVVNHNGVCLSQKQEPRLCLIQPFIDLRQRIMVIKAKG MEPIEVPLEENSERTQIRQSRVCADRVSTYDCGEKISSWLSTFFGRPCHLI KQSSNSQRNAKKKHGKDQLPGTMATLSLVNEAQYLLINTSSILELHRQLNT SDENGKEELFSLKDLSLRFRANIIINGKRAFEEEKWDEISIGSLRFQVLGPC HRCQMICIDQQTGQRNQHVFQKLSESRETKVNFGMYLMHASLDLSSPCF LSVGSQVLPVLKENVEGHDLPASEKHQDVTS [SEQ ID NO:461]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
MOCS1	4337	0124615	A0A024RD17, Q9NZB8	MAARPLSRMLRRLLRSSARSCSSGAPVTQPCPGESARAASEEVSRRRQFLREHAAPFSAFLTDSFGRQHSYLRISLTEKCNLRCQYCMPEEGVPLTPKANLLTTEEILTLARLFVKEGIDKIRLTGGEPLIRPDVVDIVAQLQRLEGLRTIGVTTNGINLARLLPQLQKAGLSAINISLDTLVPAKFEFIVRRKGFHKVMEGIHKAIELGYNPVKVNCVVMRGLNEDELLDFAALTEGLPLDVRFIEYMPFDGNKWNFKKMVSYKEMLDTVRQQWPELEKVPEEESSTAKAFKIPGFQGQISFITSMSEHFCGTCNRLRITADGNLKVCLFGNSEVSLRDHLRAGASEQELLRIIGAAVGRKKRQHAGMFSISQMKNRPMILIELFLMFPNSPPANPSIFSWDPLHVQGLRPRMSFSSQVATLWKGCRVPQTPPLAQQRLGSGSFQRHYTSRADSDANSKCLSPGSWASAAPSGPQLTSEQLTHVDSEGRAAMVDVGRKPDTERVAVASAVVLLGPVAFKLVQQNQLKKGDALVVAQLAGVQAAKVTSQLIPLCHHVALSHIQVQLELDSTRHAVKIQASCRARGPTGVEMEALTSAAVAALTLYDMCKAVSRDIVLEEIKLISKTGGQRGDFHRA [SEQ ID NO:462]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
PNP	4860	0198805	P00491, V9HWH6	MENGYTYEDYKNTAEWLLSHTKHRPQVAIICGSGLGGLTDKLTQAQIFDYGEIPNFPRSTVPGHAGRLVFGFLNGRACVMMQGRFHMYEGYPLWKVTFPVRVFHLLGVDTLVVTNAAGGLNPKFEVGDIMLIRDHINLPGFSGQNPLRGPNDERFGDRFPAMSDAYDRTMRQRALSTWKQMGEQRELQEGTYVMVAGPSFETVAECRVLQKLGADAVGMSTVPEVIVARHCGLRVFGFSLITNKVIMDYESLEKANHEEVLAAGKQAAQKLEQFVSILMASIPLPDKAS [SEQ ID NO:463]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
XDH	7498	0158125	P47989	MTADKLVFFVNGRKVVEKNADPETTLLAYLRRKLGLSGTKLGCGEGGCGACTVMLSKYDRLQNKIVHFSANACLAPICSLHHVAVTTVEGIGSTKTRLHPVQERIAKSHGSQCGFCTPGIVMSMYTLLRNQPEPTMEEIENAFQGNLCRCTGYRPILQGFRTFARDGGCCGGDGNNPNCCMNQKKDHSVSLSPSLFKPEEFTPLDPTQEPIFPPELLRLKDTPRKQLRFEGERVTWIQASTLKELLDLKAQHPDAKLVVGNTEIGIEMKFKNMLFPMIVCPAWIPELNSVEHGPDGISFGAACPLSIVEKTLVDAVAKLPAQKTEVFRGVLEQLRWFAGKQVKSVASVGGNIITASPISDLNPVFMASGAKLTLVSRGTRRTVQMDHTFFPGYRKTLLSPEEILLSIEIPYSREGEYFSAFKQASRREDDIAKVTSGMRVLFKPGTTEVQELALCYGGMANRTISALKTTQRQLSKLWKEELLQDVCAGLAEELHLPPDAPGGMVDFRCTLTLSFFFKFYLTVLQKLGQENLEDKCGKLDPTFASATLLFQKDPPA	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DVQLFQEVPKGQSEEDMVGRPLPHLAADMQASGEAVYCDDIPRYENELS LRLVTSTRAHAKIKSIDTSEAKKVPGFVCFISADDVPGSNITGICNDETVFAK DKVTCVGHIIGAVVADTPEHTQRAAQGVKITYEELPAIITIEDAIKNNSFYGP ELKIEKGDLKKGFSEADNVVSGEIYIGGQEHFYLETHCTIAVPKGEAGEME LFVSTQNTMKTQSFVAKMLGVPANRIVVRVKRMGGGFGGKETRSTVVST AVALAAYKTGRPVRCMLDRDEDMLITGGRHPFLARYKVGFMKTGTVVALE VDHFSNVGNTQDLSQSIMERALFHMDNCYKIPNIRGTGRLCKTNLPSNTAF RGFGGPQGMLIAECWMSEVAVTCGMPAEEVRRKNLYKEGDLTHFNQKLE GFTLPRCWEECLASSQYHARKSEVDKFNKENCWKKRGLCIIPTKFGISFTV PFLNQAGALLHVYTDGSVLLTHGGTEMGQGLHTKMVQVASRALKIPTSKIY ISETSTNTVPNTSPTAASVSADLNGQAVYAACQTILKRLEPYKKKNPSGSW EDWVTAAYMDTVSLSATGFYRTPNLGYSFETNSGNPFHYFSYGVACSEV EIDCLTGDHKNLRTDIVMDVGSSLNPAIDIGQVEGAFVQGLGLFTLEELHYS PEGSLHTRGPSTYKIPAFGSIPIEFRVSLLRDCPNKKAIYASKAVGEPPLFLA ASIFFAIKDAIRAARAQHTGNNVKELFRLDSPATPEKIRNACVDKFTTLCVT GVPENCKPWSVRV [SEQ ID NO:464]		
SUOX	6821	0139531	A0A024RB79, P51687	MLLLHRAVVLRLQQACRLKSIPSRICIQACSTNDSFQPQRPSLTFSGDNSS TQGWRVMGTLLGLGAVLAYQDHRCRAAQESTHIYTKEEVSSHTSPETGI WVTLGSEVFDVTEFVDLHPGGPSKLMLAAGGPLEPFWALYAVHNQSHVR ELLAQYKIGELNPEDKVAPTVETSDPYADDPVRHPALKVNSQRPFNAEPP PELLTENYITPNPIFFTRNHLPVPNLDPDTYRLHVVGAPGGQSLSLSLDDLH NFPRYEITVTLQCAGNRRSEMTQVKEVKGLEWRTGAISTARWAGARLCD VLAQAGHQLCETEAHVCFEGLDSDPTGTAYGASIPLARAMDPEAEVLLAY EMNGQPLPRDHGFPVRVVVPGVVGARHVKWLGRVSVQPEESYSHWQR RDYKGFSPSVDWETVDFDSAPSIQELPVQSAITEPRDGETVESGEVTIKGY AWSGGGRAVIRVDVSLDGGLTWQVAKLDGEEQRPRKAWAWRLWQLKA PVPAGQKELNIVCKAVDDGYNVQPDTVAPIWNLRGVLSNAWHRVHVYVS P[SEQ ID NO:465]	Transtorno do Metabolismo da Purina	Transtorno do Metabolismo da Purina
OGDH	4967	0105953	A0A140VJQ5, Q02218, B4E3E9, E9PCR7, E9PDF2	MFHLRTCAAKLRPLTASQTVKTFSQNRPAAARTFQQIRCYSAPVAAEPFLS GTSSNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTAYQSPLPLS RGSLAAVAHAQSLVEAQPNVDKLVEDHLAVQSLIRAYQIRGHHVAQLDPL	Deficiência de 2-Cetoglutarato Desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GILDADLDSSVPADIISSTDKLGFYGLDESDLDKVFHLPTTTFIGGQESALPL REIIRRLEMAYCQHIGVEFMFINDLEQCQWIRQKFETPGIMQFTNEEKRTLL ARLVRSTRFEEFLQRKWSSEKRFGLEGCEVLIPALKTIIDKSSENGVDYVIM GMPHRGRLNVLANVIRKELEQIFCQFDSKLEAADEGSGDVKYHLGMYHRR INRVTDRNITLSLVANPSHLEAADPVVMGKTKAEQFYCGDTEGKKVMSILL HGDAAFAGQGIVYETFHLSDLPSYTTHGTVHVVVNNQIGFTTDPRMARSS PYPTDVARVVNAPIFHVNSDDPEAVMYVCKVAAEWRSTFHKDVVVDLVCY RRNGHNEMDEPMFTQPLMYKQIRKQKPVLQKYAELLVSQGVVNQPEYEE EISKYDKICEEAFARSKDEKILHIKHWLDSPWPGFFTLDGQPRSMSCPSTG LTEDILTHIGNVASSVPVENFTIHGGLSRILKTRGEMVKNRTVDWALAEYM AFGSLLKEGIHIRLSGQDVERGTFSHRHHVLHDQNVDKRTCIPMNHLWPN QAPYTVCNSSLSEYGVLGFELGFAMASPNALVLWEAQFGDFHNTAQCIID QFICPGQAKWVRQNGIVLLLPHGMEGMGPEHSSARPERFLQMCNDDPDV LPDLKEANFDINQLYDCNWVVVNCSTPGNFFHVLRRQILLPFRKPLIIFTPK SLLRHPEARSSFDEMLPGTHFQRVIPEDGPAAQNPENVKRLLFCTGKVYY DLTRERKARDMVGQVAITRIEQLSPFPFDLLLKEVQKYPNAELAWCQEEH KNQGYYDYVKPRLRTTISRAKPVWYAGRDPAAAPATGNKKTHLTELQRLL DTAFDLDVFKNFS [SEQ ID NO:466]		
SLC25A19	60386	0125454	Q5JPC1, Q9HC21	MVGYDPKPDGRNNTKFQVAVAGSVSGLVTRALISPFDVIKIRFQLQHERLS RSDPSAKYHGILQASRQILQEEGPTAFWKGHVPAQILSIGYGAVQFLSFEM LTELVHRGSVYDAREFSVHFVCGGLAACMATLTVHPVDVLRTRFAAQGEP KVYNTLRHAVGTMYRSEGPQVFYKGLAPTLIAIFPYAGLQFSCYSSLKHLY KWAIPAEGKKNENLQNLLCGSGAGVISKTLTYPLDLFKKRLQVGGFEHAR AAFGQVRRYKGLMDCAKQVLQKEGALGFFKGLSPSLLKAALSTGFMFFSY EFFCNVFHCMNRTASQR [SEQ ID NO:467]	Deficiência de 2-Cetoglutarato Desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO
DHTKD1	55526	0181192	Q96HY7	MASATAAAARRGLGRALPLFWRGYQTERGVYGYRPRKPESREPQGALE RPPVDHGLARLVTVYCEHGHKAAKINPLFTGQALLENVPEIQALVQTLQGP FHTAGLLNMGKEEASLEEVLVYLNQIYCGQISIETSQLQSQDEKDWFAKRF EELQKETFTTEERKHLSKLMLESQEFDHFLATKFSTVKRYGGEGAESMMG FFHELLKMSAYSGITDVIIGMPHRGRLNLLTGLLQFPPELMFRKMRGLSEF PENFSATGDVLSHLTSSVDLYFGAHHPLHVTMLPNPSHLEAVNPVAVGKT	Deficiência de 2-Cetoglutarato Desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				RGRQQSRQDGDYSPDNSAQPGDRVICLQVHGDASFCGQGIVPETFTLSN LPHFRIGGSVHLIVNNQLGYTTPAERGRSSLYCSDIGKLVGCAIIHVNGDSP EEVVRATRLAFEYQRQFRKDVIIDLLCYRQWGHNELDEPFYTNPIMYKIIRA RKSIPDTYAEHLIAGGLMTQEEVSEIKSSYYAKLNDHLNNMAHYRPPALNL QAHWQGLAQPEAQITTWSTGVPLDLLRFVGMKSVEVPRELQMHSHLLKT HVQSRMEKMMDGIKLDWATAEALALGSLLAQGFNVRLSGQDVGRGTFSQ RHAIVVCQETDDTYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFEYGMSIES PKLLPLWEAQFGDFFNGAQIIFDTFISGGEAKWLLQSGIVILLPHGYDGAGP DHSSCRIERFLQMCDSAEEGVDGDTVNMFVVHPTTPAQYFHLLRRQMVR NFRKPLIVASPKMLLRLPAAVSTLQEMAPGTTFNPVIGDSSVDPKKVKTLV FCSGKHFYSLVKQRESLGAKKHDFAIIRVEELCPFPLDSLQQEMSKYKHVK DHIWSQEEPQNMGPWSFVSPRFEKQLACKLRLVGRPPLPVPAVGIGTVHL HQHEDILAKTFA [SEQ ID NO:468]		
SLC13 A5	284111	0141485	Q68D44, Q86YT5	MASALSYVSKFKSFVILFVTPLLLLPLVILMPAKFVRCAYVIILMAIYWCTEVI PLAVTSLMPVLLFPLFQILDSRQVCVQYMKDTNMLFLGGLIVAVAVERWNL HKRIALRTLLWVGAKPARLMLGFMGVTALLSMWISNTATTAMMVPIVEAIL QQMEATSAATEAGLELVDKGKAKELPGSQVIFEGPTLGQQEDQERKRLC KAMTLCICYAASIGGTATLTGTGPNVVLLGQMNELFPDSKDLVNFASWFAF AFPNMLVMLLFAWLWLQFVYMRFNFKKSWGCGLESKKNEKAALKVLQEE YRKLGPLSFAEINVLICFFLLVILWFSRDPGFMPGWLTVAWVEGETKYVSD ATVAIFVATLLFIVPSQKPKFNFRSQTEEERKTPFYPPPLLDWKVTQEKVP WGIVLLLGGGFALAKGSEASGLSVWMGKQMEPLHAVPPAAITLILSLLVAV FTECTSNVATTTLFLPIFASMSRSIGLNPLYIMLPCTLSASFAFMLPVATPPN AIVFTYGHLKVADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDHFPDWAN VTHIET [SEQ ID NO:469]	Deficiência do Transportador de Citrato	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCI-DO TRICARBO-XÍLICO
FH	2271	0091483	A0A0S2Z4C3, P07954	MYRALRLLARSRPLVRAPAAALASAPGLGGAAVPSFWPPNAARMASQNS FRIEYDTFGELKVPNDKYYGAQTVRSTMNFKIGGVTERMPTPVIKAFGILK RAAAEVNQDYGLDPKIANAIMKAADEVAEGKLNDHFPLVVWQTGSGTQTN MNVNEVISNRAIEMLGGELGSKIPVHPNDHVNKSQSSNDTFPTAMHIAAAI EVHEVLLPGLQKLHDALDAKSKEFAQIIKIGRTHTQDAVPLTLGQEFSGYV QQVKYAMTRIKAAMPRIYELAAGGTAVGTGLNTRIGFAEKVAAKVAALTGL	Deficiência de Fumarase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCI-DO TRICARBO-XÍLICO

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				PFVTAPNKFEALAAHDALVELSGAMNTTACSLMKIANDIRFLGSGPRSGLG ELILPENEPGSSIMPGKVNPTQCEAMTMVAAQVMGNHVAVTVGGSNGHF ELNVFKPMMIKNVLHSARLLGDASVSFTENCVVGIQANTERINKLMNESLM LVTALNPHIGYDKAAKIAKTAHKNGSTLKETAIELGYLTAEQFDEWVKPKD MLGPK [SEQ ID NO:470]		
DLAT	1737	0150768	P10515, Q86YI5	MWRVCARRAQNVAPWAGLEARWTALQEVPGTPRVTSRSGPAPARRNSV TTGYGGVRALCGWTPSSGATPRNRLLLQLLGSPGRRYYSLPPHQKVPLP SLSPTMQAGTIARWEKKEGDKINEGDLIAEVETDKATVGFESLEECYMAKI LVAEGTRDVPIGAIICITVGKPEDIEAFKNYTLDSSAAPTPQAAPAPTPAATA SPPTPSAQAPGSSYPPHMQVLLPALSPTMTMGTVQRWEKKVGEKLSEGD LLAEIETDKATIGFEVQEEGYLAKILVPEGTRDVPLGTPLCIIVEKEADISAFA DYRPTEVTDLKPQVPPPTPPPVAAVPPTPQPLAPTPSAPCPATPAGPKGR VFVSPLAKKLAVEKGIDLTQVKGTGPDGRITKKDIDSFVPSKVAPAPAAVVP PTGPGMAPVPTGVFTDIPISNIRRVIAQRLMQSKQTIPHYYLSIDVNMGEVL LVRKELNKILEGRSKISVNDFIIKASALACLKVPEANSSWMDTVIRQNHVVD VSVAVSTPAGLITPIVFNAHIKGVETIANDVVSLATKAREGKLQPHEFQGGT FTISNLGMFGIKNFSAIINPPQACILAIGASEDKLVPADNEKGFDVASMMSV TLSCDHRVVDGAVGAQWLAEFRKYLEKPITMLL [SEQ ID NO:471]	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO
MPC1	51660	0060762	Q5TI65, Q9Y5U8	MAGALVRKAADYVRSKDFRDYLMSTHFWGPVANWGLPIAAINDMKKSPEI ISGRMTFALCCYSLTFMRFAYKVQPRNWLLFACHATNEVAQLIQGGRLIKH EMTKTASA [SEQ ID NO:472]	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO
PDHA 1	5160	0131828	A0A024RBX9, P08559	MRKMLAAVSRVLSGASQKPASRVLVASRNFANDATFEIKKCDLHRLEEGP PVTTVLTREDGLKYYRMMQTVRRMELKADQLYKQKIIRGFCHLCDGQEAC CVGLEAGINPTDHLITAYRAHGFTFTRGLSVREILAELTGRKGGCAKGKGG SMHMYAKNFYGGNGIVGAQVPLGAGIALACKYNGKDEVCLTLYGDGAAN QGQIFEAYNMAALWKLPCIFICENNRYGMGTSVERAAASTDYYKRGDFIP GLRVDGMDILCVREATRFAAAYCRSGKGPILMELQTYRYHGHSMSDPGV	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				SYRTREEIQEVRSKSDPIMLLKDRMVNSNLASVEELKEIDVEVRKEIEDAAQ FATADPEPPLEELGYHIYSSDPPFEVRGANQWIKFKSVS [SEQ ID NO:473]		
PDHB	5162	0168291	P11177	MAAVSGLVRRPLREVSGLLKRRFHWTAPAALQVTVRDAINQGMDEELER DEKVFLLGEEVAQYDGAYKVSRGLWKKYGDKRIIDTPISEMGFAGIAVGAA MAGLRPICEFMTFNFSMQAIDQVINSAAKTYYMSGGLQPVPIVFRGPNGA SAGVAAQHSQCFAAWYGHCPGLKVVSPWNSEDAKGLIKSAIRDNNPVVV LENELMYGVPFEFPPEAQSKDFLIPIGKAKIERQGTHITVVSHSRPVGHCLE AAAVLSKEGVECEVINMRTIRPMDMETIEASVMKTNHLVTVEGGWPQFGV GAEICARIMEGPAFNFLDAPAVRVTGADVPMPYAKILEDNSIPQVKDIIFAIK KTLNI [SEQ ID NO:474]	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO
PDHX	8050	0110435	O00330	MAASWRLGCDPRLLRYLVGFPGRRSVGLVKGALGWSVSRGANWRWFH STQWLRGDPIKILMPSLSPTMEEGNIVKWLKKEGEAVSAGDALCEIETDKA VVTLDASDDGILAKIVVEEGSKNIRLGSLIGLIVEEGEDWKHVEIPKDVGPPP PVSKPSEPRPSPEPQISIPVKKEHIPGTLRFRLSPAARNILEKHSLDASQGT ATGPRGIFTKEDALKLVQLKQTGKITESRPTPAPTATPTAPSPLQATAGPS YPRPVIPPVSTPGQPNAVGTFTEIPASNIRRVIAKRLTESKSTVPHAYATAD CDLGAVLKVRQDLVKDDIKVSVNDFIIKAAAVTLKQMPDVNVSWDGEGPK QLPFIDISVAVATDKGLLTPIIKDAAAKGIQEIADSVKALSKKARDGKLLPEEY QGGSFSISNLGMFGIDEFTAVINPPQACILAVGRFRPVLKLTEDEEGNAKL QQRQLITVTMSSDSRVVDDELATRFLKSFKANLENPIRLA [SEQ ID NO:475]	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO
PDP1	54704	0164951	Q9P0J1, Q6P1N1, A0A024R9C0	MPAPTQLFFPLIRNCELSRIYGTACYCHHKHLCCSSSYIPQSRLRYTPHPA YATFCRPKENWWQYTQGRRYASTPQKFYLTPPQVNSILKANEYSFKVPEF DGKNVSSILGFDSNQLPANAPIEDRRSAATCLQTRGMLLGVFDGHAGCAC SQAVSERLFYYIAVSLLPHETLLEIENAVESGRALLPILQWHKHPNDYFSKE ASKLYFNSLRTYWQELIDLNTGESTDIDVKEALINAFKRLDNDISLEAQVGD PNSFLNYLVLRVAFSGATACVAHVDGVDLHVANTGDSRAMLGVQEEDGS WSAVTLSNDHNAQNERELERLKLEHPKSEAKSVVKQDRLLGLLMPFRAFG DVKFKWSIDLQKRVIESGPDQLNDNEYTKFIPPNYHTPPYLTAEPEVTYHR LRPQDKFLVLATDGLWETMHRQDVVRIVGEYLTGMHHQQPIAVGGYKVTL	Deficiência de piruvato desidrogenase	METABOLISMO DE PIRUVATO E DEFEITO DO CICLO DE ÁCIDO TRICARBOXÍLICO

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GQMHGLLTERRTKMSSVFEDQNAATHLIRHAVGNNEFGTVDHERLSKML SLPEELARMYRDDITIIVVQFNSHVVGAYQNQE [SEQ ID NO:476]		
ABCC 2	1244	0023839	Q92887	MLEKFCNSTFWNSSFLDSPEADLPLCFEQTVLVWIPLGYLWLLAPWQLLH VYKSRTKRSSTTKLYLAKQVFVGFLLILAAIELALVLTEDSGQATVPAVRYT NPSLYLGTWLLVLLIQYSRQWCVQKNSWFLSLFWILSILCGTFQFQTLIRTL LQGDNSNLAYSCLFFISYGFQILILIFSAFSENNESSNNPSSIASFLSSITYSW YDSIILKGYKRPLTLEDVWEVDEEMKTKTLVS KFETHMKRELQKARRALQRRQEKSSQQNSGARLPGLNKNQSQSQDALVL EDVEKKKKKSGTKKDVPKSWLMKALFKTFYMVLLKSFLLKLVNDIFTFVSP QLLKLLISFASDRDTYLWIGYLCAILLFTAALIQSFCLQCYFQLCFKLGVKVR TAIMASVYKKALTLSNLARKEYTVGETVNLMSVDAQKLMDVTNFMHMLWS SVLQIVLSIFFLWRELGPSVLAGVGVMVLVIPINAILSTKSKTIQVKNMKNKD KRLKIMNEILSGIKILKYFAWEPSFRDQVQNLRKKELKNLLAFSQLQCVVIFV FQLTPVLVSVVTFSVYVLVDSNNILDAQKAFTSITLFNILRFPLSMLPMMISS MLQASVSTERLEKYLGGDDLDTSAIRHDCNFDKAMQFSEASFTWEHDSE ATVRDVNLDIMAGQLVAVIGPVGSGKSSLISAMLGEMENVHGHITIKGTTA YVPQQSWIQNGTIKDNILFGTEFNEKRYQQVLEACALLPDLEMLPGGDLAE IGEKGINLSGGQKQRISLARATYQNLDIYLLDDPLSAVDAHVGKHIFNKVLG PNGLLKGKTRLLVTHSMHFLPQVDEIVVLGNGTIVEKGSYSALLAKKGEFA KNLKTFLRHTGPEEEATVHDGSEEEDDDYGLISSVEEIPEDAASITMRREN SFRRTLSRSSRSNGRHLKSLRNSLKTRNVNSLKEDEELVKGQKLIKKEFIE TGKVKFSIYLEYLQAIGLFSIFFILAFVMNSVAFIGSNLWLSAWTSDSKIFNS TDYPASQRDMRVGVYGALGLAQGIFVFIAHFWSAFGFVHASNILHKQLLN NILRAPMRFFDTTPTGRIVNRFAGDISTVDDTLPQSLRSWITCFLGIISTLVMI CMATPVFTIIVIPLGIIYVSVQMFYVSTSRQLRRLDSVTRSPIYSHFSETVSG LPVIRAFEHQQRFLKHNEVRIDTNQKCVFSWITSNRWLAIRLELVGNLTVFF SALMMVIYRDTLSGDTVGFVLSNALNITQTLNWLVRMTSEIETNIVAVERIT EYTKVENEAPWVTDKRPPPDWPSKGKIQFNNYQVRYRPELDLVLRGITCD IGSMEKIGVVGRTGAGKSSLTNCLFRILEAAGGQIIIDGVDIASIGLHDLREK LTIIPQDPILFSGSLRMNLDPFNNYSDEEIWKALELAHLKSFVASLQLGLSH EVTEAGGNLSIGQRQLLCLGRALLRKSKILVLDEATAAVDLETDNLIQTTIQ		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				NEFAHCTVITIAHRLHTIMDSDKVMVLDNGKIIECGSPEELLQIPGPFYFMAK EAGIENVNSTKF [SEQ ID NO:477]		
SLCO 1B1	10599	0134538	A0A024RAU7, Q05CV5, Q9Y6L6	MDQNQHLNKTAEAQPSENKKTRYCNGLKMFLAALSLSFIAKTLGAIIMKSSI IHIERRFEISSSLVGFIDGSFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGIG GVLTALPHFFMGYYRYSKETNINSSENSTSTLSTCLINQILSLNRASPEIVG KGCLKESGSYMWIYVFMGNMLRGIGETPIVPLGLSYIDDFAKEGHSSLYLG ILNAIAMIGPIIGFTLGSLFSKMYVDIGYVDLSTIRITPTDSRWVGAWWLNFL VSGLFSIISSIPFFFLPQTPNKPQKERKASLSLHVLETNDEKDQTANLTNQG KNITKNVTGFFQSFKSILTNPLYVMFVLLTLLQVSSYIGAFTYVFKYVEQQY GQPSSKANILLGVITIPIFASGMFLGGYIIKKFKLNTVGIAKFSCFTAVMSLSF YLLYFFILCENKSVAGLTMTYDGNNPVTSHRDVPLSYCNSDCNCDESQWE PVCGNNGITYISPCLAGCKSSSGNKKPIVFYNCSCLEVTGLQNRNYSAHLG ECPRDDACTRKFYFFVAIQVLNLFFSALGGTSHVMLIVKIVQPELKSLALGF HSMVIRALGGILAPIYFGALIDTTCIKWSTNNCGTRGSCRTYNSTSFSRVYL GLSSMLRVSSLVLYIILIYAMKKKYQEKDINASENGSVMDEANLESLNKNKH FVPSAGADSETHC [SEQ ID NO:478]	Síndrome de Rotor	
SLCO 1B3	28234	0111700	B3KP78, Q9NPD5	MDQHQHLNKTAESASSEKKKTRRCNGFKMFLAALSFSYIAKALGGIIMKISI TQIERRFDISSSLAGLIDGSFEIGNLLVIVFVSYFGSKLHRPKLIGIGCLLMGT GSILTSLPHFFMGYYRYSKETHINPSENSTSSLSTCLINQTLSFNGTSPEIVE KDCVKESGSHMWIYVFMGNMLRGIGETPIVPLGISYIDDFAKEGHSSLYLG SLNAIGMIGPVIGFALGSLFAKMYVDIGYVDLSTIRITPKDSRWVGAWWLGF LVSGLFSIISSIPFFFLPKNPNKPQKERKISLSLHVLKTNDDRNQTANLTNQG KNVTKNVTGFFQSLKSILTNPLYVIFLLLTLLQVSSFIGSFTYVFKYMEQQYG QSASHANFLLGIITIPTVATGMFLGGFIIKKFKLSLVGIAKFSFLTSMISFLFQL LYFPLICESKSVAGLTLTYDGNNSVASHVDVPLSYCNSECNCDESQWEPV CGNNGITYLSPCLAGCKSSSGIKKHTVFYNCSCVEVTGLQNRNYSAHLGE CPRDNTCTRKFFIYVAIQVINSLFSATGGTTFILLTVKIVQPELKALAMGFQS MVIRTLGGILAPIYFGALIDKTCMKWSTNSCGAQGACRIYNSVFFGRVYLG LSIALRFPALVLYIVFIFAMKKKFQGKDTKASDNERKVMDEANLEFLNNGEH FVPSAGTDSKTCNLDMQDNAAAN[SEQ ID NO:479]	Síndrome de Rotor	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
HFE2	148738	0168509	Q6ZVN8, A8K466, A0A024R4F5	MGEPGQSPSPRSSHGSPPTLSTLTLLLLLCGHAHSQCKILRCNAEYVSSTL SLRGGGSSGALRGGGGGGRGGGVGSGGLCRALRSYALCTRRTARTCR GDLAFHSAVHGIEDLMIQHNCSRQGPTAPPPPRGPALPGAGSGLPAPDPC DYEGRFSRLHGRPPGFLHCASFGDPHVRSFHHHFHTCRVQGAWPLLDN DFLFVQATSSPMALGANATATRKLTIIFKNMQECIDQKVYQAEVDNLPVAF EDGSINGGDRPGGSSLSIQTANPGNHVEIQAAYIGTTIIIRQTAGQLSFSIKV AEDVAMAFSAEQDLQLCVGGCPPSQRLSRSERNRRGAITIDTARRLCKEG LPVEDAYFHSCVFDVLISGDPNFTVAAQAALEDARAFLPDLEKLHLFPSDA GVPLSSATLLAPLLSGLFVLWLCIQ [SEQ ID NO:480]	Hemocromatose, tipo 2A	
ADAM TS13	11093	0160323,0281244	Q76LX8	MHQRHPRARCPPLCVAGILACGFLLGCWGPSHFQQSCLQALEPQAVSSY LSPGAPLKGRPPSPGFQRQRQRQRRAAGGILHLELLVAVGPDVFQAHQE DTERYVLTNLNIGAELLRDPSLGAQFRVHLVKMVILTEPEGAPNITANLTSS LLSVCGWSQTINPEDDTDPGHADLVLYITRFDLELPDGNRQVRGVTQLGG ACSPTWSCLITEDTGFDLGVTIAHEIGHSFGLEHDGAPGSGCGPSGHVMA SDGAAPRAGLAWSPCSRRQLLSLLSAGRARCVWDPPRPQPGSAGHPPD AQPGLYYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSC SRLLVPLLDGTECGVEKWCSKGRCRSLVELTPIAAVHGRWSSWGPRSPC SRSCGGGVVTRRRQCNNPRPAFGGRACVGADLQAEMCNTQACEKTQLE FMSQQCARTDGQPLRSSPGGASFYHWGAAVPHSQGDALCRHMCRAIGE SFIMKRGDSFLDGTRCMPSGPREDGTLSLCVSGSCRTFGCDGRMDSQQ VWDRCQVCGGDNSTCSPRKGSFTAGRAREYVTFLTVTPNLTSVYIANHR PLFTHLAVRIGGRYVVAGKMSISPNTTYPSLLEDGRVEYRVALTEDRLPRL EEIRIWGPLQEDADIQVYRRYGEEYGNLTRPDITFTYFQPKPRQAWVWAA VRGPCSVSCGAGLRWVNYSCLDQARKELVETVQCQGSQQPPAWPEACV LEPCPPYWAVGDFGPCSASCGGGLRERPVRCVEAQGSLLKTLPPARCRA GAQQPAVALETCNPQPCPARWEVSEPSSCTSAGGAGLALENETCVPGAD GLEAPVTEGPGSVDEKLPAPEPCVGMSCPPGWGHLDATSAGEKAPSPW GSIRTGAQAAHVWTPAAGSCSVSCGRGLMELRFLCMDSALRVPVQEELC GLASKPGSRREVCQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHG EDDGEEILLDTQCQGLPRPEPQEACSLEPCPPRWKVMSLGPCSASCGLG TARRSVACVQLDQGQDVEVDEAACAALVRPEASVPCLIADCTYRWHVGT	Púrpura trombocitopênica trombótica congênita devido à deficiência de ADAMTS-13	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				WMECSVSCGDGIQRRRDTCLGPQAQAPVPADFCQHLPKPVTVRGCWAG PCVGQGTPSLVPHEEAAAPGRTTATPAGASLEWSQARGLLFSPAPQPRR LLPGPQENSVQSSACGRQHLEPTGTIDMRGPGQADCAVAIGRPLGEVVTL RVLESSLNCSAGDMLLLWGRLTWRKMCRKLLDMTFSSKTNTLVVRQRCG RPGGGVLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGC RLFINVAPHARIAIHALATNMGAGTEGANASYILIRDTHSLRTTAFHGQQVL YWESESSQAEMEFSEGFLKAQASLRGQYWTLQSWVPEMQDPQSWKGK EGT [SEQ ID NO:481]		
PYGM	5837	0068976	P11217	MSRPLSDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATPR DYYFALAHTVRDHLVGRWIRTQQHYYEKDPKRIYYLSLEFYMGRTLQNTM VNLALENACDEATYQLGLDMEELEEIEEDAGLGNGGLGRLAACFLDSMAT LGLAAYGYGIRYEFGIFNQKISGGWQMEEADDWLRYGNPWEKARPEFTL PVHFYGHVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMRLWSA KAPNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKELRLKQEY FVVAATLQDIIRRFKSSKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPEL MRILVDLERMDWDKAWDVTVRTCAYTNHTVLPEALERWPVHLLETLLPRH LQIIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLCIAGSHA VNGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVLCNPGLAEVI AERIGEDFISDLDQLRKLLSFVDDEAFIRDVAKVKQENKLKFAAYLEREYKV HINPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKREPNKFFVPRTVMIGGK AAPGYHMAKMIIRLVTAIGDVVNHDPAVGDRLRVIFLENYRVSLAEKVIPAA DLSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEEAGEENFFI FGMRVEDVDKLDQRGYNAQEYYDRIPELRQVIEQLSSGFFSPKQPDLFKD IVNMLMHHDRFKVFADYEDYIKCQEKVSALYKNPREWTRMVIRNIATSGKF SSDRTIAQYAREIWGVEPSRQRLPAPDEAI[SEQ ID NO:482]	Doença de McArdle	
COL1A2	1278	0164692	A0A0S2Z3H5, P08123	MLSFVDTRTLLLLAVTLCLATCQSLQEETVRKGPAGDRGPRGERGPPGPP GRDGEDGPTGPPGPPGPPGPPGPPGLGGNFAAQYDGKGVGLGPGPMGLMG PRGPPGAAGAPGPQGFQGPAGEPGEPGQTGPAGARGPAGPPGKAGED GHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKGIRGHNGLDGLKGQP GAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDGS VGPVGPAGPIGSAGPPGFPGAPGPKGEIGAVGNAGPAGPAGPRGEVGLP	Síndrome de Ehlers-Danlos, tipo cardíaco-valvular	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GLSGPVGPPGNPGANGLTGAKGAAGLPGVAGAPGLPGPRGIPGPVGAA GATGARGLVGEPGPAGSKGESGNKGEPGSAGPQGPPGPSGEEGKRGP NGEAGSAGPPGPPGLRGSPGSRGLPGADGRAGVMGPPGSRGASGPAG VRGPNGDAGRPGEPGLMGPRGLPGSPGNIGPAGKEGPVGLPGIDGRPG PIGPAGARGEPGNIGFPGPKGPTGDPGKNGDKGHAGLAGARGAPGPDG NNGAQGPPGPQGVQGGKGEQGPPGPPGFQGLPGPSGPAGEVGKPGER GLHGEFGLPGPAGPRGERGPPGESGAAGPTGPIGSRGPSGPPGPDGNK GEPGVVGAVGTAGPSGPSGLPGERGAAGIPGGKGEKGEPGLRGEIGNPG RDGARGAPGAVGAPGPAGATGDRGEAGAAGPAGPAGPRGSPGERGEV GPAGPNGFAGPAGAAGQPGAKGERGAKGPKGENGVVGPTGPVGAAGP AGPNGPPGPAGSRGDGGPPGMTGFPGAAGRTGPPGPSGISGPPGPPGP AGKEGLRGPRGDQGPVGRTGEVGAVGPPGFAGEKGPSGEAGTAGPPGT PGPQGLLGAPGILGLPGSRGERGLPGVAGAVGEPGPLGIAGPPGARGPP GAVGSPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNIGP VGAAGAPGPHGPVGPAGKHGNRGETGPSGPVGPAGAVGPRGPSGPQGI RGDKGEPGEKGPRGLPGLKGHNGLQGLPGIAGHHGDQGAPGSVGPAGP RGPAGPSGPAGKDGRTGHPGTVGPAGIRGPQGHQGPAGPPGPPGPPG PPGVSGGGYDFGYDGDFYRADQPRSAPSLRPKDYEVDATLKSLNNQIETL LTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFS TGETCIRAQPENIPAKNWYRSSKDKKHVWLGETINAGSQFEYNVEGVTSK EMATQLAFMRLLANYASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVE LVAEGNSRFTYTVLVDGCSKKTNEWGKTIIEYKTNKPSRLPFLDIAPLDIGG ADQEFFVDIGPVCFK [SEQ ID NO:483]		
TNFRSF11B	4982	0164761	O00300	MNNLLCCALVFLDISIKWTTQETFPPKYLHYDEETSHQLLCDKCPPGTYLK QHCTAKWKTVCAPCPDHYYTDSWHTSDECLYCSPVCKELQYVKQECNR THNRVCECKEGRYLEIEFCLKHRSCPPGFGVVQAGTPERNTVCKRCPDG FFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKCGID VTLCEEAFFRFAVPTKFTPNWLSVLVDNLPGTKVNAESVERIKRQHSSQE QTFQLLKLWKHQNKDQDIVKKIIQDIDLCENSVQRHIGHANLTFEQLRSLME SLPGKKVGAEDIEKTIKACKPSDQILKLLSLWRIKNGDQDTLKGLMHALKHS KTYHFPKTVTQSLKKTIRFLHSFTMYKLYQKLFLEMIGNQVQSVKISCL [SEQ ID NO:484]	Doença de Paget Juvenil	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
TSC1	7248	0165699	Q86WV8, Q92574, X5D9D2, Q32NF0	MAQQANVGELLAMLDSPMLGVRDDVTAVFKENLNSDRGPMLVNTLVDYY LETSSQPALHILTTLQEPHDKHLLDRINEYVGKAATRLSILSLLGHVIRLQPS WKHKLSQAPLLPSLLKCLKMDTDVVVLTTGVLVLITMLPMIPQSGKQHLLD FFDIFGRLSSWCLKKPGHVAEVYLVHLHASVYALFHRLYGMYPCNFVSFL RSHYSMKENLETFEEVVKPMMEHVRIHPELVTGSKDHELDPRRWKRLET HDVVIECAKISLDPTEASYEDGYSVSHQISARFPHRSADVTTSPYADTQNS YGCATSTPYSTSRLMLLNMPGQLPQTLSSPSTRLITEPPQATLWSPSMVC GMTTPPTSPGNVPPDLSHPYSKVFGTTAGGKGTPLGTPATSPPPAPLCHS DDYVHISLPQATVTPPRKEERMDSARPCLHRQHHLLNDRGSEEPPGSKG SVTLSDLPGFLGDLASEEDSIEKDKEEAAISRELSEITTAEAEPVVPRGGFD SPFYRDSLPGSQRKTHSAASSSQGASVNPEPLHSSLDKLGPDTPKQAFTP IDLPCGSADESPAGDRECQTSLETSIFTPSPCKIPPPTRVGFGSGQPPPYD HLFEVALPKTAHHFVIRKTEELLKKAKGNTEEDGVPSTSPMEVLDRLIQQG ADAHSKELNKLPLPSKSVDWTHFGGSPPSDEIRTLRDQLLLLHNQLLYERF KRQQHALRNRRLLRKVIKAAALEEHNAAMKDQLKLQEKDIQMWKVSLQKE QARYNQLQEQRDTMVTKLHSQIRQLQHDREEFYNQSQELQTKLEDCRNM IAELRIELKKANNKVCHTELLLSQVSQKLSNSESVQQQMEFLNRQLLVLGE VNELYLEQLQNKHSDTTKEVEMMKAAYRKELEKNRSHVLQQTQRLDTSQ KRILELESHLAKKDHLLLEQKKYLEDVKLQARGQLQAAESRYEAQKRITQV FELEILDLYGRLEKDGLLKKLEEEKAEAAEAAEERLDCCNDGCSDSMVGH NEEASGHNGETKTPRPSSARGSSGSRGGGGSSSSSSELSTPEKPPHQR AGPFSSRWETTMGEASASIPTTVGSLPSSKSFLGMKARELFRNKSESQCD EDGMTSSLSESLKTELGKDLGVEAKIPLNLDGPHPSPPTPDSVGQLHIMDY NETHHEHS [SEQ ID NO:485]	Esclerose tuberosa	
TSC2	7249	0103197	P49815, X5D7Q2, B3KWH7, Q5HYF7, H3BMQ0, X5D2U8	MAKPTSKDSGLKEKFKILLGLGTPRPNPRSAEGKQTEFIITAEILRELSMEC GLNNRIRMIGQICEVAKTKKFEEHAVEALWKAVADLLQPERPLEARHAVLA LLKAIVQGQGERLGVLRALFFKVIKDYPSNEDLHERLEVFKALTDNGRHITY LEEELADFVLQWMDVGLSSEFLLVLVNLVKFNSCYLDEYIARMVQMICLLC VRTASSVDIEVSLQVLDAVVCYNCLPAESLPLFIVTLCRTINVKELCEPCWK LMRNLLGTHLGHSAIYNMCHLMEDRAYMEDAPLLRGAVFFVGMALWGAH RLYSLRNSPTSVLPSFYQAMACPNEVVSYEIVLSITRLIKKYRKELQVVAWD	Esclerose tuberosa	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ILLNIIERLLQQLQTLDSPELRTIVHDLLTTVEELCDQNEFHGSQERYFELVE RCADQRPESSLLNLISYRAQSIHPAKDGWIQNLQALMERFFRSESRGAVRI KVLDVLSFVLLINRQFYEEELINSVVISQLSHIPEDKDHQVRKLATQLLVDLA EGCHTHHFNSLLDIIEKVMARSLSPPPELEERDVAAYSASLEDVKTAVLGLL VILQTKLYTLPASHATRVYEMLVSHIQLHYKHSYTLPIASSIRLQAFDFLLLL RADSLHRLGLPNKDGVVRFSPYCVCDYMEPERGSEKKTSGPLSPPTGPP GPAPAGPAVRLGSVPYSLLFRVLLQCLKQESDWKVLKLVLGRLPESLRYK VLIFTSPCSVDQLCSALCSMLSGPKTLERLRGAPEGFSRTDLHLAVVPVLT ALISYHNYLDKTKQREMVYCLEQGLIHRCASQCVVALSICSVEMPDIIIKALP VLVVKLTHISATASMAVPLLEFLSTLARLPHLYRNFAAEQYASVFAISLPYTN PSKFNQYIVCLAHHVIAMWFIRCRLPFRKDFVPFITKGLRSNVLLSFDDTPE KDSFRARSTSLNERPKSLRIARPPKQGLNNSPPVKEFKESSAAEAFRCRSI SVSEHVVRSRIQTSLTSASLGSADENSVAQADDSLKNLHLELTETCLDMM ARYVFSNFTAVPKRSPVGEFLLAGGRTKTWLVGNKLVTVTTSVGTGTRSL LGLDSGELQSGPESSSSPGVHVRQTKEAPAKLESQAGQQVSRGARDRV RSMSGGHGLRVGALDVPASQFLGSATSPGPRTAPAAKPEKASAGTRVPV QEKTNLAAYVPLLTQGWAEILVRRPTGNTSWLMSLENPLSPFSSDINNMP LQELSNALMAAERFKEHRDTALYKSLSVPAASTAKPPPLPRSNTVASFSSL YQSSCQGQLHRSVSWADSAVVMEEGSPGEVPVLVEPPGLEDVEAALGM DRRTDAYSRSSSVSSQEEKSLHAEELVGRGIPIERVVSSEGGRPSVDLSF QPSQPLSKSSSSPELQTLQDILGDPGDKADVGRLSPEVKARSQSGTLDGE SAAWSASGEDSRGQPEGPLPSSSPRSPSGLRPRGYTISDSAPSRRGKRV ERDALKSRATASNAEKVPGINPSFVFLQLYHSPFFGDESNKPILLPNESQS FERSVQLLDQIPSYDTHKIAVLYVGEGQSNSELAILSNEHGSYRYTEFLTGL GRLIELKDCQPDKVYLGGLDVCGEDGQFTYCWHDDIMQAVFHIATLMPTK DVDKHRCDKKRHLGNDFVSIVYNDSGEDFKLGTIKGQFNFVHVIVTPLDYE CNLVSLQCRKDMEGLVDTSVAKIVSDRNLPFVARQMALHANMASQVHHS RSNPTDIYPSKWIARLRHIKRLRQRICEEAAYSNPSLPLVHPPSHSKAPAQT PAEPTPGYEVGQRKRLISSVEDFTEFV[SEQ ID NO:486]		
DHCR7	1717	0172893	A0A024R5F7, Q9UBM7	MAAKSQPNIPKAKSLDGVTNDRTASQGQWGRAWEVDWFSLASVIFLLLFA PFIVYYFIMACDQYSCALTGPVVDIVTGHARLSDIWAKTPPITRKAAQLYTL WVTFQVLLYTSLPDFCHKFLPGYVGGIQEGAVTPAGVVNKYQINGLQAWL	Síndrome de Smith-Lemli-Opitz	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LTHLLWFANAHLLSWFSPTIIFDNWIPLLWCANILGYAVSTFAMVKGYFFPT SARDCKFTGNFFYNYMMGIEFNPRIGKWFDFKLFFNGRPGIVAWTLINLSF AAKQRELHSHVTNAMVLVNVLQAIYVIDFFWNETWYLKTIDICHDHFGWYL GWGDCVWLPYLYTLQGLYLVYHPVQLSTPHAVGVLLLGLVGYYIFRVANH QKDLFRRTDGRCLIWGRKPKVIECSYTSADGQRHHSKLLVSGFWGVARH FNYVGDLMGSLAYCLACGGGHLLPYFYIIYMAILLTHRCLRDEHRCASKYG RDWERYTAAVPYRLLPGIF [SEQ ID NO:487]		
PGK1	5230	0102144	P00558, V9HWF4	MSLSNKLTLDKLDVKGKRVVMRVDFNVPMKNNQITNNQRIKAAVPSIKFCL DNGAKSVVLMSHLGRPDGVPMPDKYSLEPVAVELKSLLGKDVLFLKDCVG PEVEKACANPAAGSVILLENLRFHVEEEGKGKDASGNKVKAEPAKIEAFRA SLSKLGDVYVNDAFGTAHRAHSSMVGVNLPQKAGGFLMKKELNYFAKAL ESPERPFLAILGGAKVADKIQLINNMLDKVNEMIIGGGMAFTFLKVLNNMEI GTSLFDEEGAKIVKDLMSKAEKNGVKITLPVDFVTADKFDENAKTGQATVA SGIPAGWMGLDCGPESSKKYAEAVTRAKQIVWNGPVGVFEWEAFARGTK ALMDEVVKATSRGCITIIGGGDTATCCAKWNTEDKVSHVSTGGGASLELLE GKVLPGVDALSNI [SEQ ID NO:488]	D-glicericacidemia	
VLDLR	7436	0147852	P98155, Q5VVF5	MGTSALWALWLLLALCWAPRESGATGTGRKAKCEPSQFQCTNGRCITLL WKCDGDEDCVDGSDEKNCVKKTCAESDFVCNNGQCVPSRWKCDGDPD CEDGSDESPEQCHMRTCRIHEISCGAHSTQCIPVSWRCDGENDCDSGED EENCGNITCSPDEFTCSSGRCISRNFVCNGQDDCSDGSDELDCAPPTCG AHEFQCSTSSCIPISWVCDDDADCSDQSDESLEQCGRQPVIHTKCPASEI QCGSGECIHKKWRCDGDPDCKDGSDEVNCPSRTCRPDQFECEDGSCIH GSRQCNGIRDCVDGSDEVNCKNVNQCLGPGKFKCRSGECIDISKVCNQE QDCRDWSDEPLKECHINECLVNNGGCSHICKDLVIGYECDCAAGFELIDR KTCGDIDECQNPGICSQICINLKGGYKCECSRGYQMDLATGVCKAVGKEP SLIFTNRRDIRKIGLERKEYIQLVEQLRNTVALDADIAAQKLFWADLSQKAIF SASIDDKVGRHVKMIDNVYNPAAIAVDWVYKTIYWTDAASKTISVATLDGT KRKFLFNSDLREPASIAVDPLSGFVYWSDWGEPAKIEKAGMNGFDRRPLV TADIQWPNGITLDLIKSRLYWLDSKLHMLSSVDLNGQDRRIVLKSLEFLAHP LALTIFEDRVYWIDGENEAVYGANKFTGSELATLVNNLNDAQDIIVYHELVQ PSGKNWCEEDMENGGCEYLCLPAPQINDHSPKYTCSCPSGYNVEENGR	Síndrome de Desequilíbrio	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DCQSTATTVTYSETKDTNTTEISATSGLVPGGINVTTAVSEVSVPPKGTSA AWAILPLLLLVMAAVGGYLMWRNWQHKNMKSMNFDNPVYLKTTEEDLSI DIGRHSASVGHTYPAISVVSTDDDLA [SEQ ID NO:489]		
KYNU	8942	0115919	Q16719	MEPSSLELPADTVQRIAAELKCHPTDERVALHLDEEDKLRHFRECFYIPKIQ DLPPVDLSLVNKDENAIYFLGNSLGLQPKMVKTYLEEELDKWAKIAAYGHE VGKRPWITGDESIVGLMKDIVGANEKEIALMNALTVNLHLLMLSFFKPTPKR YKILLEAKAFPSDHYAIESQLQLHGLNIEESMRMIKPREGEETLRIEDILEVIE KEGDSIAVILFSGVHFYTGQHFNIPAITKAGQAKGCYVGFDLAHAVGNVEL YLHDWGVDFACWCSYKYLNAGAGGIAGAFIHEKHAHTIKPALVGWFGHEL STRFKMDNKLQLIPGVCGFRISNPPILLVCSLHASLEIFKQATMKALRKKSV LLTGYLEYLIKHNYGKDKAATKKPVVNIITPSHVEERGCQLTITFSVPNKDVF QELEKRGVVCDKRNPNGIRVAPVPLYNSFHDVYKFTNLLTSILDSAETKN [SEQ ID NO: 490]	Encefalopatia devido à hidroxi-quinureninúria	
F5	2153	0198734	P12259	MFPGCPRLWVLVVLGTSWVGWGSQGTEAAQLRQFYVAAQGISWSYRPE PTNSSLNLSVTSFKKIVYREYEPYFKKEKPQSTISGLLGPTLYAEVGDIIKVH FKNKADKPLSIHPQGIRYSKLSEGASYLDHTFPAEKMDDAVAPGREYTYE WSISEDSGPTHDDPPCLTHIYYSHENLIEDFNSGLIGPLLICKKGTLTEGGT QKTFDKQIVLLFAVFDESKSWSQSSSLMYTVNGYVNGTMPDITVCAHDHIS WHLLGMSSGPELFSIHFNGQVLEQNHHKVSAITLVSATSTTANMTVGPEG KWIISSLTPKHLQAGMQAYIDIKNCPKKTRNLKKITREQRRHMKRWEYFIAA EEVIWDYAPVIPANMDKKYRSQHLDNFSNQIGKHYKKVMYTQYEDESFTK HTVNPNMKEDGILGPIIRAQVRDTLKIVFKNMASRPYSIYPHGVTFSPYEDE VNSSFTSGRNNTMIRAVQPGETYTYKWNILEFDEPTENDAQCLTRPYYSD VDIMRDIASGLIGLLLICKSRSLDRRGIQRAADIEQQAVFAVFDENKSWYLE DNINKFCENPDEVKRDDPKFYESNIMSTINGYVPESITTLGFCFDDTVQWH FCSVGTQNEILTIHFTGHSFIYGKRHEDTLTLFPMRGESVTVTMDNVGTWM LTSMNSSPRSKKLRLKFRDVKCIPDDDEDSYEIFEPPESTVMATRKMHDRL EPEDEESDADYDYQNRLAAALGIRSFRNSSLNQEEEEFNLTALALENGTEF VSSNTDIIVGSNYSSPSNISKFTVNNLAEPQKAPSHQQATTAGSPLRHLIGK NSVLNSSTAEHSSPYSEDPIEDPLQPDVTGIRLLSLGAGEFKSQEHAKHKG PKVERDQAAKHRFSWMKLLAHKVGRHLSQDTGSPSGMRPWEDLPSQDT	Deficiência de Fator V	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				GSPSRMRPWKDPPSDLLLLKQSNSSKILVGRWHLASEKGSYEIIQDTDED TAVNNWLISPQNASRAWGESTPLANKPGKQSGHPKFPRVRHKSLQVRQD GGKSRLKKSQFLIKTRKKKKEKHTHHAPLSPRTFHPLRSEAYNTFSERRLK HSLVLHKSNETSLPTDLNQTLPSMDFGWIASLPDHNQNSSNDTGQASCPP GLYQTVPPEEHYQTFPIQDPDQMHSTSDPSHRSSSPELSEMLEYDRSHK SFPTDISQMSPSSEHEVWQTVISPDLSQVTLSPELSQTNLSPDLSHTTLSP ELIQRNLSPALGQMPISPDLSHTTLSPDLSHTTLSLDLSQTNLSPELSQTNL SPALGQMPLSPDLSHTTLSLDFSQTNLSPELSHMTLSPELSQTNLSPALGQ MPISPDLSHTTLSLDFSQTNLSPELSQTNLSPALGQMPLSPDPSHTTLSLD LSQTNLSPELSQTNLSPDLSEMPLFADLSQIPLTPDLDQMTLSPDLGETDL SPNFGQMSLSPDLSQVTLSPDISDTTLLPDLSQISPPPDLDQIFYPSESSQS LLLQEFNESFPYPDLGQMPSPSSPTLNDTFLSKEFNPLVIVGLSKDGTDYIE IIPKEEVQSSEDDYAEIDYVPYDDPYKTDVRTNINSSRDPDNIAAWYLRSNN GNRRNYYIAAEEISWDYSEFVQRETDIEDSDDIPEDTTYKKVVFRKYLDST FTKRDPRGEYEEHLGILGPIIRAEVDDVIQVRFKNLASRPYSLHAHGLSYEK SSEGKTYEDDSPEWFKEDNAVQPNSSYTYVWHATERSGPESPGSACRA WAYYSAVNPEKDIHSGLIGPLLICQKGILHKDSNMPMDMREFVLLFMTFDE KKSWYYEKKSRSSWRLTSSEMKKSHEFHAINGMIYSLPGLKMYEQEWVR LHLLNIGGSQDIHVVHFHGQTLLENGNKQHQLGVWPLLPGSFKTLEMKAS KPGWWLLNTEVGENQRAGMQTPFLIMDRDCRMPMGLSTGIISDSQIKASE FLGYWEPRLARLNNGGSYNAWSVEKLAAEFASKPWIQVDMQKEVIITGIQ TQGAKHYLKSCYTTEFYVAYSSNQINWQIFKGNSTRNVMYFNGNSDASTI KENQFDPPIVARYIRISPTRAYNRPTLRLELQGCEVNGCSTPLGMENGKIE NKQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNNKQWLEID LLKIKKITAIITQGCKSLSSEMYVKSYTIHYSEQGVEWKPYRLKSSMVDKIFE GNTNTKGHVKNFFNPPIISRFIRVIPKTWNQSIALRLELFGCDIY [SEQ ID NO:491]		
C3	718	0125730	B4DR57, P01024, V9HWA9	MGPTSGPSLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGD VPVTVTVHDFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRN KFVTVQATFGTQVVEKVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHK LLPVGRTVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDIPELVNMGQWKI RAYYENSPQQVFSTEFEVKEYVLPSFEVIVEPTEKFYYIYNEKGLEVTITAR	Síndrome hemolítico-urêmica atípica com anomalia C3	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				FLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEVVLSRKVLLDG VQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPIVTSPYQIHFTKTPK YFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSI NTHPSQKPLSITVRTKKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVL RTELRPGETLNVNFLLRMDRAHEAKIRYYTYLIMNKGRLLKAGRQVREPG QDLVVLPLSITTDFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCVGS LVVKSGQSEDRQPVPGQQMTLKIEGDHGARVVLVAVDKGVFVLNKKNKL TQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTAQRAELQ CPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQR RTRFISLGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIV SRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTWEILAVSMSDKK GICVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRV ELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVIVPLKTGLQEVEVKAAV YHHFISDGVRKSLKVVPEGIRMNKTVAVRTLDPERLGREGVQKEDIPPADL SDQVPDTESETRILLQGTPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIG MTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLAFRQPSSAFA AFVKRAPSTWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVF QEDAPVIHQEMIGGLRNNNEKDMALTAFVLISLQEAKDICEEQVNSLPGSIT KAGDFLEANYMNLQRSYTVAIAGYALAQMGRLKGPLLNKFLTTAKDKNRW EDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQRYYGGGYGST QATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLL RSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKP APETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQ LANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKVHQYFNVELI QPGAVKVYAYYNLEESCTRFYHPEKEDGKLNKLCRDELCRCAEENCFIQK SDDKVTLEERLDKACEPGVDYVYKTRLVKVQLSNDFDEYIMAIEQTIKSGS DEVQVGQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSYIIGK DTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN [SEQ ID NO:492]		
COL4A1	1282	0187498	A5PKV2, F5H5K0, P02462	MGPRLSVWLLLLPAALLLHEEHSRAAAKGGCAGSGCGKCDCHGVKGQK GERGLPGLQGVIGFPGMQGPEGPQGPPGQKGDTGEPGLPGTKGTRGPP GASGYPGNPGLPGIPGQDGPPGPPGIPGCNGTKGERGPLGPPGLPGFAG	Hematúria familial autossômica dominante - tortu-	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				NPGPPGLPGMKGDPGEILGHVPGMLLKGERGFPGIPGTPGPPGLPGLQG PVGPPGFTGPPGPPGPPGPPGEKGQMGLSFQGPKGDKGDQGVSGPPG VPGQAQVQEKGDFATKGEKGQKGEPGFQGMPGVGEKGEPGKPGPRGK PGKDGDKGEKGSPGFPGEPGYPGLIGRQGPQGEKGEAGPPGPPGIVIGT GPLGEKGERGYPGTPGPRGEPGPKGFPGLPGQPGPPGLPVPGQAGAPG FPGERGEKGDRGFPGTSLPGPSGRDGLPGPPGSPGPPGQPGYTNGIVE CQPGPPGDQGPPGIPGQPGFIGEIGEKGQKGESCLICDIDGYRGPPGPQG PPGEIGFPGQPGAKGDRGLPGRDGVAGVPGPQGTPGLIGQPGAKGEPG EFYFDLRLKGDKGDPGFPGQPGMPGRAGSPGRDGHPGLPGPKGSPGSV GLKGERGPPGGVGFPGSRGDTGPPGPPGYGPAGPIGDKGQAGFPGGPG SPGLPGPKGEPGKIVPLPGPPGAEGLPGSPGFPGPQGDRGFPGTPGRPG LPGEKGAVGQPGIGFPGPPGPKGVDGLPGDMGPPGTPGRPGFNGLPGN PGVQGQKGEPGVGLPGLKGLPGLPGIPGTPGEKGSIGVPGVPGEHGAIG PPGLQGIRGEPGPPGLPGSVGSPGVPGIGPPGARGPPGGQGPPGLSGPP GIKGEKGFPGFPGLDMPGPKGDKGAQGLPGITGQSGLPGLPGQQGAPGI PGFPGSKGEMGVMGTPGQPGSPGPVGAPGLPGEKGDHGFPGSSGPRG DPGLKGDKGDVGLPGKPGSMDKVDMGSMKGQKGDQGEKGQIGPIGEKG SRGDPGTPGVPGKDGQAGQPGQPGPKGDPGISGTPGAPGLPGPKGSVG GMGLPGTPGEKGVPGIPGPQGSPGLPGDKGAKGEKGQAGPPGIGIPGLR GEKGDQGIAGFPGSPGEKGEKGSIGIPGMPGSPGLKGSPGSVGYPGSPG LPGEKGDKGLPGLDGIPGVKGEAGLPGTPGPTGPAGQKGEPGSDGIPGS AGEKGEPGLPGRGFPGFPGAKGDKGSKGEVGFPGLAGSPGIPGSKGEQ GFMGPPGPQGQPGLPGSPGHATEGPKGDRGPQGQPGLPGLPGPMGPP GLPGIDGVKGDKGNPGWPGAPGVPGPKGDPGFQGMPGIGGSPGITGSK GDMGPPGVPGFQGPKGLPGLQGIKGDQGDQGVPGAKGLPGPPGPPGPY DIIKGEPGLPGPEGPPGLKGLQGLPGPKGQQGVTGLVGIPGPPGIPGFDG APGQKGEMGPAGPTGPRGFPGPPGPDGLPGSMGPPGTPSVDHGFLVTR HSQTIDDPQCPSGTKILYHGYSLLYVQGNERAHGQDLGTAGSCLRKFSTM PFLFCNINNVCNFASRNDYSYWLSTPEPMPMSMAPITGENIRPFISRCAVC EAPAMVMAVHSQTIQIPPCPSGWSSLWIGYSFVMHTSAGAEGSGQALAS PGSCLEEFRSAPFIECHGRGTCNYYANAYSFWLATIERSEMFKKPTPSTLK AGELRTHVSRCQVCMRRT [SEQ ID NO:493]	osidade arteriolar retiniana - contra-turas	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
CFH	3075	0000971	A0A024R962, P08603, A0A0D9SG88	MRLLAKIICLMLWAICVAEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKC RPGYRSLGNVIMVCRKGEWVALNPLRKCQKRPCGHPGDTPFGTFTLTGG NVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPVTAP ENGKIVSSAMEPDREYHFGQAVRFVCNSGYKIEGDEEMHCSDDGFWSKE KPKCVEISCKSPDVINGSPISQKIIYKENERFQYKCNMGYEYSERGDAVCT ESGWRPLPSCEEKSCDNPYIPNGDYSPLRIKHRTGDEITYQCRNGFYPAT RGNTAKCTSTGWIPAPRCTLKPCDYPDIKHGGLYHENMRRPYFPVAVGKY YSYYCDEHFETPSGSYWDHIHCTQDGWSPAVPCLRKCYFPYLENGYNQN YGRKFVQGKSIDVACHPGYALPKAQTTVTCMENGWSPTPRCIRVKTCSKS SIDIENGFISESQYTYALKEKAKYQCKLGYVTADGETSGSITCGKDGWSAQ PTCIKSCDIPVFMNARTKNDFTWFKLNDTLDYECHDGYESNTGSTTGSIVC GYNGWSDLPICYERECELPKIDVHLVPDRKKDQYKVGEVLKFSCKPGFTIV GPNSVQCYHFGLSPDLPICKEQVQSCGPPPELLNGNVKEKTKEEYGHSEV VEYYCNPRFLMKGPNKIQCVDGEWTTLPVCIVEESTCGDIPELEHGWAQL SSPPYYYGDSVEFNCSESFTMIGHRSITCIHGVWTQLPQCVAIDKLKKCKS SNLIILEEHLKNKKEFDHNSNIRYRCRGKEGWIHTVCINGRWDPEVNCSMA QIQLCPPPPQIPNSHNMTTTLNYRDGEKVSVLCQENYLIQEGEEITCKDGR WQSIPLCVEKIPCSQPPQIEHGTINSSRSSQESYAHGTKLSYTCEGGFRIS EENETTCYMGKWSSPPQCEGLPCKSPPEISHGVVAHMSDSYQYGEEVTY KCFEGFGIDGPAIAKCLGEKWSHPPSCIKTDCLSLPSFENAIPMGEKKDVY KAGEQVTYTCATYYKMDGASNVTCINSRWTGRPTCRDTSCVNPPTVQNA YIVSRQMSKYPSGERVRYQCRSPYEMFGDEEVMCLNGNWTEPPQCKDS TGKCGPPPPIDNGDITSFPLSVYAPASSVEYQCQNLYQLEGNKRITCRNG QWSEPPKCLHPCVISREIMENYNIALRWTAKQKLYSRTGESVEFVCKRGY RLSSRSHTLRTTCWDGKLEYPTCAKR [SEQ ID NO:494]	Síndrome hemolítica urêmica atípica	
SLC12A2	6558	0064651	P55011, Q53ZR1, B7ZM24	MEPRPTAPSSGAPGLAGVGETPSAAALAAARVELPGTAVPSVPEDAAPAS RDGGGVRDEGPAAAGDGLGRPLGPTPSQSRFQVDLVSENAGRAAAAAA AAAAAAAAAGAGAGAKQTPADGEASGESEPAKGSEEAKGRFRVNFVDPA ASSSAEDSLSDAAGVGVDGPNVSFQNGGDTVLSEGSSLHSGGGGGSGH HQHYYYDTHTNTYYLRTFGHNTMDAVPRIDHYRHTAAQLGEKLLRPSLAE LHDELEKEPFEDGFANGEESTPTRDAVVTYTAESKGVVKFGWIKGVLVRC	Síndrome de Bartter tipo I (neonatal)	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				MLNIWGVMLFIRLSWIVGQAGIGLSVLVIMMATVVTTITGLSTSAIATNGFVR GGGAYYLISRSLGPEFGGAIGLIFAFANAVAVAMYVVGFAETVVELLKEHSI LMIDEINDIRIIGAITVVILLGISVAGMEWEAKAQIVLLVILLLAIGDFVIGTFIPL ESKKPKGFFGYKSEIFNENFGPDFREEETFFSVFAIFFPAATGILAGANISG DLADPQSAIPKGTLLAILITTLVYVGIAVSVGSCVVRDATGNVNDTIVTELTN CTSAACKLNFDFSSCESSPCSYGLMNNFQVMSMVSGFTPLISAGIFSATLS SALASLVSAPKIFQALCKDNIYPAFQMFAKGYGKNNEPLRGYILTFLIALGFI LIAELNVIAPIISNFFLASYALINFSVFHASLAKSPGWRPAFKYYNMWISLLG AILCCIVMFVINWWAALLTYVIVLGLYIYVTYKKPDVNWGSSTQALTYLNAL QHSIRLSGVEDHVKNFRPQCLVMTGAPNSRPALLHLVHDFTKNVGLMICG HVHMGPRRQAMKEMSIDQAKYQRWLIKNKMKAFYAPVHADDLREGAQYL MQAAGLGRMKPNTLVLGFKKDWLQADMRDVDMYINLFHDAFDIQYGVVVI RLKEGLDISHLQGQEELLSSQEKSPGTKDVVVSVEYSKKSDLDTSKPLSEK PITHKVEEEDGKTATQPLLKKESKGPIVPLNVADQKLLEASTQFQKKQGKN TIDVWWLFDDGGLTLLIPYLLTTKKKWKDCKIRVFIGGKINRIDHDRRAMAT LLSKFRIDFSDIMVLGDINTKPKKENIIAFEEIIEPYRLHEDDKEQDIADKMKE DEPWRITDNELELYKTKTYRQIRLNELLKEHSSTANIIVMSLPVARKGAVSS ALYMAWLEALSKDLPPILLVRGNHQSVLTFYS [SEQ ID NO: 495]		
GK	2710	0198814	B4DH54, P32189	MAASKKAVLGPLVGAVDQGTSSTRFLVFNSKTAELLSHHQVEIKQEFPRE GWVEQDPKEILHSVYECIEKTCEKLGQLNIDISNIKAIGVSNQRETTVVWDK ITGEPLYNAVVWLDLRTQSTVESLSKRIPGNNNFVKSKTGLPLSTYFSAVK LRWLLDNVRKVQKAVEEKRALFGTIDSWLIWSLTGGVNGGVHCTDVTNAS RTMLFNIHSLEWDKQLCEFFGIPMEILPNVRSSSEIYGLMKISHSVKAGALE GVPISGCLGDQSAALVGQMCFQIGQAKNTYGTGCFLLCNTGHKCVFSDH GLLTTVAYKLGRDKPVYYALEGSVAIAGAVIRWLRDNLGIIKTSEEIEKLAKE VGTSYGCYFVPAFSGLYAPYWEPSARGIICGLTQFTNKCHIAFAALEAVCF QTREILDAMNRDCGIPLSHLQVDGGMTSNKILMQLQADILYIPVVKPSMPE TTALGAAMAAGAAEGVGVWSLEPEDLSAVTMERFEPQINAEESEIRYSTW KKAVMKSMGWVTTQSPESGDPSIFCSLPLGFFIVSSMVMLIGARYISGIP [SEQ ID NO: 496]	Deficiência de glicerol quinase	
SFTP	6440	0168484	A0A0A0MTC9,	MDVGSKEVLMESPPDYSAAPRGRFGIPCCPVHLKRLLIVVVVVVLIVVVIVG	Desconforto respi-	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
C			P11686, A0A0S2Z4Q0, E5RI64	ALLMGLHMSQKHTEMVLEMSIGAPEAQQRLALSEHLVTTATFSIGSTGLVV YDYQQLLIAYKPAPGTCCYIMKIAPESIPSLEALNRKVHNFQMECSLQAKPA VPTSKLGQAEGRDAGSAPSGGDPAFLGMAVNTLCGEVPLYYI [SEQ ID NO:497]	ratório crônico com deficiência de metabolismo de tensoativo	
CRTA P	10491	0170275	O75718	MEPGRRGAAALLALLCVACALRAGRAQYERYSFRSFPRDELMPLESAYR HALDKYSGEHWAESVGYLEISLRLHRLLRDSEAFCHRNCSAAPQPEPAAG LASYPELRLFGGLLRRAHCLKRCKQGLPAFRQSQPSREVLADFQRREPYK FLQFAYFKANNLPKAIAAAHTFLLKHPDDEMMKRNMAYYKSLPGAEDYIKD LETKSYESLFIRAVRAYNGENWRTSITDMELALPDFFKAFYECLAACEGSR EIKDFKDFYLSIADHYVEVLECKIQCEENLTPVIGGYPVEKFVATMYHYLQF AYYKLNDLKNAAPCAVSYLLFDQNDKVMQQNLVYYQYHRDTWGLSDEHF QPRPEAVQFFNVTTLQKELYDFAKENIMDDDEGEVVEYVDDLLEETS [SEQ ID NO: 498]	Osteogenesis Imperfecta VII	
P3H1	64175	0117385	Q32P28	MAVRALKLLTTLLAVVAAASQAEVESEAGWGMVTPDLLFAEGTAAYARGD WPGVVLSMERALRSRAALRALRLRCRTQCAADFPWELDPDWSPSPAQA SGAAALRDLSFFGGLLRRAACLRRCLGPPAAHSLSEEMELEFRKRSPYNY LQVAYFKINKLEKAVAAAHTFFVGNPEHMEMQQNLDYYQTMSGVKEADF KDLETQPHMQEFRLGVRLYSEEQPQEAVPHLEAALQEYFVAYEECRALCE GPYDYDGYNYLEYNADLFQAITDHYIQVLNCKQNCVTELASHPSREKPFE DFLPSHYNYLQFAYYNIGNYTQAVECAKTYLLFFPNDEVMNQNLAYYAAM LGEEHTRSIGPRESAKEYRQRSLLEKELLFFAYDVFGIPFVDPDSWTPEEVI PKRLQEKQKSERETAVRISQEIGNLMKEIETLVEEKTKESLDVSRLTREGG PLLYEGISLTMNSKLLNGSQRVVMDGVISDHECQELQRLTNVAATSGDGY RGQTSPHTPNEKFYGVTVFKALKLGQEGKVPLQSAHLYYNVTEKVRRIME SYFRLDTPLYFSYSHLVCRTAIEEVQAERKDDSHPVHVDNCILNAETLVCV KEPPAYTFRDYSAILYLNGDFDGGNFYFTELDAKTVTAEVQPQCGRAVGF SSGTENPHGVKAVTRGQRCAIALWFTLDPRHSERDRVQADDLVKMLFSP EEMDLSQEQPLDAQQGPPEPAQESLSGSESKPKDEL [SEQ ID NO: 499]	Osteogênesis Imperfecta VIII	
COL7 A1	1294	0114270	Q02388, Q59F16	MTLRLLVAALCAGILAEAPRVRAQHRERVTCTRLYAADIVFLLDGSSSIGRS NFREVRSFLEGLVLPFSGAASAQGVRFATVQYSDDPRTEFGLDALGSGG	Epidermólise bolhosa distrófica	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				DVIRAIRELSYKGGNTRTGAAILHVADHVFLPQLARPGVPKVCILITDGKSQ DLVDTAAQRLKGQGVKLFAVGIKNADPEELKRVASQPTSDFFFFVNDFSIL RTLLPLVSRRVCTTAGGVPVTRPPDDSTSAPRDLVLSEPSSQSLRVQWTA ASGPVTGYKVQYTPLTGLGQPLPSERQEVNVPAGETSVRLRGLRPLTEYQ VTVIALYANSIGEAVSGTARTTALEGPELTIQNTTAHSLLVAWRSVPGATGY RVTWRVLSGGPTQQQELGPGQGSVLLRDLEPGTDYEVTVSTLFGRSVGP ATSLMARTDASVEQTLRPVILGPTSILLSWNLVPEARGYRLEWRRETGLEP PQKVVLPSDVTRYQLDGLQPGTEYRLTLYTLLEGHEVATPATVVPTGPELP VSPVTDLQATELPGQRVRVSWSPVPGATQYRIIVRSTQGVERTLVLPGSQ TAFDLDDVQAGLSYTVRVSARVGPREGSASVLTVRREPETPLAVPGLRVV VSDATRVRVAWGPVPGASGFRISWSTGSGPESSQTLPPDSTATDITGLQP GTTYQVAVSVLRGREEGPAAVIVARTDPLGPVRTVHVTQASSSSVTITWT RVPGATGYRVSWHSAHGPEKSQLVSGEATVAELDGLEPDTEYTVHVRAH VAGVDGPPASVVVRTAPEPVGRVSRLQILNASSDVLRITWVGVTGATAYR LAWGRSEGGPMRHQILPGNTDSAEIRGLEGGVSYSVRVTALVGDREGTP VSIVVTTPPEAPPALGTLHVVQRGEHSLRLRWEPVPRAQGFLLHWQPEG GQEQSRVLGPELSSYHLDGLEPATQYRVRLSVLGPAGEGPSAEVTARTE SPRVPSIELRVVDTSIDSVTLAWTPVSRASSYILSWRPLRGPGQEVPGSPQ TLPGISSSQRVTGLEPGVSYIFSLTPVLDGVRGPEASVTQTPVCPRGLADV VFLPHATQDNAHRAEATRRVLERLVLALGPLGPQAVQVGLLSYSHRPSPL FPLNGSHDLGIILQRIRDMPYMDPSGNNLGTAVVTAHRYMLAPDAPGRRQ HVPGVMVLLVDEPLRGDIFSPIREAQASGLNVVMLGMAGADPEQLRRLAP GMDSVQTFFAVDDGPSLDQAVSGLATALCQASFTTQPRPEPCPVYCPKG QKGEPGEMGLRGQVGPPGDPGLPGRTGAPGPQGPPGSATAKGERGFP GADGRPGSPGRAGNPGTPGAPGLKGSPGLPGPRGDPGERGPRGPKGE PGAPGQVIGGEGPGLPGRKGDPGPSGPPGPRGPLGDPGPRGPPGLPGT AMKGDKGDRGERGPPGPGEGGIAPGEPGLPGLPGSPGPQGPVGPPGKK GEKGDSEDGAPGLPGQPGSPGEQGPRGPPGAIGPKGDRGFPGPLGEAG EKGERGPPGPAGSRGLPGVAGRPGAKGPEGPPGPTGRQGEKGEPGRP GDPAVVGPAVAGPKGEKGDVGPAGPRGATGVQGERGPPGLVLPGDPGP KGDPGDRGPIGLTGRAGPPGDSGPPGEKGDPGRPGPPGPVGPRGRDGE VGEKGDEGPPGDPGLPGKAGERGLRGAPGVRGPVGEKGDQGDPGEDG	autossômica recessiva	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				RNGSPGSSGPKGDRGEPGPPGPPGRLVDTGPGAREKGEPGDRGQEGP RGPKGDPGLPGAPGERGIEGFRGPPGPQGDPGVRGPAGEKGDRGPPGL DGRSGLDGKPGAAGPSGPNGAAGKAGDPGRDGLPGLRGEQGLPGPSG PPGLPGKPGEDGKPGLNGKNGEPGDPGEDGRKGEKGDSGASGREGRD GPKGERGAPGILGPQGPPGLPGPVGPPGQGFPGVPGGTGPKGDRGETG SKGEQGLPGERGLRGEPGSVPNVDRLLETAGIKASALREIVETWDESSGS FLPVPERRRGPKGDSGEQGPPGKEGPIGFPGERGLKGDRGDPGPQGPP GLALGERGPPGPSGLAGEPGKPGIPGLPGRAGGVGEAGRPGERGERGE KGERGEQGRDGPPGLPGTPGPPGPPGPKVSVDEPGPGLSGEQGPPGLK GAKGEPGSNGDQGPKGDRGVPGIKGDRGEPGPRGQDGNPGLPGERGM AGPEGKPGLQGPRGPPGPVGGHGDPGPPGAPGLAGPAGPQGPSGLKG EPGETGPPGRGLTGPTGAVGLPGPPGPSGLVGPQGSPGLPGQVGETGK PGAPGRDGASGKDGDRGSPGVPGSPGLPGPVGPKGEPGPTGAPGQAV VGLPGAKGEKGAPGGLAGDLVGEPGAKGDRGLPGPRGEKGEAGRAGEP GDPGEDGQKGAPGPKGFKGDPGVGVPGSPGPPGPPGVKGDLGLPGLP GAPGVVGFPGQTGPRGEMGQPGPSGERGLAGPPGREGIPGPLGPPGPP GSVGPPGASGLKGDKGDPGVGLPGPRGERGEPGIRGEDGRPGQEGPR GLTGPPGSRGERGEKGDVGSAGLKGDKGDSAVILGPPGPRGAKGDMGE RGPRGLDGDKGPRGDNGDPGDKGSKGEPGDKGSAGLPGLRGLLGPQG QPGAAGIPGDPGSPGKDGVPGIRGEKGDVGFMGPRGLKGERGVKGACG LDGEKGDKGEAGPPGRPGLAGHKGEMGEPGVPGQSGAPGKEGLIGPKG DRGFDGQPGPKGDQGEKGERGTPGIGGFPGPSGNDGSAGPPGPPGSV GPRGPEGLQGQKGERGPPGERVVGAPGVPGAPGERGEQGRPGPAGPR GEKGEAALTEDDIRGFVRQEMSQHCACQGQFIASGSRPLPSYAADTAGS QLHAVPVLRVSHAEEEERVPPEDDEYSEYSEYSVEEYQDPEAPWDSDDP CSLPLDEGSCTAYTLRWYHRAVTGSTEACHPFVYGGCGGNANRFGTREA CERRCPPRVVQSQGTGTAQD [SEQ ID NO: 500]		
PKLR	5313	0143627	P30613	MSIQENISSLQLRSWVSKSQRDLAKSILIGAPGGPAGYLRRASVAQLTQEL GTAFFQQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIIATIGPASRSVER LKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFAGSPLSYRPVAIAL DTKGPEIRTGILQGGPESEVELVKGSQVLVTVDPAFRTRGNANTVWVDYP NIVRVVPVGGRIYIDDGLISLVVQKIGPEGLVTQVENGGVLGSRKGVNLPG	Deficiência de piruvato quinase	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				AQVDLPGLSEQDVRDLRFGVEHGVDIVFASFVRKASDVAAVRAALGPEGH GIKIISKIENHEGVKRFDEILEVSDGIMVARGDLGIEIPAEKVFLAQKMMIGR CNLAGKPVVCATQMLESMITKPRPTRAETSDVANAVLDGADCIMLSGETA KGNFPVEAVKMQHAIAREAEAAVYHRQLFEELRRAAPLSRDPTEVTAIGAV EAAFKCCAAAIIVLTTTGRSAQLLSRYRPRAAVIAVTRSAQAARQVHLCRG VFPLLYREPPEAIWADDVDRRVQFGIESGKLRGFLRVGDLVIVVTGWRPG SGYTNIMRVLSIS [SEQ ID NO: 501]		
TAL-DO1	6888	0177156	A0A140VK56, P37837	MSSSPVKRQRMESALDQLKQFTTVVADTGDFHAIDEYKPQDATTNPSLILA AAQMPAYQELVEEAIAYGRKLGGSQEDQIKNAIDKLFVLFGAEILKKIPGRV STEVDARLSFDKDAMVARARRLIELYKEAGISKDRILIKLSSTWEGIQAGKE LEEQHGIHCNMTLLFSFAQAVACAEAGVTLISPFVGRILDWHVANTDKKSY EPLEDPGVKSVTKIYNYYKKFSYKTIVMGASFRNTGEIKALAGCDFLTISPK LLGELLQDNAKLVPVLSAKAAQASDLEKIHLDEKSFRWLHNEDQMAVEKL SDGIRKFAADAVKLERMLTERMFNAENGK [SEQ ID NO: 502]	Deficiência de transaldolase	
TF	7018	0091513	A0PJA6, P02787, Q06AH7	MRLAVGALLVCAVLGLCLAVPDKTVRWCAVSEHEATKCQSFRDHMKSVIP SDGPSVACVKKASYLDCIRAIAANEADAVTLDAGLVYDAYLAPNNLKPVVA EFYGSKEDPQTFYYAVAVVKKDSGFQMNQLRGKKSCHTGLGRSAGWNIP IGLLYCDLPEPRKPLEKAVANFFSGSCAPCADGTDFPQLCQLCPGCGCST LNQYFGYSGAFKCLKDGAGDVAFVKHSTIFENLANKADRDQYELLCLDNT RKPVDEYKDCHLAQVPSHTVVARSMGGKEDLIWELLNQAQEHFGKDKSK EFQLFSSPHGKDLLFKDSAHGFLKVPPRMDAKMYLGYEYVTAIRNLREGT CPEAPTDECKPVKWCALSHHERLKCDEWSVNSVGKIECVSAETTEDCIAK IMNGEADAMSLDGGFVYIAGKCGLVPVLAENYNKSDNCEDTPEAGYFAIA VVKKSASDLTWDNLKGKKSCHTAVGRTAGWNIPMGLLYNKINHCRFDEFF SEGCAPGSKKDSSLCKLCMGSGLNLCEPNNKEGYYGYTGAFRCLVEKGD VAFVKHQTVPQNTGGKNPDPWAKNLNEKDYELLCLDGTRKPVEEYANCH LARAPNHAVVTRKDKEACVHKILRQQQHLFGSNVTDCSGNFCLFRSETKD LLFRDDTVCLAKLHDRNTYEKYLGEEYVKAVGNLRKCSTSSLLEACTFRRP [SEQ ID NO: 503]	Atransferrinemia (hipotransferrine-mia familiar)	
EPCA	4072	0119888	P16422	MAPPQVLAFGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQCQCTS	Displasia epitelial	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
M				VGAQNTVICSKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNDGLYDPDC DESGLFKAKQCNGTSMCWCVNTAGVRRTDKDTEITCSERVRTYWIIIELKH KAREKPYDSKSLRTALQKEITTRYQLDPKFITSILYENNVITIDLVQNSSQKT QNDVDIADVAYYFEKDVKGESLFHSKKMDLTVNGEQLDLDPGQTLIYYVD EKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMAKYEKAEIKEM GEMHRELNA [SEQ ID NO: 504]	intestinal	
VHL	7428	0134086	A0A024R2F2, P40337, A0A0S2Z4K1	MPRRAENWDEAEVGAEEAGVEEYGPEEDGGEESGAEESGPEESGPEEL GAEEEMEAGRPRPVLRSVNSREPSQVIFCNRSPRVVLPVWLNFDGEPQP YPTLPPGTGRRIHSYRGHLWLFRDAGTHDGLLVNQTELFVPSLNVDGQPI FANITLPVYTLKERCLQVVRSLVKPENYRRLDIVRSLYEDLEDHPNVQKDL ERLTQERIAHQRMGD [SEQ ID NO: 505]	Eritrocitose familiar tipo 2; doença de von Hippel Lindau	
GC	2638	0145321	P02774	MKRVLVLLLAVAFGHALERGRDYEKNKVCKEFSHLGKEDFTSLSLVLYSR KFPSGTFEQVSQLVKEVVSLTEACCAEGADPDCYDTRTSALSAKSCESNS PFPVHPGTAECCTKEGLERKLCMAALKHQPQEFPTYVEPTNDEICEAFRK DPKEYANQFMWEYSTNYGQAPLSLLVSYTKSYLSMVGSCCTSASPTVCF LKERLQLKHLSLLTTLSNRVCSQYAAYGEKKSRLSNLIKLAQKVPTADLED VLPLAEDITNILSKCCESASEDCMAKELPEHTVKLCDNLSTKNSKFEDCCQ EKTAMDVFVCTYFMPAAQLPELPDVELPTNKDVCDPGNTKVMDKYTFELS RRTHLPEVFLSKVLEPTLKSLGECCDVEDSTTCFNAKGPLLKKELSSFIDK GQELCADYSENTFTEYKKKLAERLKAKLPDATPTELAKLVNKHSDFASNC CSINSPPLYCDSEIDAELKNIL [SEQ ID NO: 506]	Deficiência de vitamina D	
SER-PINA1	5265	0197249,0277377	E9KL23, P01009	MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHPTFNKI TPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILE GLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKF LEDVKKLYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTV FALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQH CKKLSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFLENEDRRS ASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLSKAVHK AVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGK VVNPTQK [SEQ ID NO: 507]	Deficiência de alfa-1 antitripsina	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
ABCC6	368	0091262,0275331	O95255	MAAPAEPCAGQGVWNQTEPEPAATSLLSLCFLRTAGVWVPPMYLWVLG PIYLLFIHHHGRGYLRMSPLFKAKMVLGFALIVLCTSSVAVALWKIQQGTPE APEFLIHPTVWLTTMSFAVFLIHTERKKGVQSSGVLFGYWLLCFVLPATNA AQQASGAGFQSDPVRHLSTYLCLSLVVAQFVLSCLADQPPFFPEDPQQS NPCPETGAAFPSKATFWWVSGLVWRGYRRPLRPKDLWSLGRENSSEEL VSRLEKEWMRNRSAARRHNKAIAFKRKGGSGMKAPETEPFLRQEGSQW RPLLKAIWQVFHSTFLLGTLSLIISDVFRFTVPKLLSLFLEFIGDPKPPAWKG YLLAVLMFLSACLQTLFEQQNMYRLKVLQMRLRSAITGLVYRKVLALSSGS RKASAVGDVVNLVSVDVQRLTESVLYLNGLWLPLVWIVVCFVYLWQLLGP SALTAIAVFLSLLPLNFFISKKRNHHQEEQMRQKDSRARLTSSILRNSKTIKF HGWEGAFLDRVLGIRGQELGALRTSGLLFSVSLVSFQVSTFLVALVVFAVH TLVAENAMNAEKAFVTLTVLNILNKAQAFLPFSIHSLVQARVSFDRLVTFLC LEEVDPGVVDSSSSGSAAGKDCITIHSATFAWSQESPPCLHRINLTVPQGC LLAVVGPVGAGKSSLLSALLGELSKVEGFVSIEGAVAYVPQEAWVQNTSV VENVCFGQELDPPWLERVLEACALQPDVDSFPEGIHTSIGEQGMNLSGG QKQRLSLARAVYRKAAVYLLDDPLAALDAHVGQHVFNQVIGPGGLLQGTT RILVTHALHILPQADWIIVLANGAIAEMGSYQELLQRKGALMCLLDQARQPG DRGEGETEPGTSTKDPRGTSAGRRPELRRERSIKSVPEKDRTTSEAQTEV PLDDPDRAGWPAGKDSIQYGRVKATVHLAYLRAVGTPLCLYALFLFLCQQ VASFCRGYWLSLWADDPAVGGQQTQAALRGGIFGLLGCLQAIGLFASMA AVLLGGARASRLLFQRLLWDVVRSPISFFERTPIGHLLNRFSKETDTVDVDI PDKLRSLLMYAFGLLEVSLVVAVATPLATVAILPLFLLYAGFQSLYVVSSCQ LRRLESASYSSVCSHMAETFQGSTVVRAFRTQAPFVAQNNARVDESQRIS FPRLVADRWLAANVELLGNGLVFAAATCAVLSKAHLSAGLVGFSVSAALQ VTQTLQWVVRNWTDLENSIVSVERMQDYAWTPKEAPWRLPTCAAQPPW PQGGQIEFRDFGLRYRPELPLAVQGVSFKIHAGEKVGIVGRTGAGKSSLA SGLLRLQEAAEGGIWIDGVPIAHVGLHTLRSRISIIPQDPILFPGSLRMNLDL LQEHSDEAIWAALETVQLKALVASLPGQLQYKCADRGEDLSVGQKQLLCL ARALLRKTQILILDEATAAVDPGTELQMQAMLGSWFAQCTVLLIAHRLRSV MDCARVLVMDKGQVAESGSPAQLLAQKGLFYRLAQESGLV [SEQ ID NO: 508]	Pseudoxantoma elástico	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
F8	2157	0185010	P00451	MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFP PRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYD TVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPG GSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCRE GSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPK MHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRN HRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPE EPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMA YTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDV RPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYS SFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSW YLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAY WYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGL WILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIE PRSFSQNSRHPSTRQKQFNATTIPENDIEKTDPWFAHRTPMPKIQNVSSS DLLMLLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAIDSNNSLSEMTHFR PQLHHSGDMVFTPESGLQLRLNEKLGTTAATELKKLDFKVSSTSNNLISTIP SDNLAAGTDNTSSLGPPSMPVHYDSQLDTTLFGKKSSPLTESGGPLSLSE ENNDSKLLESGLMNSQESSWGKNVSSTESGRLFKGKRAHGPALLTKDNA LFKVSISLLKTNKTSNNSATNRKTHIDGPSLLIENSPSVWQNILESDTEFKKV TPLIHDRMLMDKNATALRLNHMSNKTTSSKNMEMVQQKKEGPIPPDAQN PDMSFFKMLFLPESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSVEG QNFLSEKNKVVVGKGEFTKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQ EKKIQEEIEKKETLIQENVVLPQIHTVTGTKNFMKNLFLLSTRQNVEGSYDG AYAPVLQDFRSLNDSTNRTKKHTAHFSKKGEEENLEGLGNQTKQIVEKYA CTTRISPNTSQQNFVTQRSKRALKQFRLPLEETELEKRIIVDDTSTQWSKN MKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHSIPQANRSPLPIAKVSS FPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQGAKKNNLSLAI LTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLPKTSGKVELLPK VHIYQKDLFPTETSNGSPGHLDLVEGSLLQGTEGAIKWNEANRPGKVPFL RVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPEKTAFKKKDTI	Hemofilia A	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				LSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQR EITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFI AAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYR GELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGA EPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHS GLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCR APCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSN ENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLI GEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLA RLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIM YSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHY SIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPS KARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTS MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR YLRIHPQSWVHQIALRMEVLGCEAQDLY [SEQ ID NO: 509]		
F9	2158	0101981	P00740	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKL EEFVQGNLERECMEEKCSFEEAREVFENTERTTEFWKQYVDGDQCESNP CLNGGSCKDDINSYECWCPFGFEGKNCELDVTCNIKNGRCEQFCKNSAD NKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQTSKLTRAETVFPDV DYVNSTEAETILDNITQSTQSFNDFTRVVGGEDAKPGQFPWQVVLNGKVD AFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRIIPH HNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSGYVSG WGRVFHKGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGR DSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWI KEKTKLT [SEQ ID NO:510]	Hemofilia B	
ApoB	338	0084674	P04114	MDPPRPALLALLALPALLLLLLAGARAEEEMLENVSLVCPKDATRFKHLRK YTYNYEAESSSGVPGTADSRSATRINCKVELEVPQLCSFILKTSQCTLKEV YGFNPEGKALLKKTKNSEEFAAAMSRYELKLAIPEGKQVFLYPEKDEPTYIL NIKRGIISALLVPPETEEAKQVLFLDTVYGNCSTHFTVKTRKGNVATEISTER DLGQCDRFKPIRTGISPLALIKGMTRPLSTLISSSQSCQYTLDAKRKHVAEA ICKEQHLFLPFSYKNKYGMVAQVTQTLKLEDTPKINSRFFGEGTKKMGLAF	Hipercolesterole-mia familiar	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ESTKSTSPPKQAEAVLKTLQELKKLTISEQNIQRANLFNKLVTELRGLSDEA VTSLLPQLIEVSSPITLQALVQCGQPQCSTHILQWLKRVHANPLLIDVVTYL VALIPEPSAQQLREIFNMARDQRSRATLYALSHAVNNYHKTNPTGTQELLD IANYLMEQIQDDCTGDEDYTYLILRVIGNMGQTMEQLTPELKSSILKCVQST KPSLMIQKAAIQALRKMEPKDKDQEVLLQTFLDDASPGDKRLAAYLMLMR SPSQAINKIVQILPWEQNEQVKNFVASHIANILNSEELDIQDLKKLVKEALKE SQLPTVMDFRKFSRNYQLYKSVSLPSLDPASAKIEGNLIFDPNNYLPKESM LKTTLTAFGFASADLIEIGLEGKGFEPTLEALFGKQGFFPDSVNKALYWVN GQVPDGVSKVLVDHFGYTKDDKHEQDMVNGIMLSVEKLIKDLKSKEVPEA RAYLRILGEELGFASLHDLQLLGKLLLMGARTLQGIPQMIGEVIRKGSKNDF FLHYIFMENAFELPTGAGLQLQISSSGVIAPGAKAGVKLEVANMQAELVAK PSVSVEFVTNMGIIIPDFARSGVQMNTNFFHESGLEAHVALKAGKLKFIIPS PKRPVKLLSGGNTLHLVSTTKTEVIPPLIENRQSWSVCKQVFPGLNYCTSG AYSNASSTDSASYYPLTGDTRLELELRPTGEIEQYSVSATYELQREDRALV DTLKFVTQAEGAKQTEATMTFKYNRQSMTLSSEVQIPDFDVDLGTILRVND ESTEGKTSYRLTLDIQNKKITEVALMGHLSCDTKEERKIKGVISIPRLQAEAR SEILAHWSPAKLLLQMDSSATAYGSTVSKRVAWHYDEEKIEFEWNTGTNV DTKKMTSNFPVDLSDYPKSLHMYANRLLDHRVPQTDMTFRHVGSKLIVAM SSWLQKASGSLPYTQTLQDHLNSLKEFNLQNMGLPDFHIPENLFLKSDGR VKYTLNKNSLKIEIPLPFGGKSSRDLKMLETVRTPALHFKSVGFHLPSREFQ VPTFTIPKLYQLQVPLLGVLDLSTNVYSNLYNWSASYSGGNTSTDHFSLRA RYHMKADSVVDLLSYNVQGSGETTYDHKNTFTLSYDGSLRHKFLDSNIKF SHVEKLGNNPVSKGLLIFDASSSWGPQMSASVHLDSKKKQHLFVKEVKID GQFRVSSFYAKGTYGLSCQRDPNTGRLNGESNLRFNSSYLQGTNQITGR YEDGTLSLTSTSDLQSGIIKNTASLKYENYELTLKSDTNGKYKNFATSNKM DMTFSKQNALLRSEYQADYESLRFFSLLSGSLNSHGLELNADILGTDKINS GAHKATLRIGQDGISTSATTNLKCSLLVLENELNAELGLSGASMKLTTNGR FREHNAKFSLDGKAALTELSLGSAYQAMILGVDSKNIFNFKVSQEGLKLSN DMMGSYAEMKFDHTNSLNIAGLSLDFSSKLDNIYSSDKFYKQTVNLQLQP YSLVTTLNSDLKYNALDLTNNGKLRLEPLKLHVAGNLKGAYQNNEIKHIYAI SSAALSASYKADTVAKVQGVEFSHRLNTDIAGLASAIDMSTNYNSDSLHFS NVFRSVMAPFTMTIDAHTNGNGKLALWGEHTGQLYSKFLLKAEPLAFTFS		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				HDYKGSTSHHLVSRKSISAALEHKVSALLTPAEQTGTWKLKTQFNNNEYS QDLDAYNTKDKIGVELTGRTLADLTLLDSPIKVPLLLSEPINIIDALEMRDAV EKPQEFTIVAFVKYDKNQDVHSINLPFFETLQEYFERNRQTIIVVLENVQRN LKHINIDQFVRKYRAALGKLPQQANDYLNSFNWERQVSHAKEKLTALTKKY RITENDIQIALDDAKINFNEKLSQLQTYMIQFDQYIKDSYDLHDLKIAIANIIDEI IEKLKSLDEHYHIRVNLVKTIHDLHLFIENIDFNKSGSSTASWIQNVDTKYQI RIQIQEKLQQLKRHIQNIDIQHLAGKLKQHIEAIDVRVLLDQLGTTISFERINDI LEHVKHFVINLIGDFEVAEKINAFRAKVHELIERYEVDQQIQVLMDKLVELAH QYKLKETIQKLSNVLQQVKIKDYFEKLVGFIDDAVKKLNELSFKTFIEDVNKF LDMLIKKLKSFDYHQFVDETNDKIREVTQRLNGEIQALELPQKAEALKLFLE ETKATVAVYLESLQDTKITLIINWLQEALSSASLAHMKAKFRETLEDTRDRM YQMDIQQELQRYLSLVGQVYSTLVTYISDWWTLAAKNLTDFAEQYSIQDW AKRMKALVEQGFTVPEIKTILGTMPAFEVSLQALQKATFQTPDFIVPLTDLRI PSVQINFKDLKNIKIPSRFSTPEFTILNTFHIPSFTIDFVEMKVKIIRTIDQMLN SELQWPVPDIYLRDLKVEDIPLARITLPDFRLPEIAIPEFIIPTLNLNDFQVPDL HIPEFQLPHISHTIEVPTFGKLYSILKIQSPLFTLDANADIGNGTTSANEAGIA ASITAKGESKLEVLNFDFQANAQLSNPKINPLALKESVKFSSKYLRTEHGSE MLFFGNAIEGKSNTVASLHTEKNTLELSNGVIVKINNQLTLDSNTKYFHKLNI PKLDFSSQADLRNEIKTLLKAGHIAWTSSGKGSWKWACPRFSDEGTHES QISFTIEGPLTSFGLSNKINSKHLRVNQNLVYESGSLNFSKLEIQSQVDSQH VGHSVLTAKGMALFGEGKAEFTGRHDAHLNGKVIGTLKNSLFFSAQPFEIT ASTNNEGNLKVRFPLRLTGKIDFLNNYALFLSPSAQQASWQVSARFNQYK YNQNFSAGNNENIMEAHVGINGEANLDFLNIPLTIPEMRLPYTIITTPPLKDF SLWEKTGLKEFLKTTKQSFDLSVKAQYKKNKHRHSITNPLAVLCEFISQSIK SFDRHFEKNRNNALDFVTKSYNETKIKFDKYKAEKSHDELPRTFQIPGYTV PVVNVEVSPFTIEMSAFGYVFPKAVSMPSFSILGSDVRVPSYTLILPSLELP VLHVPRNLKLSLPDFKELCTISHIFIPAMGNITYDFSFKSSVITLNTNAELFNQ SDIVAHLLSSSSSVIDALQYKLEGTTRLTRKRGLKLATALSLSNKFVEGSHN STVSLTTKNMEVSVATTTKAQIPILRMNFKQELNGNTKSKPTVSSSMEFKY DFNSSMLYSTAKGAVDHKLSLESLTSYFSIESSTKGDVKGSVLSREYSGTI ASEANTYLNSKSTRSSVKLQGTSKIDDIWNLEVKENFAGEATLQRIYSLWE HSTKNHLQLEGLFFTNGEHTSKATLELSPWQMSALV		

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				QVHASQPSSFHDFPDLGQEVALNANTKNQKIRWKNEVRIHSGSFQSQVEL SNDQEKAHLDIAGSLEGHLRFLKNIILPVYDKSLWDFLKLDVTTSIGRRQHL RVSTAFVYTKNPNGYSFSIPVKVLADKFIIPGLKLNDLNSVLVMPTFHVPFT DLQVPSCKLDFREIQIYKKLRTSSFALNLPTLPEVKFPEVDVLTKYSQPEDS LIPFFEITVPESQLTVSQFTLPKSVSDGIAALDLNAVANKIADFELPTIIVPEQ TIEIPSIKFSVPAGIVIPSFQALTARFEVDSPVYNATWSASLKNKADYVETVL DSTCSSTVQFLEYELNVLGTHKIEDGTLASKTKGTFAHRDFSAEYEEDGKY EGLQEWEGKAHLNIKSPAFTDLHLRYQKDKKGISTSAASPAVGTVGMDMD EDDDFSKWNFYYSPQSSPDKKLTIFKTELRVRESDEETQIKVNWEEEAAS GLLTSLKDNVPKATGVLYDYVNKYHWEHTGLTLREVSSKLRRNLQNNAE WVYQGAIRQIDDIDVRFQKAASGTTGTYQEWKDKAQNLYQELLTQEGQA SFQGLKDNVFDGLVRVTQEFHMKVKHLIDSLIDFLNFPRFQFPGKPGIYTR EELCTMFIREVGTVLSQVYSKVHNGSEILFSYFQDLVITLPFELRKHKLIDVI SMYRELLKDLSKEAQEVFKAIQSLKTTEVLRNLQDLLQFIFQLIEDNIKQLKE MKFTYLINYIQDEINTIFSDYIPYVFKLLKENLCLNLHKFNEFIQNELQEASQE LQQIHQYIMALREEYFDPSIVGWTVKYYELEEKIVSLIKNLLVALKDFHSEYI VSASNFTSQLSSQVEQFLHRNIQEYLSILTDPDGKGKEKIAELSATAQEIIKS QAIATKKIISDYHQQFRYKLQDFSDQLSDYYEKFIAESKRLIDLSIQNYHTFLI YITELLKKLQSTTVMNPYMKLAPGELTIIL [SEQ ID NO: 511]		
PCSK 9	255738	0169174	Q8NBP7	MGTVSSRRSWWPLPLLLLLLLLLGPAGARAQEDEDGDYEELVLALRSEED GLAEAPEHGTTATFHRCAKDPWRLPGTYVVVLKEETHLSQSERTARRLQA QAARRGYLTKILHVFHGLLPGFLVKMSGDLLELALKLPHVDYIEEDSSVFA QSIPWNLERITPPRYRADEYQPPDGGSLVEVYLLDTSIQSDHREIEGRVMV TDFENVPEEDGTRFHRQASKCDSHGTHLAGVVSGRDAGVAKGASMRSL RVLNCQGKGTVSGTLIGLEFIRKSQLVQPVGPLVVLLPLAGGYSRVLNAAC QRLARAGVVLVTAAGNFRDDACLYSPASAPEVITVGATNAQDQPVTLGTL GTNFGRCVDLFAPGEDIIGASSDCSTCFVSQSGTSQAAAHVAGIAAMMLS AEPELTLAELRQRLIHFSAKDVINEAWFPEDQRVLTPNLVAALPPSTHGAG WQLFCRTVWSAHSGPTRMATAVARCAPDEELLSCSSFSRSGKRRGERM EAQGGKLVCRAHNAFGGEGVYAIARCCLLPQANCSVHTAPPAEASMGTR VHCHQQGHVLTGCSSHWEVEDLGTHKPPVLRPRGQPNQCVGHREASIH	Hipercolesterole-mia familiar	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				ASCCHAPGLECKVKEHGIPAPQEQVTVACEEGWTLTGCSALPGTSHVLG AYAVDNTCVVRSRDVSTTGSTSEGAVTAVAICCRSRHLAQASQELQ [SEQ ID NO: 512]		
LDL-RAP1	26119	0157978	B3KR97, Q5SW96	MDALKSAGRALIRSPSLAKQSWGGGGRHRKLPENWTDTRETLLEGMLFS LKYLGMTLVEQPKGEELSAAAIKRIVATAKASGKKLQKVTLKVSPRGIILTD NLTNQLIENVSIYRISYCTADKMHDKVFAYIAQSQHNQSLECHAFLCTKRK MAQAVTLTVAQAFKVAFEFWQVSKEEKEKRDKASQEGGDVLGARQDCTP SLKSLVATGNLLDLEETAKAPLSTVSANTTNMDEVPRPQALSGSSVVWEL DDGLDEAFSRLAQSRTNPQVLDTGLTAQDMHYAQCLSPVDWDKPDSSGT EQDDLFSF [SEQ ID NO: 513]	Hipercolesterole-mia familiar	
ABCG 5	64240	0138075	Q9H222	MGDLSSLTPGGSMGLQVNRGSQSSLEGAPATAPEPHSLGILHASYSVSH RVRPWWDITSCRQQWTRQILKDVSLYVESGQIMCILGSSGSGKTTLLDAM SGRLGRAGTFLGEVYVNGRALRREQFQDCFSYVLQSDTLLSSLTVRETLH YTALLAIRRGNPGSFQKKVEAVMAELSLSHVADRLIGNYSLGGISTGERRR VSIAAQLLQDPKVMLFDEPTTGLDCMTANQIVVLLVELARRNRIVVLTIHQP RSELFQLFDKIAILSFGELIFCGTPAEMLDFFNDCGYPCPEHSNPFDFYMDL TSVDTQSKEREIETSKRVQMIESAYKKSAICHKTLKNIERMKHLKTLPMVPF KTKDSPGVFSKLGVLLRRVTRNLVRNKLAVITRLLQNLIMGLFLLFFVLRVR SNVLKGAIQDRVGLLYQFVGATPYTGMLNAVNLFPVLRAVSDQESQDGLY QKWQMMLAYALHVLPFSVVATMIFSSVCYWTLGLHPEVARFGYFSAALLA PHLIGEFLTLVLLGIVQNPNIVNSVVALLSIAGVLVGSGFLRNIQEMPIPFKIIS YFTFQKYCSEILVVNEFYGLNFTCGSSNVSVTTNPMCAFTQGIQFIEKTCP GATSRFTMNFLILYSFIPALVILGIVVFKIRDHLISR [SEQ ID NO: 514]	Sitosterolemia	
ABCG 8	64241	0143921	Q9H221	MAGKAAEERGLPKGATPQDTSGLQDRLFSSESDNSLYFTYSGQPNTLEV RDLNYQVDLASQVPWFEQLAQFKMPWTSPSCQNSCELGIQNLSFKVRSG QMLAIIGSSGCGRASLLDVITGRGHGGKIKSGQIWINGQPSSPQLVRKCVA HVRQHNQLLPNLTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQ CADTRVGNMYVRGLSGGERRRVSIGVQLLWNPGILILDEPTSGLDSFTAH NLVKTLSRLAKGNRLVLISLHQPRSDIFRLFDLVLLMTSGTPIYLGAAQHMV QYFTAIGYPCPRYSNPADFYVDLTSIDRRSREQELATREKAQSLAALFLEK	Sitosterolemia	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				VRDLDDFLWKAETKDLDEDTCVESSVTPLDTNCLPSPTKMPGAVQQFTTL IRRQISNDFRDLPTLLIHGAEACLMSMTIGFLYFGHGSIQLSFMDTAALLFMI GALIPFNVILDVISKCYSERAMLYYELEDGLYTTGPYFFAKILGELPEHCAYII IYGMPTYWLANLRPGLQPFLLHFLLVWLVVFCCRIMALAAAALLPTFHMAS FFSNALYNSFYLAGGFMINLSSLWTVPAWISKVSFLRWCFEGLMKIQFSRR TYKMPLGNLTIAVSGDKILSVMELDSYPLYAIYLIVIGLSGGFMVLYYVSLRFI KQKPSQDW [SEQ ID NO: 515]		
LCAT	3931	0213398	A0A140VK24, P04180	MGPPGSPWQWVTLLLGLLLPPAAPFWLLNVLFPPHTTPKAELSNHTRPVIL VPGCLGNQLEAKLDKPDVVNWMCYRKTEDFFTIWLDLNMFLPLGVDCWI DNTRVVYNRSSGLVSNAPGVQIRVPGFGKTYSVEYLDSSKLAGYLHTLVQ NLVNNGYVRDETVRAAPYDWRLEPGQQEEYYRKLAGLVEEMHAAYGKP VFLIGHSLGCLHLLYFLLRQPQAWKDRFIDGFISLGAPWGGSIKPMLVLAS GDNQGIPIMSSIKLKEEQRITTTSPWMFPSRMAWPEDHVFISTPSFNYTGR DFQRFFADLHFEEGWYMWLQSRDLLAGLPAPGVEVYCLYGVGLPTPRTYI YDHGFPYTDPVGVLYEDGDDTVATRSTELCGLWQGRQPQPVHLLPLHGI QHLNMVFSNLTLEHINAILLGAYRQGPPASPTASPEPPPPE [SEQ ID NO:516]	Deficiência de lecitina colesterol aciltransferase	
SPINK 5	11005	0133710	Q9NQ38	MKIATVSVLLPLALCLIQDAASKNEDQEMCHEFQAFMKNGKLFCPQDKKFF QSLDGIMFINKCATCKMILEKEAKSQKRARHLARAPKATAPTELNCDDFKK GERDGDFICPDYYEAVCGTDGKTYDNRCALCAENAKTGSQIGVKSEGEC KSSNPEQDVCSAFRPFVRDGRLGCTRENDPVLGPDGKTHGNKCAMCAE LFLKEAENAKREGETRIRRNAEKDFCKEYEKQVRNGRLFCTRESDPVRGP DGRMHGNKCALCAEIFKQRFSEENSKTDQNLGKAEEKTKVKREIVKLCSQ YQNQAKNGILFCTRENDPIRGPDGKMHGNLCSMCQAYFQAENEEKKKAE ARARNKRESGKATSYAELCSEYRKLVRNGKLACTRENDPIQGPDGKVHG NTCSMCEVFFQAEEEEKKKKEGKSRNKRQSKSTASFEELCSEYRKSRKN GRLFCTRENDPIQGPDGKMHGNTCSMCEAFFQQEERARAKAKREAAKEI CSEFRDQVRNGTLICTREHNPVRGPDGKMHGNKCAMCASVFKLEEEEKK NDKEEKGKVEAEKVKREAVQELCSEYRHYVRNGRLPCTRENDPIEGLDG KIHGNTCSMCEAFFQQEAKEKERAEPRAKVKREAEKETCDEFRRLLQNG KLFCTRENDPVRGPDGKTHGNKCAMCKAVFQKENEERKRKEEEDQRNA	Síndrome de Netherton	

Gene	Número de acesso Entrez	Nº(s) de Acesso ao Gene Ensembl (ENSG0000+ nº mostrado)	Número de Acesso à Proteína(s) Uniprot	Sequência de Aminoácidos (primeiro Número de Acesso Uniprot) [SEQ ID NO]	Doença/Distúrbio	Categoria
				AGHGSSGGGGGNTQDECAEYREQMKNGRLSCTRESDPVRDADGKSYN NQCTMCKAKLEREAERKNEYSRSRSNGTGSESGKDTCDEFRSQMKNGK LICTRESDPVRGPDGKTHGNKCTMCKEKLEREAAEKKKKEDEDRSNTGE RSNTGERSNDKEDLCREFRSMQRNGKLICTRENNPVRGPYGKMHINKCA MCQSIFDREANERKKKDEEKSSSKPSNNAKDECSEFRNYIRNNELICPRE NDPVHGADGKFYTNKCYMCRAVFLTEALERAKLQEKPSHVRASQEEDSP DSFSSLDSEMCKDYRVLPRIGYLCPKDLKPVCGDDGQTYNNPCMLCHEN LIRQTNTHIRSTGKCEESSTPGTTAASMPPSDE [SEQ ID NO: 517]		
GNE	10020	0159921	Q9Y223	MEKNGNNRKLRVCVATCNRADYSKLAPIMFGIKTEPEFFELDVVVLGSHLI DDYGNTYRMIEQDDFDINTRLHTIVRGEDEAAMVESVGLALVKLPDVLNRL KPDIMIVHGDRFDALALATSAALMNIRILHIEGGEVSGTIDDSIRHAITKLAHY HVCCTRSAEQHLISMCEDHDRILLAGCPSYDKLLSAKNKDYMSIIRMWLGD DVKSKDYIVALQHPVTTDIKHSIKMFELTLDALISFNKRTLVLFPNIDAGSKE MVRVMRKKGIEHHPNFRAVKHVPFDQFIQLVAHAGCMIGNSSCGVREVG AFGTPVINLGTRQIGRETGENVLHVRDADTQDKILQALHLQFGKQYPCSKI YGDGNAVPRILKFLKSIDLQEPLQKKFCFPPVKENISQDIDHILETLSALAVD LGGTNLRVAIVSMKGEIVKKYTQFNPKTYEERINLILQMCVEAAAEAVKLNC RILGVGISTGGRVNPREGIVLHSTKLIQEWNSVDLRTPLSDTLHLPVWVDN DGNCAALAERKFGQGKGLENFVTLITGTGIGGGIIHQHELIHGSSFCAAELG HLVVSLDGPDCSCGSHGCIEAYASGMALQREAKKLHDEDLLLVEGMSVPK DEAVGALHLIQAAKLGNAKAQSILRTAGTALGLGVVNILHTMNPSLVILSGV LASHYIHIVKDVIRQQALSSVQDVDVVVSDLVDPALLGAASMVLDYTTRRIY [SEQ ID NO:518]	Miopatia de corpo de inclusão 2	

[00545] Em algumas modalidades, o agente de proteína é diferente de um fator de coagulação, por exemplo, diferente de Fator VII ou Fator IX. Em algumas modalidades, o agente de proteína é diferente de uma proteína repórter, por exemplo, proteína fluorescente, por exemplo, GFP ou luciferase. Em alguns, o agente de proteína é diferente de um receptor de superfície celular, um receptor de NGF, galactocerebrosidase, gp91 phox, IFN-alfa, TK, GCV e antígeno autoimune, citocina, inibidor de angiogênese ou agente anticâncer ou um fragmento ou variante disso.

ELEMENTOS ISOLANTES

[00546] Em algumas modalidades, um fusossoma, retroviral ou vetor lentiviral ou VLP compreende adicionalmente um ou mais elementos isolantes, por exemplo, um elemento de isolante descrito no presente documento. Elementos isolantes podem contribuir para proteger sequências expressas de lentivírus, por exemplo, polipeptídeos terapêuticos, de efeitos de sítio de integração, que podem ser mediados por elementos de ação cis presentes no DNA genômico e levar à expressão desregulada de sequências transferidas (por exemplo, efeito de posição; consultar, por exemplo, Burgess-Beusse et al, 2002, Proc. Natl. Acad. Sci., USA, 99: 16.433; e Zhan et al, 2001, Hum. Genet., 109: 471) ou expressão desregulada de sequências endógenas adjacentes as sequências transferidas. Em algumas modalidades, os vetores de transferência compreendem um ou mais elementos isolantes da LTR 3', e mediante integração do provírus no genoma hospedeiro, o provírus compreende um ou mais isolantes na LTR 5' e/ou LTR 3', em virtude de duplicar a LTR 3'. Isolantes adequados incluem, porém sem limitação, isolante de β-globina de frango (consultar Chung et al, 1993. Cell 74:505; Chung et al, 1997. N4S 94:575; e Bell et al., 1999. Cell 98:387, incorporados no presente documento a título de referência) ou um isolante de um locus de β-globina humana, tal como HS4 de gali-

nha. Em algumas modalidades, o isolante se liga ao fator de ligação CCCTC (CTCF). Em algumas modalidades, o isolante é um isolante de barreira. Em algumas modalidades, o isolante é um isolante de bloqueio de intensificador. Consultar, por exemplo, Emery et al., *Human Gene Therapy*, 2011, e em Browning e Trobridge, Biomedicines, 2016, ambos incluídos em sua totalidade a título de referência.

[00547] Em algumas modalidades, os isolantes no ácido nucleico retroviral reduzem a genotoxicidade nas células receptoras. A genotoxicidade pode ser medida, por exemplo, conforme descrito em Cesana et al, “Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo” Mol Ther. Abril de 2014;22(4):774 a 785. doi: 10.1038/mt.2014.3. Publicação Eletrônica de 20 de janeiro de 2014.

FUSOSSOMAS DERIVADOS DE CÉLULAS

[00548] A presente divulgação fornece, em alguns aspectos, um fusossoma que compreende:

[00549] (a) uma bicamada lipídica,

[00550] (b) um lúmen (por exemplo, que compreende citosol) cercado pela bicamada lipídica;

[00551] (c) um fusogênio exógeno ou superexpresso, por exemplo, em que o fusogênio está disposto na bicamada lipídica,

[00552] em que o fusossoma é derivado de uma célula de origem; e

[00553] em que o fusossoma tem inativação nuclear parcial ou completa (por exemplo, remoção nuclear).

[00554] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00555] (a) uma bicamada lipídica,

[00556] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00557] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00558] (d) um ácido nucleico, por exemplo, um ácido nucleico que compreende um gene de carga útil; e

[00559] em que o fusossoma não compreende um núcleo;

[00560] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total;

[00561] em que:

[00562] (i) quando a pluralidade de fusossomas são contatados com uma população de células que compreendem células-alvo e células não alvo, a carga está presente em pelo menos 10 vezes mais em células-alvo do que em células não alvo ou células de referência, ou

[00563] (ii) os fusossomas da pluralidade são fundidos com uma célula-alvo a uma taxa mais alta do que com uma célula não alvo ou célula de referência em pelo menos 50%;

[00564] em que a célula-alvo é escolhida a partir de uma célula endotelial sinusoidal do fígado, colangiócito, célula estrelada, célula apresentadora de antígeno residente no fígado (por exemplo, célula de Kupffer), linfócito imune residente no fígado (por exemplo, célula T, célula B ou célula NK) ou fibroblasto portal.

[00565] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00566] (a) uma bicamada lipídica,

[00567] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00568] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00569] (d) um ácido nucleico que compreende um gene de carga

útil que codifica um agente exógeno da Tabela 5,

[00570] em que o fusossoma não compreende um núcleo; e

[00571] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total.

[00572] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00573] (a) uma bicamada lipídica,

[00574] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00575] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00576] (d) um ácido nucleico que compreende um gene de carga útil, em que o ácido nucleico compreende um NTCSRE ligado operacionalmente ao gene de carga útil, em que o NTCSRE compreende uma sequência de reconhecimento de miRNA específico de células não hepáticas, por exemplo, um miRNA específico de células não hepáticas sequência de reconhecimento ligada por um miRNA presente em uma célula hematopoiética ou uma célula dendrítica plasmocitoide (pDC), por exemplo, por um miRNA presente em uma célula hematopoiética ou uma célula dendrítica plasmocitoide (pDC) em um nível mais alto do que em uma célula hepática, por exemplo, uma sequência de reconhecimento de miRNA específica para células não hepáticas ligada por um miRNA da Tabela 4; e

[00577] em que o fusossoma não compreende um núcleo; e

[00578] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total.

[00579] Em algumas modalidades, o miRNA está presente em uma célula não alvo (por exemplo, uma célula hematopoiética ou um pDC)

em um nível pelo menos 10, 100, 1.000 ou 10.000 vezes superior ao nível do miRNA presente na célula-alvo (por exemplo, uma célula hepática, por exemplo, uma célula hepática descrita no presente documento). Em algumas modalidades, o miRNA não está detectavelmente presente em uma célula-alvo (por exemplo, uma célula hepática). Em algumas modalidades, o miRNA não está presente na célula-alvo (por exemplo, uma célula hepática).

[00580] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00581] (a) uma bicamada lipídica,

[00582] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00583] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00584] (d) um ácido nucleico que compreende um gene de carga útil, em que o ácido nucleico compreende um promotor ligado operacionalmente ao gene de carga útil, em que o promotor é um promotor específico do fígado, por exemplo, é um promotor específico para uma célula endotelial sinusoidal hepática, colangiócito, célula estrelada, célula apresentadora de antígeno residente no fígado (por exemplo, célula de Kupffer), linfócito imune residente no fígado (por exemplo, célula T, célula B ou célula NK) ou fibroblasto portal;

[00585] em que o fusossoma não compreende um núcleo; e

[00586] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total.

[00587] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusosso-

mas da pluralidade compreendem:

[00588] (a) uma bicamada lipídica,

[00589] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00590] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00591] (d) um ácido nucleico que compreende um gene de carga útil, em que o ácido nucleico compreende um promotor que tem a sequência da Tabela 3 ou uma sequência que tem pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com o mesmo;

[00592] em que o fusossoma não compreende um núcleo; e

[00593] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total.

[00594] A presente divulgação fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00595] (a) uma bicamada lipídica,

[00596] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00597] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00598] (d) um ácido nucleico que compreende:

[00599] (i) um gene de carga útil;

[00600] (ii) um NTCSRE ligado operacionalmente ao gene de carga útil, por exemplo, em que o NTCSRE compreende uma sequência de reconhecimento de miRNA específico de células não hepáticas, por exemplo, uma sequência de reconhecimento de miRNA específicos de células não hepáticas ligada por um miRNA da Tabela 4, e

[00601] (iii) opcionalmente, um elemento regulatório específico de célula-alvo positivo, por exemplo, um elemento regulatório específico de célula do fígado positivo (por exemplo, um promotor específico da célula hepática) ligado operacionalmente ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula-alvo em relação a um fusossoma de outra forma similar sem o elemento regulatório específico de célula-alvo positivo;

[00602] em que o fusossoma não compreende um núcleo; e

[00603] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total.

[00604] Em algumas modalidades, um ou mais dos seguintes está presente:

[00605] i) o fusossoma compreende ou é compreendido por um citobiológico;

[00606] ii) o fusossoma compreende uma célula enucleada;

[00607] iii) o fusossoma compreende um núcleo inativado;

[00608] iv) o fusossoma é fundido com uma célula-alvo a uma taxa mais alta que com uma célula não alvo, por exemplo, em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes, por exemplo, em um ensaio do Exemplo 43;

[00609] v) o fusossoma é fundido com uma célula-alvo a uma taxa mais alta do que com outros fusossomas, por exemplo, por pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, ou 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes, por exemplo, em um ensaio do Exemplo 43;

[00610] vi) o fusossoma é fundido com células-alvo a uma taxa de modo que um agente no fusossoma é entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% de células-alvo após

24, 48 ou 72 horas, por exemplo, em um ensaio do Exemplo 43;

[00611] vii) o fusogênio está presente em um número de cópias de pelo menos, ou não mais do que, 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias, por exemplo, conforme medido por um ensaio do Exemplo 27;

[00612] viii) o fusossoma compreende um agente terapêutico em um número de cópias de pelo menos, ou no máximo, 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias, por exemplo, conforme medido por um ensaio de Exemplo 89;

[00613] ix) a razão entre o número de cópias do fusogênio e o número de cópias do agente terapêutico está entre 1.000.000:1 e 100.000:1, 100.000:1 e 10.000:1, 10.000:1 e 1.000:1, 1.000:1 e 100:1, 100:1 e 50:1, 50:1 e 20:1, 20:1 e 10:1, 10:1 e 5:1, 5:1 e 2:1, 2:1 e 1:1, 1:1 e 1:2, 1:2 e 1:5, 1:5 e 1:10, 1:10 e 1:20, 1:20 e 1:50, 1:50 e 1:100, 1:100 e 1:1.000, 1:1.000 e 1:10.000, 1:10.000 e 1:100.000 ou 1:100.000 e 1:1.000.000;

[00614] x) o fusossoma compreende uma composição lipídica substancialmente semelhante à da célula de origem ou em que um ou mais dentre CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM e TAG está dentro de 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70% ou 75% do nível de lipídio correspondente na célula de origem;

[00615] xi) o fusossoma compreende uma composição proteômica semelhante à da célula de origem, por exemplo, com o uso de um ensaio de Exemplo 88;

[00616] xii) o fusossoma compreende uma razão entre lipídios para

proteínas que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um ensaio do Exemplo 41;

[00617] xiii) o fusossoma compreende uma razão entre proteínas e ácidos nucleicos (por exemplo, DNA) que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um ensaio do Exemplo 42;

[00618] xiv) o fusossoma compreende uma razão entre lipídios para ácidos nucleicos (por exemplo, DNA) que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um ensaio do Exemplo 92;

[00619] xv) o fusossoma tem meia-vida em um indivíduo, por exemplo, em um camundongo, que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% da meia-vida de uma célula de referência, por exemplo, a célula de origem, por exemplo, por um ensaio do Exemplo 61;

[00620] xvi) o fusossoma transporta glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) através de uma membrana, por exemplo, em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40 %, 50%, 60%, 70%, 80%, 90%, 100% a mais (por exemplo, de aproximadamente 11,6% a mais) que um controle negativo, por exemplo, um fusossoma semelhante na ausência de glicose, por exemplo, conforme medido com o uso de um ensaio do Exemplo 51;

[00621] xvii) o fusossoma compreende atividade esterase no lúmen que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% da atividade da esterase em uma célula de referência, por exemplo, a célula de origem ou um fibroblasto embrionário de camundongo, por exemplo, com o uso de um ensaio do Exemplo 52;

[00622] xviii) o fusossoma compreende um nível de atividade meta-

bólica que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90 % ou 100% da atividade da citrato sintase em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito no Exemplo 54;

[00623] xix) o fusossoma compreende um nível de respiração (por exemplo, taxa de consumo de oxigênio) que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70 %, 80%, 90% ou 100% do nível de respiração em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito em Exemplo 55;

[00624] xx) o fusossoma compreende um nível de marcação de Anexina-V de no máximo 18.000, 17.000, 16.000, 15.000, 14.000, 13.000, 12.000, 11.000 ou 10.000 MFI, por exemplo, com o uso de um ensaio do Exemplo 56, ou em que o fusossoma compreende um nível de marcação de Anexina-V pelo menos 5%, 10%, 20%, 30%, 40% ou 50% mais baixo que o nível de marcação de Anexina-V de um fusossoma semelhante tratado com menadiona no ensaio do Exemplo 56, ou em que o fusossoma compreende um nível de marcação de Anexina-V pelo menos 5%, 10%, 20%, 30%, 40% ou 50% menor que o nível de marcação de Anexina-V de um macrófago tratado com menadiona no ensaio do Exemplo 56,

[00625] xxi) o fusossoma tem um nível de conteúdo de miRNA de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou maior que a célula de origem, por exemplo, por um ensaio do Exemplo 34;

[00626] xxii) o fusossoma tem uma razão entre proteína solúvel:não solúvel dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80 %, 90% ou maior do que a célula de origem, por exemplo, dentro de 1% a 2%, 2% a 3%, 3% a 4%, 4% a 5%, 5% a 10%, 10% a 20%, 20% a 30%, 30% a 40%, 40% a 50%, 50% a 60%, 60% a

70%, 70% a 80% ou 80% a 90% daquele da célula de origem, por exemplo, por um ensaio do Exemplo 39;

[00627] xxiii) o fusossoma tem um nível de LPS inferior a 5%, 1%, 0,5%, 0,01%, 0,005%, 0,0001%, 0,00001% ou menos do teor de LPS da célula de origem, por exemplo, conforme medido por espectrometria de massa, por exemplo, em um ensaio do Exemplo 40.

[00628] xxiv) o fusossoma tem capacidade para transdução de sinal, por exemplo, transmitir um sinal extracelular, por exemplo, fosforilação AKT em resposta à insulina ou absorção de glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) em resposta à insulina, por exemplo, por pelo menos 1 %, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% a mais que um controle negativo, por exemplo, um fusossoma de outra forma similar na ausência de insulina, por exemplo, com o uso de um ensaio do Exemplo 50.

[00629] xxv) o fusossoma tem como alvo um tecido, por exemplo, fígado, pulmões, coração, baço, pâncreas, trato gastrointestinal, rim, testículos, ovários, cérebro, órgãos reprodutivos, sistema nervoso central, sistema nervoso periférico, músculo esquelético, endotélio, ouvido interno ou olho, quando administrado a um indivíduo, por exemplo, um camundongo, por exemplo, em que pelo menos 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% dos fusossomas em uma população de fusossomas administrados estão presentes no tecido-alvo após 24, 48 ou 72 horas, por exemplo, por um ensaio do Exemplo 65;

[00630] xxvi) o fusossoma tem nível de sinalização justacrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% a mais que o nível de sinalização justacrina induzida por uma célula de referência, por exemplo, a célula de origem ou uma célula estromal da medula óssea (BMSC), por exemplo, por

um ensaio do Exemplo 57.

[00631] xxvii) o fusossoma tem nível de sinalização parácrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% a mais que o nível de sinalização parácrina induzida por uma célula de referência, por exemplo, a célula de origem ou um macrófago, por exemplo, por um ensaio do Exemplo 58;

[00632] xxviii) o fusossoma polimeriza a actina em um nível dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% em comparação ao nível de actina polimerizada em uma célula de referência, por exemplo, a célula de origem ou uma célula C2C12, por exemplo, pelo ensaio do Exemplo 59;

[00633] xxix) o fusossoma tem um potencial de membrana dentro de aproximadamente 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% do potencial de membrana de uma célula de referência, por exemplo, a célula de origem ou uma célula C2C12, por exemplo, por um ensaio do Exemplo 60 ou em que o fusossoma tem um potencial de membrana de aproximadamente -20 a -150 mV, -20 a -50 mV, -50 a -100 mV ou -100 a -150 mV;

[00634] xxx) o fusossoma tem capacidade para extravasamento dos vasos sanguíneos, por exemplo, a uma taxa de pelo menos 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, ou 90% da taxa de extravasamento da célula de origem ou de uma célula do mesmo tipo que a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 45, por exemplo, em que a célula de origem é um neutrófilo, linfócito, célula B, macrófago ou célula NK;

[00635] xxxi) o fusossoma tem capacidade para atravessar uma membrana celular, por exemplo, uma membrana celular endotelial ou a barreira hematoencefálica;

[00636] xxxii) o fusossoma tem capacidade para secretar uma proteína, por exemplo, a uma taxa de pelo menos 1%, 2%, 3%, 4%, 5%,

10%, 20%, 30%, 40%, 50%, 60%, 70 %, 80%, 90% ou 100% a mais que uma célula de referência, por exemplo, um fibroblasto embrionário de camundongo, por exemplo, com o uso de um ensaio do Exemplo 49;

[00637] xxxiii) o fusossoma atende a um padrão farmacêutico ou de boas práticas de fabricação (GMP);

[00638] xxxiv) o fusossoma foi produzido de acordo com as boas práticas de fabricação (GMP);

[00639] xxxv) o fusossoma tem um nível de patógeno abaixo de um valor de referência predeterminado, por exemplo, está substancialmente livre de patógenos;

[00640] xxxvi) o fusossoma tem um nível de contaminante abaixo de um valor de referência predeterminado, por exemplo, é substancialmente livre de contaminantes;

[00641] xxxvii) o fusossoma tem baixa imunogenicidade, por exemplo, conforme descrito no presente documento;

[00642] xxxviii) a célula de origem é selecionada a partir de um neutrófilo, um granulócito, uma célula-tronco mesenquimal, uma célula-tronco da medula óssea, uma célula-tronco pluripotente induzida, uma célula-tronco embrionária, um mieloblasto, um mioblasto, um hepatócito ou um neurônio, por exemplo, célula neuronal retiniana; ou

[00643] xxxix) a célula de origem é outra que não uma célula 293, célula HEK, célula endotelial humana ou uma célula epitelial humana, monócito, macrófago, célula dendrítica ou célula-tronco.

[00644] A presente divulgação também fornece, em alguns aspectos, um fusossoma que compreende:

[00645] a) uma bicamada lipídica e um lúmen que é miscível com uma solução aquosa, por exemplo, água, em que o fusossoma é derivado de uma célula de origem,

[00646] b) um fusogênio exógeno ou superexpresso disposto na

bicamada lipídica, e

[00647] c) uma organela, por exemplo, um número terapeuticamente eficaz de organelas, dispostas no lúmen.

[00648] Em algumas modalidades, um ou mais dos seguintes está presente:

[00649] i) a célula de origem é selecionada a partir de uma célula endotelial, um macrófago, um neutrófilo, um granulócito, um leucócito, uma célula-tronco (por exemplo, uma célula-tronco mesenquimal, uma célula-tronco da medula óssea, uma célula-tronco pluripotente induzida, uma célula-tronco embrionária), um mieloblasto, um mioblasto, um hepatócito ou um neurônio, por exemplo, célula neuronal retiniana;

[00650] ii) a organela é selecionada a partir de um aparelho de Golgi, lisossoma, retículo endoplasmático, mitocôndria, vacúolo, endossoma, acrossoma, autofagossomo, centríolo, glicosomo, glioxissomo, hidrogenossoma, melanossoma, mitossomo, cnidocisto, peroxissoma, proteassoma, vesícula e grânulo de tensão;

[00651] iii) o fusossoma tem um tamanho maior que 5 um, 10 um, 20 um, 50 um ou 100 um;

[00652] i) o fusossoma, ou uma composição ou preparação compreendendo uma pluralidade de fusossomas, tem uma densidade diferente de entre 1,08 g/ml e 1,12 g/ml, por exemplo, o fusossoma tem uma densidade de 1,25 g/ml +/- 0,05, por exemplo, conforme medido por um ensaio de Exemplo 31;

[00653] iv) o fusossoma não é capturado pelo sistema de eliminação em circulação ou pelas células de Kupffer no seio do fígado;

[00654] v) a célula de origem é diferente de uma célula 293;

[00655] vi) a célula de origem não é transformada ou imortalizada;

[00656] vii) a célula de origem é transformada ou imortalizada com o uso de um método diferente de imortalização mediada por adenovírus, por exemplo, imortalizada por mutação espontânea ou expressão

de telomerase;

[00657] viii) o fusogênio é diferente de VSVG, uma proteína SNARE ou uma proteína granular secretora;

[00658] ix) o fusossoma não compreende Cre ou GFP, por exemplo, EGFP;

[00659] x) o fusossoma compreende ainda uma proteína exógena diferente de Cre ou GFP, por exemplo, EGFP

[00660] xi) o fusossoma compreende adicionalmente um ácido nucleico exógeno (por exemplo, RNA, por exemplo, mRNA, miRNA ou siRNA) ou uma proteína exógena (por exemplo, um anticorpo, por exemplo, um anticorpo), por exemplo, no lúmen; ou

[00661] xii) o fusossoma não compreende mitocôndrias.

[00662] A presente divulgação também fornece, em alguns aspectos, um fusossoma que compreende:

[00663] (a) uma bicamada lipídica,

[00664] (b) um lúmen (por exemplo, que compreende citosol) rodeado pela bicamada lipídica,

[00665] (c) um fusogênio exógeno ou superexpresso, por exemplo, em que o fusogênio está disposto na bicamada lipídica, e

[00666] (d) um núcleo funcional,

[00667] em que o fusossoma é derivado de uma célula de origem.

[00668] Em algumas modalidades, um ou mais dos seguintes está presente:

[00669] i) a célula de origem é diferente de uma célula dendrítica ou célula tumoral, por exemplo, a célula de origem é selecionada a partir de uma célula endotelial, um macrófago, um neutrófilo, um granulócito, um leucócito, uma célula-tronco (por exemplo, uma célula-tronco mesenquimal, um célula-tronco da medula óssea, uma célula-tronco pluripotente induzida, uma célula-tronco embrionária), um mieloblasto, um mioblasto, um hepatócito ou um neurônio, por exemplo, célula neuro-

nal retiniana;

[00670] ii) o fusogênio é diferente de uma glicoproteína fusogênica;

[00671] iii) o fusogênio é uma proteína de mamífero diferente da fertilina-beta,

[00672] iv) o fusossoma tem baixa imunogenicidade, por exemplo, conforme descrito no presente documento;

[00673] v) o fusossoma atende a um padrão farmacêutico ou de boas práticas de fabricação (GMP);

[00674] vi) o fusossoma foi produzido de acordo com as boas práticas de fabricação (GMP);

[00675] vii) o fusossoma tem um nível de patógeno abaixo de um valor de referência predeterminado, por exemplo, está substancialmente livre de patógenos; ou

[00676] viii) o fusossoma tem um nível de contaminante abaixo de um valor de referência predeterminado, por exemplo, é substancialmente livre de contaminantes.

[00677] A presente divulgação também fornece, em alguns aspectos, uma composição de fusossomas que compreendem uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem:

[00678] (a) uma bicamada lipídica,

[00679] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00680] (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica,

[00681] (d) uma carga; e

[00682] em que o fusossoma não compreende um núcleo;

[00683] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total;

[00684] em que a pluralidade de fusossomas, quando em contato

com uma população de células-alvo na presença de um inibidor de endocitose, e quando em contato com uma população de células-alvo de referência não tratada com o inibidor de endocitose, entrega a carga a pelo menos 30% do número de células na população de células-alvo em comparação à população de células-alvo de referência.

[00685] A presente divulgação também fornece, em alguns aspectos, uma composição de fusossomas que compreende uma pluralidade de fusossomas derivados de uma célula de origem, e em que os fusossomas da pluralidade compreendem:

[00686] (a) uma bicamada lipídica,

[00687] (b) um lúmen que compreende citosol, em que o lúmen é rodeado pela bicamada lipídica;

[00688] (c) um fusogênio exógeno ou superexpresso realvejado disposto na bicamada lipídica;

[00689] (d) uma carga; e

[00690] em que o fusossoma não compreende um núcleo;

[00691] em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total;

[00692] em que:

[00693] (i) quando a pluralidade de fusossomas é colocada em contato com uma população de células que compreende células-alvo e células não alvo, a carga está presente em pelo menos 2 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes mais células-alvo do que células não alvo, ou

[00694] (ii) os fusossomas da pluralidade é fundidom a uma taxa mais alta com uma célula-alvo do que com uma célula não alvo em pelo menos 50%.

[00695] A presente divulgação também fornece, em alguns aspectos, uma composição de fusossomas que compreende uma pluralidade de fusossomas derivados de uma célula de origem, e em que os

fusossomas da pluralidade compreendem:

[00696] (a) uma bicamada lipídica,

[00697] (b) um lúmen cercado pela bicamada lipídica;

[00698] (c) um fusogênio exógeno ou superexpresso, em que o fusogênio está disposto na bicamada lipídica; e

[00699] (d) uma carga;

[00700] em que o fusossoma não compreende um núcleo; e

[00701] em que um ou mais dentre (por exemplo, pelo menos 2, 3, 4 ou 5):

[00702] i) o fusogênio está presente em um número de cópias de pelo menos 1.000 cópias;

[00703] ii) o fusossoma compreende um agente terapêutico em um número de cópias de pelo menos 1.000 cópias;

[00704] iii) o fusossoma compreende um lipídio em que um ou mais dentre CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM e TAG está dentro de 75 % do nível de lipídio correspondente na célula de origem;

[00705] iv) o fusossoma compreende uma composição proteômica semelhante à da célula de origem;

[00706] v) o fusossoma tem capacidade para transdução de sinal, por exemplo, transmitir um sinal extracelular, por exemplo, fosforilação AKT em resposta à insulina, ou absorção de glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) em resposta à insulina, por exemplo, por pelo menos 10% a mais do que um controle negativo, por exemplo, um fusossoma semelhante na ausência de insulina;

[00707] o fusossoma tem como alvo um tecido, por exemplo, fígado, pulmões, coração, baço, pâncreas, trato gastrointestinal, rim, testículos, ovários, cérebro, órgãos reprodutivos, sistema nervoso central, sistema nervoso periférico, músculo esquelético, endotélio, ouvido interno ou olho, quando administrado a um indivíduo, por exemplo, um

camundongo, por exemplo, em que pelo menos 0,1%, ou 10%, dos fusossomas em uma população de fusossomas administrados estão presentes no tecido-alvo após 24 horas; ou

[00708] a célula de origem é selecionada a partir de um neutrófilo, um granulócito, uma célula-tronco mesenquimal, uma célula-tronco da medula óssea, uma célula-tronco pluripotente induzida, uma célula-tronco embrionária, um mieloblasto, um mioblasto, um hepatócito ou um neurônio, por exemplo, célula neuronal retiniana.

[00709] A presente divulgação também fornece, em alguns aspectos, uma composição farmacêutica que compreende a composição de fusossoma descrita no presente documento e um carreador farmaceuticamente aceitável.

[00710] A presente divulgação também fornece, em determinados aspectos, um método para administrar uma composição de fusossoma a um indivíduo (por exemplo, um indivíduo humano), um tecido-alvo, ou uma célula, que compreende administrar ao indivíduo uma composição de fusossoma que compreende uma pluralidade de fusossomas descritos no presente documento, ou colocar o tecido-alvo ou, a célula, em contato com a mesma, uma composição de fusossomas descrita no presente documento ou uma composição farmacêutica descrita no presente documento, desse modo, administrando a composição de fusossomas ao indivíduo.

[00711] A presente divulgação também fornece, em determinados aspectos, um método para entregar um agente terapêutico (por exemplo, um polipeptídeo, um ácido nucleico, um metabólito, uma organela ou uma estrutura subcelular) a um indivíduo, a um tecido-alvo ou a uma célula que compreende administrar ao indivíduo uma composição de fusossomas que compreende uma pluralidade de fusossomas descritos no presente documento, ou colocar o tecido-alvo, ou uma célula, em contato com a mesma, uma composição de fusossomas descrita

no presente documento ou uma composição farmacêutica descrita no presente documento, em que a composição de fusossomas é administrada em uma quantidade e/ou tempo para que o agente terapêutico seja administrado.

[00712] A presente divulgação também fornece, em determinados aspectos, um método para entregar uma função a um indivíduo a um tecido-alvo, ou a uma célula, que compreende administrar ao indivíduo uma pluralidade de fusossomas descritos no presente documento, ou colocar o tecido-alvo ou a célula em contato com a mesma, uma composição de fusossoma que compreende uma pluralidade de fusossomas descritos no presente documento, uma composição de fusossomas descrita no presente documento ou uma composição farmacêutica descrita no presente documento, em que a composição de fusossomas é administrada em uma quantidade e/ou tempo para que a função seja entregue.

[00713] Nas modalidades, um ou mais dentre:

[00714] i) a célula de origem é diferente de uma célula 293;

[00715] ii) a célula de origem não é transformada ou imortalizada;

[00716] iii) a célula de origem é transformada ou imortalizada com o uso de um método diferente da imortalização mediada por adenovírus, por exemplo, imortalizada por mutação espontânea ou expressão de telomerase;

[00717] iv) o fusogênio é diferente de VSVG, uma proteína SNARE ou uma proteína granular secretora;

[00718] v) o agente terapêutico é diferente de Cre ou EGFP;

[00719] vi) o agente terapêutico é um ácido nucleico (por exemplo, RNA, por exemplo, mRNA, miRNA ou siRNA) ou uma proteína exógena (por exemplo, um anticorpo, por exemplo, um anticorpo), por exemplo, no lúmen; ou

[00720] vii) o fusossoma não compreende mitocôndrias.

[00721] Nas modalidades, um ou mais dentre:

[00722] i) a célula de origem é diferente de uma célula 293 ou HEK;

[00723] ii) a célula de origem não é transformada ou imortalizada;

[00724] iii) a célula de origem é transformada ou imortalizada com o uso de um método diferente da imortalização mediada por adenovírus, por exemplo, imortalizada por mutação espontânea ou expressão de telomerase;

[00725] iv) o fusogênio não é um fusogênio viral; ou

[00726] v) o fusossoma tem um tamanho diferente de entre 40 e 150 nm, por exemplo, maior que 150 nm, 200 nm, 300 nm, 400 nm ou 500 nm.

[00727] Nas modalidades, um ou mais dentre:

[00728] i) o agente terapêutico é uma proteína solúvel expressa pela célula de origem;

[00729] ii) o fusogênio é diferente de TAT, TAT-HA2, HA-2, gp41, peptídeo beta-amiloide de Alzheimer, uma proteína do vírus Sendai ou peptídeo negativo anfipático (WAE 11);

[00730] iii) o fusogênio é um fusogênio de mamífero;

[00731] iv) o fusossoma compreende em seu lúmen um polipeptídeo selecionado a partir de uma enzima, anticorpo ou polipeptídeo antiviral;

[00732] v) o fusossoma não compreende uma proteína transmembranar terapêutica exógena; ou

[00733] vi) o fusossoma não compreende CD63 ou GLUT4 ou o fusossoma compreende 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5% ou 10% ou menos de CD63 (por exemplo, de aproximadamente 0,048% ou menos), por exemplo, conforme determinado de acordo com o método descrito no Exemplo 90.

[00734] Em modalidades, o fusossoma:

[00735] i) não compreende um vírus, não é infeccioso ou não se

propaga em uma célula hospedeira;

[00736] ii) não é um vetor viral

[00737] iii) não é uma VLP (partícula semelhante a vírus);

[00738] iv) não compreende uma proteína estrutural viral, por exemplo, uma proteína derivada de gag, por exemplo, uma proteína capsidial viral, por exemplo, uma proteína capsidial viral, por exemplo, uma proteína do nucleocapsídeo viral, ou em que a quantidade de proteína capsidial viral é inferior a 10%, 5 %, 4%, 3%, 2%, 1%, 0,5%, 0,2% ou 0,1% da proteína total, por exemplo, por espectrometria de massa, por exemplo, com o uso de um ensaio do Exemplo 94;

[00739] v) não compreende uma proteína da matriz viral;

[00740] vi) não compreende uma proteína não estrutural viral; por exemplo, pol ou um fragmento ou variante deste, uma proteína de transcriptase reversa viral, uma proteína de integrase viral ou uma proteína de protease viral.

[00741] vii) não compreende ácido nucleico viral; por exemplo, RNA viral ou DNA viral;

[00742] viii) compreende menos de 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias por vesícula de uma estrutura viral proteína; ou

[00743] ix) o fusossoma não é um virossoma.

[00744] Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) menos que de aproximadamente 0,01%, 0,05%, 0,1%, 0,5%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% ou 99% da proteína capsidial viral (por exemplo, de aproximadamente 0,05% da proteína capsidial viral). Em modalidades, a proteína capsidial viral é o Complexo do Capsídeo de Lentivírus Endógeno de Coelho (RE-

LIK) com ciclofilina A. Em modalidades, a razão proteína capsidial viral:proteína total é (ou é identificada como sendo) de aproximadamente 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09 ou 0,1.

[00745] Em algumas modalidades, o fusossoma não compreende (ou é identificado como não compreendendo) uma proteína gag ou um fragmento ou variante da mesma, ou a quantidade de proteína gag ou fragmento ou variante do mesmo é inferior a 10%, 5%, 4%, 3%, 2%, 1%, 0,5%, 0,2% ou 0,1% da proteína total, por exemplo, por um ensaio do Exemplo 94.

[00746] Nas modalidades, a razão entre o número de cópias do fusogênio e o número de cópias de proteína estrutural viral no fusossoma é pelo menos 1.000.000:1, 100.000:1, 10.000:1, 1.000:1, 100:1, 50:1, 20:1, 10:1, 5:1 ou 1:1; ou é entre 100:1 e 50:1, 50:1 e 20:1, 20:1 e 10:1, 10:1 e 5:1 ou 1:1. Em modalidades, a razão do número de cópias do fusogênio para o número de cópias da proteína da matriz viral no fusossoma é de pelo menos 1.000.000:1, 100.000:1, 10.000:1, 1.000:1, 100:1, 50:1, 20:1, 10:1, 5:1 ou 1:1.

[00747] Nas modalidades, um ou mais dentre:

[00748] i) o fusossoma não compreende uma gota imiscível em água;

[00749] ii) o fusossoma compreende um lúmen aquoso e um exterior hidrofílico;

[00750] iii) o fusogênio é uma proteína fusogênica; ou

[00751] iv) a organela é selecionada a partir de uma mitocôndria, aparelho de Golgi, lisossoma, retículo endoplasmático, vacúolo, endossoma, acrossoma, autofagossomo, centríolo, glicosomo, glioxissomo, hidrogenossoma, melanossomo, mitossomo, cnidocisto, peroxissomo, proteassoma, vesícula e grânulo de tensão.

[00752] Nas modalidades, um ou mais dentre:

[00753] i) o fusogênio é um fusogênio de mamífero ou um fusogênio

viral;

[00754] ii) o fusossoma não foi produzido carregando o fusossoma com uma substância terapêutica ou diagnóstica;

[00755] iii) a célula de origem não estava carregada com uma substância terapêutica ou diagnóstica;

[00756] iv) o fusossoma não compreende doxorrubicina, dexametasona, ciclodextrina; polietilenoglicol, um micro-RNA, por exemplo, miR125, receptor VEGF, ICAM-1, E-selectina, óxido de ferro, uma proteína fluorescente, por exemplo, GFP ou RFP, uma nanopartícula ou um RNase ou não compreende uma forma exógena dentre qualquer uma das supracitadas; ou

[00757] v) o fusossoma compreende adicionalmente um agente terapêutico exógeno com uma ou mais modificações pós-tradução, por exemplo, glicosilação.

[00758] Em modalidades, o fusossoma é unilamelar ou multilamelar.

[00759] Nas modalidades, um ou mais dentre:

[00760] i) o fusossoma não é um exossoma;

[00761] ii) o fusossoma é uma microvesícula;

[00762] iii) o fusossoma compreende um fusogênio de não mamífero;

[00763] iv) o fusossoma foi projetado para incorporar um fusogênio;

[00764] v) o fusossoma compreende um fusogênio exógeno;

[00765] vi) o fusossoma tem um tamanho de pelo menos 80 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm ou 1500 nm, ou uma população de fusossomas tem um tamanho médio de pelo menos 80 nm, 100 nm, 200 nm, 500 nm, 1.000 nm, 1.200 nm, 1.400 nm ou 1.500 nm;

[00766] vii) o fusossoma compreende uma ou mais organelas, por exemplo, uma mitocôndria, aparelho de Golgi, lisossoma, retículo en-

doplasmático, vacúolo, endossoma, acrossoma, autofagossomo, centríolo, glicosomo, glioxissoma, hidrogenossoma, melanossomo, mitossomo, cnidocisto, peroxissomo, proteassoma e grânulo de tensão;

[00767] viii) o fusossoma compreende um citoesqueleto ou um componente do mesmo, por exemplo, actina, Arp2/3, formina, coronina, distrofina, queratina, miosina ou tubulina;

[00768] ix) o fusossoma ou uma composição ou preparação que compreende uma pluralidade de fusossomas não tem uma densidade de flotação de 1,08 a 1,22 g/ml ou tem uma densidade de pelo menos 1,18 a 1,25 g/ml ou 1,05 a 1,12 g/ml, por exemplo, em um ensaio de centrifugação de gradiente de sacarose, por exemplo, conforme descrito em Théry et al., “Isolation and characterization of exosomes from cell culture supernatants and biological fluids”. Curr Protoc Cell Biol. Abril de 2006; Capítulo 3:Unidade 3.22;

[00769] x) a bicamada lipídica é enriquecida em ceramidas ou esfingomielinas ou uma combinação das mesmas em comparação à célula de origem ou a bicamada lipídica não é enriquecida (por exemplo, está depletada) em glicolipídios, ácidos graxos livres ou fosfatidilserina ou uma combinação dos mesmos, em comparação à célula de origem;

[00770] xi) o fusossoma compreende um ligante de ligante de Fosfatidilserina (PS) ou de CD40 ou tanto um ligante de PS quanto um ligante de CD40, por exemplo, quando medido em um ensaio do Exemplo 93;

[00771] xii) o fusossoma é enriquecido para PS em comparação à célula de origem, por exemplo, em uma população de fusossomas de pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% são positivos para PS, por exemplo, por um ensaio de Kanada M, et al. (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. Proc Natl Acad Sci EUA 112:E1433 a E1442;

[00772] xiii) o fusossoma é substancialmente livre de acetilcolines-

terase (AChE) ou contém menos de 0,001, 0,002, 0,005, 0,01, 0,02, 0,05, 0,1, 0,2, 0,5, 1, 2, 5, 10, 20, 50, 100, 200, 500 ou 1.000 unidades de atividade AChE/ug de proteína, por exemplo, por um ensaio do Exemplo 53;

[00773] xiv) o fusossoma é substancialmente livre de uma proteína da família Tetraspanina (por exemplo, CD63, CD9 ou CD81), uma proteína relacionada a ESCRT (por exemplo, TSG101, CHMP4A-B ou VPS4B), Alix, TSG101, MHCI, MHCII, GP96, actinina -4, mitofilina, sintenina-1, TSG101, ADAM10, EHD4, sintenina-1, TSG101, EHD1, flotilina-1, proteínas de choque térmico de 70 kDa (HSC70/HSP73, HSP70/HSP72) ou qualquer combinação dos mesmos ou contém menos de 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 5% ou 10% de qualquer proteína marcadora exossômica individual e/ou menos de 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20% ou 25% do total de proteínas marcadoras exossômicas de qualquer uma dentre as ditas proteínas ou é enriquecido com qualquer uma ou mais dentre essas proteínas em comparação à célula de origem ou não é enriquecido para qualquer uma ou mais dentre essas proteínas, por exemplo, por um ensaio do Exemplo 90;

[00774] xv) o fusossoma compreende um nível de Gliceraldeído 3-fosfato desidrogenase (GAPDH) que está abaixo de 500, 250, 100, 50, 20, 10, 5 ou 1 ng de proteína total de GAPDH/ug ou abaixo do nível de GAPDH na célula de origem, por exemplo, inferior a 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, inferior ao nível de GAPDH por total de proteína em ng/ug na célula de origem, por exemplo, com o uso do Exemplo 37;

[00775] xvi) o fusossoma é enriquecido para uma ou mais proteínas de reticulo endoplasmático (por exemplo, calnexina), uma ou mais proteínas proteassoma ou uma ou mais proteínas mitocondriais ou qualquer combinação das mesmas, por exemplo, em que a quantidade de

calnexina é inferior a 500, 250, 100, 50, 20, 10, 5 ou 1 ng de Calnexina/ug da proteína total, ou em que o fusossoma compreende menos Calnexina por proteína total em ng/ug em comparação à célula de origem por 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, por exemplo, com o uso de um ensaio do Exemplo 38 ou 91, ou em que o teor fracionário médio da Calnexina no fusossoma é inferior a aproximadamente $1x10^{-4}$, $1,5x10^{-4}$, $2x10^{-4}$, $2,1x10^{-4}$, $2,2x10^{-4}$, $2,3x10^{-4}$; $2,4x10^{-4}$, $2,43x10^{-4}$, $2,5x10^{-4}$, $2,6x10^{-4}$, $2,7x10^{-4}$, $2,8x10^{-4}$, $2,9x10^{-4}$, $3x10^{-4}$, $3,5x10^{-4}$ ou $4x10^{-4}$, ou em que o fusossoma compreende uma quantidade de Calnexina por proteína total é inferior àquele da célula parental em aproximadamente 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99% ou mais;

[00776] xvii) o fusossoma compreende um agente exógeno (por exemplo, uma proteína exógena, mRNA ou siRNA), por exemplo, conforme medido com o uso de um ensaio do Exemplo 35; ou

[00777] xviii) o fusossoma pode ser imobilizado em uma superfície de mica por microscopia de força atômica durante pelo menos 30 minutos, por exemplo, por um ensaio de Kanada M, et al. (2015). Os destinos diferenciais de biomoléculas entregues às células-alvo por meio de vesículas extracelulares. Proc Natl Acad Sci EUA 112:E1433 a E1442.

[00778] Nas modalidades, um ou mais dentre:

[00779] i) o fusossoma é um exossoma;

[00780] ii) o fusossoma não é uma microvesícula;

[00781] iii) o fusossoma tem um tamanho inferior a 80 nm, 100 nm, 200 nm, 500 nm, 1.000 nm, 1.200 nm, 1.400 nm ou 1.500 nm ou uma população de fusossomas tem um tamanho médio inferior a 80 nm, 100 nm, 200 nm, 500 nm, 1.000 nm, 1.200 nm, 1.400 nm ou 1.500 nm;

[00782] iv) o fusossoma não compreende uma organela;

[00783] v) o fusossoma não compreende um citoesqueleto ou um

componente do mesmo, por exemplo, actina, Arp2/3, formina, coronina, distrofina, queratina, miosina ou tubulina;

[00784] vi) o fusossoma, ou uma composição ou preparação que compreende uma pluralidade de fusossomas, tem densidade de flutuação de 1,08 a 1,22 g/ml, por exemplo, em um ensaio de centrifugação de gradiente de sacarose, por exemplo, conforme descrito em Théry et al., “Isolation and characterization of exosomes from cell culture supernatants and biological fluids”. Curr Protoc Cell Biol. Abril de 2006; Capítulo 3:Unidade 3.22;

[00785] vii) a bicamada lipídica não é enriquecida (por exemplo, está depletada) para ceramidas ou esfingomielinas ou uma combinação das mesmas em comparação à célula de origem, ou a bicamada lipídica é enriquecida para glicolipídios, ácidos graxos livres ou fosfatidilserina, ou uma combinação dos mesmos, em comparação à célula de origem;

[00786] viii) o fusossoma não compreende ou está depletado em relação à célula de origem, ligante de Fosfatidilserina (PS) ou de CD40 ou tanto ligante de PS quanto ligante de CD40, por exemplo, quando medido em um ensaio do Exemplo 93;

[00787] ix) o fusossoma não é enriquecido (por exemplo, está depletado) para PS em comparação à célula de origem, por exemplo, em uma população de fusossomas inferior a 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% são positivos para PS, por exemplo, por um ensaio de Kanada M, et al. (2015). Os destinos diferenciais de biomoléculas entregues às células-alvo por meio de vesículas extracelulares. Proc Natl Acad Sci EUA 112:E1433 a E1442;

[00788] x) o fusossoma compreende acetilcolinesterase (AChE), por exemplo, pelo menos 0,001, 0,002, 0,005, 0,01, 0,02, 0,05, 0,1, 0,2, 0,5, 1, 2, 5, 10, 20, 50, 100, 200, 500 ou 1.000 AChE unidades de atividade/ug de proteína, por exemplo, por um ensaio do Exemplo 53;

[00789] xi) o fusossoma compreende uma proteína da família Tetraspanina (por exemplo, CD63, CD9 ou CD81), uma proteína relacionada com ESCRT (por exemplo, TSG101, CHMP4A-B ou VPS4B), Alix, TSG101, MHCI, MHCII, GP96, actinina-4, mitofilina, sintenina-1, TSG101, ADAM10, EHD4, sintenina-1, TSG101, EHD1, flotilina-1, proteínas de choque térmico de 70 kDa (HSC70/HSP73, HSP70/HSP72) ou qualquer combinação das mesmas, por exemplo, contém mais do que 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 5% ou 10% de qualquer proteína marcadora exossômica individual e/ou menos do que 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20% ou 25% do total de proteínas marcadoras exossômicas de qualquer uma dentre as ditas proteínas, ou é enriquecido com qualquer uma ou mais dessas proteínas em comparação à célula de origem, por exemplo, por um ensaio do Exemplo 90;

[00790] xii) o fusossoma compreende um nível de Gliceraldeído 3-fosfato desidrogenase (GAPDH) que está acima de 500, 250, 100, 50, 20, 10, 5 ou 1 ng de GAPDH/ug da proteína total ou abaixo do nível de GAPDH na célula de origem, por exemplo, pelo menos 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, maior que o nível de GAPDH por proteína total em ng/ug na célula de origem, por exemplo, com o uso de um ensaio do Exemplo 37;

[00791] xiii) o fusossoma não é enriquecido (por exemplo, é depletado para) para uma ou mais proteínas de reticulo endoplasmático (por exemplo, calnexina), uma ou mais proteínas proteassoma ou uma ou mais proteínas mitocondriais ou qualquer combinação das mesmas, por exemplo, em que a quantidade de calnexina é inferior a 500, 250, 100, 50, 20, 10, 5 ou 1 ng de Calnexina/ug da proteína total, ou em que o fusossoma compreende menos Calnexina por proteína total em ng/ug em comparação à célula de origem por 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, por exemplo,

com o uso de um ensaio do Exemplo 91 ou 91, ou em que o teor fracionário médio da Calnexina no fusossoma é inferior a aproximadamente $1x10^{-4}$, $1,5x10^{-4}$, $2x10^{-4}$, $2,1x10^{-4}$, $2,2x10^{-4}$, $2,3x10^{-4}$; $2,4x10^{-4}$, $2,43x10^{-4}$, $2,5x10^{-4}$, $2,6x10^{-4}$, $2,7x10^{-4}$, $2,8x10^{-4}$, $2,9x10^{-4}$, $3x10^{-4}$, $3,5x10^{-4}$ ou $4x10^{-4}$, ou em que o fusossoma compreende uma quantidade de Calnexina por proteína total é inferior àquele da célula parental em aproximadamente 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99% ou mais; ou

[00792] xiv) o fusossoma não pode ser imobilizado em uma superfície de mica por microscopia de força atômica por pelo menos 30 min, por exemplo, por um ensaio de Kanada M, et al. (2015). Os destinos diferenciais de biomoléculas entregues às células-alvo por meio de vesículas extracelulares. Proc Natl Acad Sci EUA 112:E1433 a E1442.

[00793] Nas modalidades, um ou mais dentre:

[00794] i) o fusossoma não compreende um VLP;

[00795] ii) o fusossoma não contém um vírus;

[00796] iii) o fusossoma não compreende um vírus competente para replicação;

[00797] iv) o fusossoma não compreende uma proteína viral, por exemplo, uma proteína estrutural viral, por exemplo, uma proteína capsidial ou uma proteína da matriz viral;

[00798] v) o fusossoma não compreende uma proteína capsidial de um vírus com envelope;

[00799] vi) o fusossoma não compreende uma proteína do nucleocapsídeo; ou

[00800] vii) o fusogênio não é um fusogênio viral.

[00801] Em modalidades, o fusossoma compreende citosol.

[00802] Nas modalidades, um ou mais dentre:

[00803] i) o fusossoma ou a célula de origem não forma um teratoma quando implantado no indivíduo, por exemplo, por um ensaio do

Exemplo 66;

[00804] ii) o fusossoma tem capacidade para quimiotaxia, por exemplo, dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais do que uma célula de referência, por exemplo, um macrófago, por exemplo, com o uso de um ensaio do Exemplo 46;

[00805] iii) o fusossoma tem capacidade para se alojar, por exemplo, no sítio de uma lesão, em que o fusossoma ou citobiológico é de uma célula humana, por exemplo, com o uso de um ensaio do Exemplo 47, por exemplo, em que a célula de origem é um neutrófilo; ou

[00806] iv) o fusossoma tem capacidade para fagocitose, por exemplo, em que a fagocitose pelo fusossoma é detectável dentro de 0,5, 1, 2, 3, 4, 5 ou 6 horas com o uso de um ensaio do Exemplo 48, por exemplo, em que a célula de origem é um macrófago.

[00807] Em modalidades, o fusossoma ou composição de fusossoma retém um, dois, três, quatro, cinco, seis ou mais dentre qualquer uma das características por 5 dias ou menos, por exemplo, 4 dias ou menos, 3 dias ou menos, 2 dias ou menos, 1 dia ou menos, por exemplo, de aproximadamente 12 a 72 horas, após a administração a um indivíduo, por exemplo, um indivíduo humano.

[00808] Em modalidades, o fusossoma tem uma ou mais dentre as seguintes características:

[00809] a) compreende uma ou mais proteínas endógenas de uma célula de origem, por exemplo, proteínas de membrana ou proteínas citosólicas;

[00810] b) compreende pelo menos 10, 20, 50, 100, 200, 500, 1.000, 2000 ou 5.000 proteínas diferentes;

[00811] c) compreende pelo menos 1, 2, 5, 10, 20, 50 ou 100 glicoproteínas diferentes;

[00812] d) pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%

ou 90% em massa das proteínas no fusossoma são proteínas de ocorrência natural;

[00813] e) compreende pelo menos 10, 20, 50, 100, 200, 500, 1.000, 2.000 ou 5.000 RNAs diferentes; ou

[00814] f) compreende pelo menos 2, 3, 4, 5, 10 ou 20 lipídios diferentes, por exemplo, selecionados de CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM e TAG.

[00815] Em modalidades, o fusossoma foi manipulado para ter, ou o fusossoma não é uma célula de ocorrência natural e tem, ou em que o núcleo não tem naturalmente, uma, duas, três, quatro, cinco ou mais dentre seguintes propriedades:

[00816] a) a inativação nuclear parcial resulta em uma redução de pelo menos 50%, 60%, 70%, 80%, 90% ou mais na função nuclear, por exemplo, uma redução na transcrição ou replicação de DNA, ou os dois, por exemplo, em que a transcrição é medida por um ensaio do Exemplo 25 e a replicação do DNA é medida por um ensaio do Exemplo 26;

[00817] b) o fusossoma não tem capacidade para transcrição ou tem atividade transcricional inferior a 1%, 2,5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% daquela da atividade transcricional de uma célula de referência, por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 25;

[00818] c) o fusossoma não tem capacidade para replicação de DNA nuclear ou tem replicação de DNA nuclear inferior a 1%, 2,5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% da replicação do DNA nuclear de uma célula de referência, por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 26;

[00819] d) o fusossoma carece de cromatina ou tem um teor de cromatina inferior a 1%, 2,5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% do o teor de cromatina de uma célula de referência,

por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 33;

[00820] e) o fusossoma não tem uma membrana nuclear ou tem menos de 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2% ou 1% da quantidade de membrana nuclear de uma célula de referência, por exemplo, a célula de origem ou uma célula Jurkat, por exemplo, por um ensaio do Exemplo 32;

[00821] f) o fusossoma não tem complexos de poros nucleares funcionais ou tem atividade de importação ou exportação nuclear reduzida, por exemplo, em pelo menos 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2% ou 1 % por um ensaio do Exemplo 32, ou o fusossoma carece de uma ou mais de uma proteína de poro nuclear, por exemplo, NUP98 ou Importina 7;

[00822] g) o fusossoma não compreende histonas ou tem níveis de histonas inferiores a 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% do nível de histona da célula de origem (por exemplo, de H1, H2a, H2b, H3 ou H4), por exemplo, por um ensaio do Exemplo 33;

[00823] h) o fusossoma compreende menos de 20, 10, 5, 4, 3, 2 ou 1 cromossomo;

[00824] i) a função nuclear é eliminada;

[00825] j) o fusossoma é uma célula de mamífero enucleada;

[00826] k) o núcleo é removido ou inativado, por exemplo, expulso por força mecânica, por radiação ou por ablação química; ou

[00827] l) o fusossoma é de uma célula de mamífero com DNA que é completamente ou parcialmente removido, por exemplo, durante a interfase ou mitose.

[00828] Em modalidades, o fusossoma compreende mtDNA ou DNA vetor. Em modalidades, o fusossoma não compreende DNA.

[00829] Em modalidades, a célula de origem é uma célula primária,

célula imortalizada ou uma linhagem celular (por exemplo, linhagem celular mielobásticas, por exemplo, C2C12). Em modalidades, o fusossoma é de uma célula de origem com um genoma modificado, por exemplo, tendo imunogenicidade reduzida (por exemplo, por edição do genoma, por exemplo, para remover uma proteína de MHC ou complexos de MHC). Em modalidades, a célula de origem é de uma cultura de células tratada com um sinal anti-inflamatório. Nas modalidades, a célula de origem é de uma cultura de células tratada com um agente imunossupressor. Em modalidades, a célula de origem é substancialmente não imunogênica, por exemplo, com o uso de um ensaio descrito no presente documento. Em modalidades, a célula de origem compreende um agente exógeno, por exemplo, um agente terapêutico. Nas modalidades, a célula de origem é uma célula recombinante.

[00830] Em modalidades, o fusossoma compreende ainda um agente exógeno, por exemplo, um agente terapêutico, por exemplo, uma proteína ou um ácido nucleico (por exemplo, um DNA, um cromossomo (por exemplo, um cromossomo artificial humano), um RNA, por exemplo, um mRNA ou miRNA). Nas modalidades, o agente exógeno está presente pelo menos ou não mais que 10, 20, 50, 100, 200, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias, por exemplo, compreendidas pelo fusossoma ou está presente em um nível médio de pelo menos ou não mais que, 10, 20, 50, 100, 200, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000 ou 1.000.000 de cópias por fusossoma. Em modalidades, o fusossoma tem um nível alterado, por exemplo, aumentado ou diminuído de uma ou mais moléculas endógenas, por exemplo, proteína ou ácido nucleico, por exemplo, devido ao tratamento da célula de mamífero com um siRNA ou enzima de edição de gene. Em modalidades, a molécula endógena

está presente em, por exemplo, um nível médio, de pelo menos, ou não mais do que, 10, 20, 50, 100, 200, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias (por exemplo, cópias compreendidas pelo fusossoma), ou está presente em um nível médio de pelo menos, ou não mais que, 10, 20, 50, 100, 200, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000 ou 1.000.000 de cópias por fusossoma. Em modalidades, a molécula endógena (por exemplo, um RNA ou proteína) está presente em uma concentração de pelo menos 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3, 5,0 x 10^3, 10^4, 5,0 x 10^4, 10^5, 5,0 x 10^5, 10^6, 5,0 x 10^6, 1,0 x 10^7, 5,0 x 10^7 ou 1,0 x 10^8, maior do que sua concentração na célula de origem.

[00831] Nas modalidades, o agente ativo é selecionado a partir de uma proteína, complexo de proteína (por exemplo, compreendendo pelo menos 2, 3, 4, 5, 10, 20 ou 50 proteínas, por exemplo, pelo menos pelo menos 2, 3, 4, 5, 10, 20 ou 50 proteínas diferentes) polipeptídeo, ácido nucleico (por exemplo, DNA, cromossomo ou RNA, por exemplo, mRNA, siRNA ou miRNA) ou molécula pequena. Em modalidades, o agente exógeno compreende uma nuclease específica de local, por exemplo, molécula Cas9, TALEN ou ZFN.

[00832] Em modalidades, o fusogênio é um fusogênio viral, por exemplo, HA, HIV-1 ENV, HHV-4, gp120 ou VSV-G. Em modalidades, o fusogênio é um fusogênio de mamífero, por exemplo, um SNARE, um Sincitina, myomaker, myomixer, myomerger ou FGFRL1. Em modalidades, o fusogênio é ativo a um pH de 4 a 5, 5 a 6, 6 a 7, 7 a 8, 8 a 9 ou 9 a 10. Em modalidades, o fusogênio não é ativo a um pH de 4 a 5, 5 a 6, 6 a 7, 7 a 8, 8 a 9 ou 9 a 10. Em modalidades, o fusossoma é fundido a uma célula-alvo na superfície da célula-alvo. Em modalidades, o fusogênio promove a fusão de uma maneira independente do

lisossoma. Em modalidades, o fusogênio é um fusogênio de proteína. Em modalidades, o fusogênio é um fusogênio lipídico, por exemplo, ácido oleico, mono-oleato de glicerol, um glicerídeo, diacilglicerol ou um ácido graxo insaturado modificado. Em modalidades, o fusogênio é um fusogênio químico, por exemplo, PEG. Em modalidades, o fusogênio é uma pequena molécula de fusogênio, por exemplo, halotano, um NSAID, como meloxicam, piroxicam, tenoxicam e clorpromazina. Nas modalidades, o fusogênio é recombinante. Em modalidades, o fusogênio é bioquimicamente incorporado, por exemplo, o fusogênio é fornecido como uma proteína purificada e entra em contato com uma bicamada lipídica sob condições que permitam a associação do fusogênio com a bicamada lipídica. Em modalidades, o fusogênio é biossinteticamente incorporado, por exemplo, expresso em uma célula de origem sob condições que permitam que o fusogênio se associe à bicamada lipídica.

[00833] Em modalidades, o fusossoma se liga a uma célula-alvo. Em modalidades, a célula-alvo é diferente de uma célula HeLa, ou a célula-alvo não é transformada ou imortalizada.

[00834] Em algumas modalidades que envolvem composições de fusossomas, a pluralidade de fusossomas é a mesma. Em algumas modalidades, a pluralidade de fusossomas é diferente. Em algumas modalidades, a pluralidade de fusossomas são de uma ou mais células de origem. Em algumas modalidades, pelo menos 50%, 60%, 70%, 80%, 90%, 95% ou 99% dos fusossomas na pluralidade têm um diâmetro dentro de 10%, 20%, 30%, 40% ou 50% do diâmetro médio dos fusossomas na composição dos fusossomas. Em algumas modalidades, pelo menos 50%, 60%, 70%, 80%, 90%, 95% ou 99% dos fusossomas na pluralidade têm um volume dentro de 10%, 20%, 30%, 40% ou 50% do volume médio dos fusossomas na composição dos fusossomas. Em algumas modalidades, a composição do fusossoma tem

menos de aproximadamente 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, variabilidade na distribuição de tamanho dentro de 10%, 50 % ou 90% da variabilidade da população de células de origem na distribuição de tamanho, por exemplo, com base em Exemplo 29. Em algumas modalidades, pelo menos 50%, 60%, 70%, 80%, 90%, 95% ou 99% dos fusossomas na pluralidade têm um número de cópias do fusogênio dentro de 10%, 20%, 30%, 40 %, 50%, 60%, 70%, 80% ou 90% do número médio de cópias fusogênicas nos fusossomas na composição de fusossomas. Em algumas modalidades, pelo menos 50%, 60%, 70%, 80%, 90%, 95% ou 99% dos fusossomas na pluralidade têm um número de cópias do agente terapêutico dentro de 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% do número médio de cópias do agente terapêutico nos fusossomas na composição dos fusossomas. Em algumas modalidades, a composição de fusossoma compreende pelo menos 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14} ou 10^{15} ou mais fusossomas. Em algumas modalidades, a composição de fusossoma está em um volume de pelo menos 1 ul, 2 ul, 5 ul, 10 ul, 20 ul, 50 ul, 100 ul, 200 ul, 500 ul, 1 ml, 2 ml, 5 ml ou 10 ml.

[00835] Em algumas modalidades, a composição do fusossoma entrega a carga a pelo menos 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% do número de células na população de células-alvo em comparação à população de células-alvo de referência.

[00836] Em algumas modalidades, a composição de fusossoma fornece pelo menos 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% da carga para a população de células-alvo em comparação à população de células-alvo de referência ou para uma população de células não alvo. Em algumas modalidades, a composição de fusossoma fornece pelo menos 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%,

97%, 98% ou 99% ou mais da carga para a população de células-alvo em comparação à população de células-alvo de referência ou para uma população de células não alvo.

[00837] Em algumas modalidades, menos de 10% da carga entra na célula por endocitose.

[00838] Em algumas modalidades, o inibidor da endocitose é um inibidor da acidificação lisossomal, por exemplo, bafilomicina A1. Em algumas modalidades, o inibidor de endocitose é um inibidor de dinamina, por exemplo, Dynasore.

[00839] Em algumas modalidades, a população de células-alvo está em um pH fisiológico (por exemplo, entre 7,3 a 7,5, por exemplo, entre 7,38 a 7,42).

[00840] Em algumas modalidades, a carga entregue é determinada com o uso de um ensaio de inibição de endocitose, por exemplo, um ensaio do Exemplo 81.

[00841] Em algumas modalidades, a carga entra na célula através de uma via independente de dinamina ou uma via independente de acidificação lisossomal, uma via independente de macropinocitose (por exemplo, em que o inibidor de endocitose é um inibidor de macropinocitose, por exemplo, 5- (N-etil-N-isopropil) amilorida (EIPA), por exemplo, a uma concentração de 25 μM), ou uma via independente de actina (por exemplo, em que o inibidor da endocitose é um inibidor da polimerização da actina é, por exemplo, Latrunculina B, por exemplo, a uma concentração de 6 μM).

[00842] Em algumas modalidades, os fusossomas da pluralidade compreendem adicionalmente uma porção química de alvejamento. Em modalidades, a porção de alvejamento é composta pelo fusogênio ou é composta por uma molécula separada.

[00843] Em algumas modalidades, quando a pluralidade de fusossomas são contatados com uma população de células que compreen-

dem células-alvo e células não alvo, a carga está presente em pelo menos 10 vezes mais em células-alvo do que em células não alvo.

[00844] Em algumas modalidades, quando a pluralidade de fusossomas é colocada em contato com uma população de células que compreende células-alvo e células não alvo, a carga está presente pelo menos 2 vezes, 5 vezes, 10 vezes, 20 vezes ou 50 vezes mais alto em células-alvo do que células não alvo e/ou a carga está presente pelo menos 2 vezes, 5 vezes, 10 vezes, 20 vezes ou 50 vezes mais em células-alvo do que células de referência.

[00845] Em algumas modalidades, os fusossomas da pluralidade é fundidom a uma taxa mais alta com uma célula-alvo do que com uma célula não alvo em pelo menos 50%.

[00846] Em algumas modalidades, o fusossoma, quando em contato com uma população de células-alvo, entrega carga a um sítio de célula-alvo diferente de um endossomo ou lisossoma, por exemplo, para o citosol. Nas modalidades, menos 50%, 40%, 30%, 20% ou 10% da carga é entregue a um endossomo ou lisossoma.

[00847] Em algumas modalidades, os fusossomas da pluralidade compreendem exossomas, microvesículas ou uma combinação dos mesmos.

[00848] Em algumas modalidades, a pluralidade de fusossomas tem um tamanho médio de pelo menos 50 nm, 100 nm, 200 nm, 500 nm, 1.000 nm, 1.200 nm, 1.400 nm ou 1.500 nm. Em outras modalidades, a pluralidade de fusossomas tem um tamanho médio inferior a 100 nm, 80 nm, 60 nm, 40 nm ou 30 nm.

[00849] Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) compreende um fusogênio de mamífero. Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) compreende um fusogênio viral. Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) é uma proteína fusogênica.

Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) compreende uma sequência escolhida a partir de uma proteína F do vírus Nipah, uma proteína F do vírus do sarampo, uma proteína F do paramixovírus da tupaia, uma proteína F do paramixovírus, uma proteína F do vírus Hendra, uma proteína F do Henipavírus, uma proteína F do Morbilivírus, uma proteína F do respirovírus, uma vírus proteína F doi Sendai, uma proteína F do rubulavírus ou uma proteína F do avulavírus ou um derivado dos mesmos.

[00850] Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) é ativo a um pH de 4 a 5, 5 a 6, 6 a 7, 7 a 8, 8 a 9 ou 9 a 10. Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) não é ativo a um pH de 4 a 5, 5 a 6, 6 a 7, 7 a 8, 8 a 9 ou 9 a 10.

[00851] Em algumas modalidades, o fusogênio está presente em um número de cópias de pelo menos 1, 2, 5 ou 10 cópias por fusossoma.

[00852] Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) compreende uma proteína G do vírus Nipah, uma proteína H do sarampo, uma proteína H do paramixovírus de tupaia, uma proteína G do paramixovírus, uma proteína H do paramixovírus, uma proteína HN do paramixovírus, um Morbilivírus H proteína, uma proteína HN de respirovírus, uma proteína HN sendai, uma proteína HN de rubulavírus, uma proteína HN de avulavírus ou um seu derivado. Em algumas modalidades, o fusogênio (por exemplo, fusogênio realvejado) compreende uma sequência escolhida dentre a proteínas F e G do vírus do sarampo, proteínas F e H do vírus do sarampo, proteínas F e H do paramixovírus da tupaia, proteínas F e G ou proteínas F e H ou proteínas F e HN do paramixovírus, proteínas F e G do vírus Hendra, proteínas F e G do Henipavírus, proteínas F e H do Morbilivírus, protéina F e HN do respirovírus, uma proteína F e HN do vírus

Sendai, proteínas F e HN do rubulavírus ou proteínas F e HN do avulavírus ou um derivado das mesmas ou qualquer combinação das mesmas.

[00853] Em algumas modalidades, a carga compreende uma proteína exógena ou um ácido nucleico exógeno. Em algumas modalidades, a carga compreende ou codifica uma proteína citosólica. Em algumas modalidades, a carga compreende ou codifica uma proteína de membrana. Em algumas modalidades, a carga compreende um agente terapêutico. Em algumas modalidades, a carga está presente em um número de cópias de pelo menos 1, 2, 5, 10, 20, 50, 100, ou 200 cópias por fusossoma (por exemplo, até aproximadamente 1.000 cópias por fusossoma). Em algumas modalidades, a razão entre o número de cópias do fusogênio (por exemplo, fusogênio realvejado) e o número de cópia da carga está entre 1.000:1 e 1:1 ou entre 500:1 e 1:1 ou entre 250:1 e 1:1 ou entre 150:1 e 1:1 ou entre 100:1 e 1:1 ou entre 75:1 e 1:1 ou entre 50:1 e 1:1 ou entre 25:1 e 1:1 ou entre 20:1 e 1:1 ou entre 15:1 e 1:1 ou entre 10:1 e 1:1 ou entre 5:1 e 1:1 ou entre 2:1 e 1:1 ou entre 1:1 e 1:2.

[00854] Em algumas modalidades, a composição de fusossoma compreende uma proteína capsidial viral ou um polipeptídeo de integração de DNA. Em algumas modalidades, a carga compreende um genoma viral.

[00855] Em algumas modalidades, a composição do fusossoma tem capacidade para entregar um ácido nucleico a uma célula-alvo, por exemplo, para modificar de forma estável o genoma da célula-alvo, por exemplo, para terapia genética.

[00856] Em algumas modalidades, a composição de fusossoma não compreende uma proteína de nucleocapsídeo viral ou a quantidade de proteína de nucleocapsídeo viral é inferior a 10%, 5%, 4%, 3%, 2%, 1%, 0,5%, 0,2% ou 0,1% da proteína total, por exemplo, por espectro-

metria de massa, por exemplo, com o uso de um ensaio do Exemplo 94.

[00857] Em modalidades, a composição de fusossoma compreende pelo menos 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14} ou 10^{15} fusossomas. Em modalidades, a composição de fusossoma compreende pelo menos 10 ml, 20 ml, 50 ml, 100 ml, 200 ml, 500 ml, 1 l, 2 l, 5 l, 10 l, 20 l ou 50 l.

[00858] Em modalidades, o fusossoma é de uma célula de mamífero com um genoma modificado, por exemplo, para reduzir a imunogenicidade (por exemplo, por edição do genoma, por exemplo, para remover uma proteína MHC ou complexos MHC). Em modalidades, a célula de origem é de uma cultura de células tratada com um sinal anti-inflamatório. Em modalidades, o método compreende adicionalmente o contato da célula de origem da etapa a) com um agente imunossupressor ou sinal anti-inflamatório, por exemplo, antes ou depois de inativar o núcleo, por exemplo, enuclear a célula.

[00859] Em um aspecto, é fornecido no presente documento uma composição de fusossomas que compreende uma pluralidade de fusossomas derivados de uma célula de origem, em que os fusossomas da pluralidade compreendem: (a) uma bicamada lipídica, (b) um lúmen que compreende citosol, em que o lúmen é cercado pela bicamada lipídica; (c) um fusogênio exógeno ou superexpresso disposto na bicamada lipídica; (d) uma carga; e em que o fusossoma não compreende um núcleo; em que a quantidade de proteína capsidial viral na composição de fusossoma é inferior a 1% da proteína total; em que a pluralidade de fusossomas, quando em contato com uma população de células-alvo na presença de um inibidor de endocitose e quando em contato com uma população de células-alvo de referência não tratada com o inibidor de endocitose, entrega a carga a pelo menos 30% do número de células na população de células-alvo em comparação à

população de células-alvo de referência.

[00860] Em modalidades, a composição de fusossoma entrega a carga a pelo menos 40%, 50%, 60%, 70% ou 80% do número de células na população de células-alvo em comparação à população de células-alvo de referência ou a uma população de células não alvo; ou entrega a carga, por exemplo, pelo menos 40%, 50%, 60%, 70% ou 80% da carga, para a população de células-alvo em comparação à população de células-alvo de referência ou para uma população de células não alvo. Em modalidades, menos de 10% da carga entra na célula por endocitose. Em modalidades, o inibidor da endocitose é um inibidor da acidificação lisossomal, por exemplo, bafilomicina A1. Em modalidades, a carga entregue é determinada com o uso de um ensaio de inibição de endocitose, por exemplo, um ensaio do Exemplo 81. Em modalidades, a carga entra na célula através de uma via independente de dinamina ou uma via independente de acidificação lisossomal, uma via independente de macropinocitose (por exemplo, em que o inibidor de endocitose é um inibidor de macropinocitose, por exemplo, 5-(N-etil-N-isopropil)amilorida (EIPA), por exemplo, a uma concentração de 25 μM), ou uma via independente de actina (por exemplo, em que o inibidor da endocitose é um inibidor da polimerização da actina é, por exemplo, Latrunculina B, por exemplo, a uma concentração de 6 μM).

[00861] As composições de fusossomas podem ser geradas a partir de células em cultura, por exemplo, células de mamíferos em cultura, por exemplo, células humanas em cultura. As células podem ser células progenitoras ou células não progenitoras (por exemplo, diferenciadas). As células podem ser células primárias ou linhagens celulares (por exemplo, um mamífero, por exemplo, humano, linhagem celular descrita no presente documento). Em modalidades, as células cultivadas são células progenitoras, por exemplo, células do estroma da medula óssea, células progenitoras adultas derivadas da medula (MA-

PCs), células progenitoras endoteliais (EPC), células blásticas, células progenitoras intermediárias formadas na zona subventricular, células-tronco neurais, músculo células-tronco, células-satélite, células-tronco do fígado, células-tronco hematopoiéticas, células do estroma da medula óssea, células-tronco epidérmicas, células-tronco embrionárias, células-tronco mesenquimais, células-tronco do cordão umbilical, células precursoras, células precursoras do músculo, mioblasto, cardiomioblasto, células precursoras neurais, células precursoras gliais, células precursoras neuronais, hepatoblastos.

[00862] Em algumas modalidades, a célula de origem é uma célula endotelial, um fibroblasto, uma célula sanguínea (por exemplo, um macrófago, um neutrófilo, um granulócito, um leucócito), uma célula-tronco (por exemplo, uma célula-tronco mensequimal, uma célula-tronco do cordão umbilical, célula-tronco da medula óssea, uma célula-tronco hematopoiética, uma célula-tronco pluripotente induzida por exemplo, uma célula-tronco pluripotente induzida derivada de uma célula do indivíduo), uma célula-tronco embrionária (por exemplo, uma célula-tronco de vesícula vitelina embrionária, placenta, cordão umbilical, pele fetal, pele adolescente, sangue, medula óssea, tecido de adipose, tecido eritropoiético, tecido hematopoiético), um mioblasto, uma célula parenquimal (por exemplo, hepatócito), uma célula alveolar, um neurônio (por exemplo, uma célula neuronal retiniana) uma célula precurosora (por exemplo, uma célula precursora retiniana, um mieloblasto, células precursoras mieloides, um timócito, um meiócito, um megacarioblasto, um promegacarioblasto, um melanoblasto, um linfoblasto, uma célula precursora da medula óssea, um normoblasto ou um angioblasto), uma célula progenitora (por exemplo, uma célula progenitora pancreática, uma célula satélite, uma célula da glia radial, uma célula estromal da medula óssea, uma célula progtenitora pancreática, uma célula progenitora endotelial, uma célula do blasto) ou uma célula

imortalizada (por exemplo, célula HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080 ou BJ).

[00863] As células cultivadas podem ser de tecido ou células epiteliais, conectivas, musculares ou nervosas e suas combinações. O fussossoma pode ser gerado a partir de células cultivadas de qualquer sistema de órgãos eucariótico (por exemplo, mamífero), por exemplo, do sistema cardiovascular (coração, vasculatura); sistema digestivo (esofago, estômago, fígado, vesícula biliar, pâncreas, intestinos, cólon, reto e ânus); sistema edócrino (hipotálamos, glândula pituitária, corpo pineal ou glândula pineal, tireoide, paratireoide, glândulas adrenais); sistema excretor (rins, uretras, bexiga); sistema linfático (linfo, linfonodos, vasos linfáticos, tonsilas, adenoides, timo, baço); sistema tegumentar (pele, cabelo, unhas); sistema muscular (por exemplo, músculo esquelético); sistema nervoso (cérebro, medula espinhal, nervos); sistema reprodutivo (ovários, útero, glândulas mamárias, testículos, vasos deferentes, vesículas seminais, próstata); sistema respiratório (faringe, laringe, traqueia, brônquios, pulmões, diafragma); sistema esquelético (osso, cartilagem) e combinações dos mesmos. Nas modalidades, as células são de um tecido altamente mitótico (por exemplo, um tecido saudável altamente mitótico, como epitélio, tecido embrionário, medula óssea, criptas intestinais). Em modalidades, a amostra de tecido é um tecido altamente metabólico (por exemplo, tecido esquelético, tecido neural, cardiomiócitos).

[00864] Em algumas modalidades, as células são de um doador jovem, por exemplo, um doador de 25 anos, 20 anos, 18 anos, 16 anos, 12 anos, 10 anos, 8 anos de idade, 5 anos de idade, 1 ano de idade ou mais jovem. Em algumas modalidades, as células são de tecido fetal.

[00865] Em algumas modalidades, as células são derivadas de um indivíduo e administradas ao mesmo indivíduo ou a um indivíduo com

uma assinatura genética semelhante (por exemplo, compatível com MHC).

[00866] Em determinadas modalidades, as células têm telômeros de tamanho médio maior que 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000 ou 10.000 nucleotídeos de comprimento (por exemplo, entre 4.000 a 10.000 nucleotídeos de comprimento, entre 6.000 e 10.000 nucleotídeos de comprimento).

AVALIAÇÃO DO TEOR DE FUSOSSOMA DA CÉLULA-ALVO

[00867] A presente divulgação também fornece, em alguns aspectos, um método de avaliação do conteúdo de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) em um indivíduo que compreende fornecer uma amostra biológica de um indivíduo que recebeu uma composição de fusossoma (por exemplo, uma composição de fusossoma descrita no presente documento) e realizar um ensaio para determinar uma ou mais propriedades da amostra biológica resultante da fusão de uma célula-alvo na amostra biológica com um fusossoma, conforme descrito no presente documento. Em alguns aspectos, a divulgação fornece um método de medição de fusão com uma célula-alvo, por exemplo, conforme descrito em Exemplo 72. Em algumas modalidades, a determinação de uma ou mais propriedades da amostra biológica compreende determinar: a presença de um fusogênio, o nível de uma carga ou carga útil ou uma atividade relacionada a uma carga ou carga útil.

[00868] Em alguns aspectos, a presente divulgação fornece um método de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) em um indivíduo que compreende fornecer uma amostra biológica de um indivíduo que recebeu uma composição de fusossoma, por exemplo, conforme descrito no presente documento e testar a amostra biológica quanto à presença de um fusogênio, por exemplo, um fusogênio descrito no

presente documento. Em alguns casos, o nível de fusogênio detectado é maior (por exemplo, pelo menos de aproximadamente 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%, 5000%, 10.000%, 50.000% ou 100.000% maior) que aquele observado em uma amostra biológica correspondente de um indivíduo que não recebeu uma composição fusossoma. Em algumas modalidades, o indivíduo é o mesmo indivíduo antes da administração da composição de fusossoma e, em algumas modalidades, o indivíduo é um indivíduo diferente.

[00869] Em alguns aspectos, a presente divulgação fornece um método de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) em um indivíduo que compreende fornecer uma amostra biológica de um indivíduo que recebeu uma composição de fusossoma, por exemplo, conforme descrito no presente documento, e testar a amostra biológica quanto à presença de uma carga ou carga útil, por exemplo, entregue por um fusossoma, conforme descrito no presente documento. Em alguns casos, o nível da carga ou carga útil detectada é maior (por exemplo, pelo menos de aproximadamente 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%, 5000%, 10.000%, 50.000% ou 100.000% maior) do que o observado em uma amostra biológica correspondente de um indivíduo que não recebeu uma composição de fusossoma. Em algumas modalidades, o indivíduo é o mesmo indivíduo antes da administração da composição de fusossoma e, em algumas modalidades, o indivíduo é um indivíduo diferente.

[00870] Em alguns aspectos, a presente divulgação fornece um método de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo em um indivíduo) que

compreende fornecer uma amostra biológica de um indivíduo que recebeu uma composição de fusossoma, por exemplo, conforme descrito no presente documento e testar a amostra biológica para alteração de uma atividade relacionada à composição de fusossoma, por exemplo, uma atividade relacionada a uma carga ou carga útil entregue pela composição de fusossoma. Em alguns casos, o nível de atividade detectado é aumentado, por exemplo, em pelo menos de aproximadamente 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100 %, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1.000%, 2.000%, 3.000%, 4.000%, 5.000%, 10.000%, 50.000% ou 100.000%, em relação àquela de uma amostra biológica correspondente de um indivíduo que não recebeu uma composição de fusossoma (por exemplo, o mesmo indivíduo antes da administração da composição de fusossoma). Em alguns casos, o nível de atividade detectado é diminuído, por exemplo, em pelo menos de aproximadamente 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100 %, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1.000%, 2.000%, 3.000%, 4.000%, 5.000%, 10.000%, 50.000% ou 100.000%, em relação àquela de uma amostra biológica correspondente de um indivíduo que não recebeu uma composição de fusossoma. Em algumas modalidades, o indivíduo é o mesmo indivíduo antes da administração da composição de fusossoma e, em algumas modalidades, o indivíduo é um indivíduo diferente.

[00871] Em um aspecto, a presente divulgação fornece um método para avaliar a fusão de fusossomas a uma célula-alvo em um indivíduo que compreende fornecer uma amostra biológica de um indivíduo que recebeu uma composição de fusossoma, por exemplo, conforme descrito no presente documento, e avaliar um nível de fusossomas não fundidos na amostra biológica.

[00872] Em algumas modalidades dos métodos para avaliar o teor

de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) que resultam na formação de uma célula receptora no indivíduo, o método compreende adicionalmente coletar a amostra biológica do indivíduo. Em modalidades, a amostra biológica inclui uma ou mais células receptoras.

[00873] Em algumas modalidades dos métodos de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) no indivíduo, o método compreende adicionalmente separar as células receptoras na amostra biológica de fusossomas não fundidos na amostra biológica, por exemplo, por centrifugação. Em algumas modalidades, o método compreende adicionalmente o enriquecimento de células receptoras em relação aos fusossomas não fundidos na amostra biológica, por exemplo, por centrifugação. Em algumas modalidades, o método compreende adicionalmente enriquecer células-alvo em relação às células não alvo na amostra biológica, por exemplo, por FACS.

[00874] Em algumas modalidades dos métodos de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) em um indivíduo, a atividade relativa à composição de fusossoma é escolhida a partir da presença ou nível de um metabólito, a presença ou nível de um biomarcador (por exemplo, um nível de proteína ou modificação pós-tradução, por exemplo, fosforilação ou clivagem).

[00875] Em algumas modalidades dos métodos de avaliação do teor de fusossoma de uma célula-alvo (por exemplo, fusão de fusossoma a uma célula-alvo) em um indivíduo, a atividade relativa à composição de fusossoma é imunogenicidade. Em modalidades, a célula-alvo é uma célula CD3+ e a amostra biológica é uma amostra de sangue coletada do indivíduo. Nas modalidades, as células sanguíneas são enriquecidas a partir da amostra de sangue, por exemplo, com o

uso de uma solução tamponada de cloreto de amônio. Em modalidades, células sanguíneas enriquecidas são incubadas com um anticorpo anti-CD3 (por exemplo, um anticorpo anti-CD3-FITC de murganho) e as células CD3+ são selecionadas, por exemplo, por classificação de células ativadas por fluorescência. Nas modalidades, células, por exemplo, células classificadas, por exemplo, células CD3+ são analisadas quanto à presença de anticorpos na superfície celular, por exemplo, por meio da marcação com um anticorpo anti-IgM. Em algumas modalidades, caso os anticorpos estejam presentes em um nível acima de um nível de referência, o indivíduo é identificado como tendo uma resposta imunológica contra células receptoras.

[00876] Em modalidades, a imunogenicidade é avaliada por um ensaio de lise celular. Em modalidades, as células receptoras da amostra biológica são coincubadas com células efetoras imunes com capacidade para lisar outras células. Em modalidades, as células efetoras imunes são do indivíduo ou de um indivíduo não administrado com a composição de fusossoma. Por exemplo, nas modalidades, imunogenicidade é avaliada por um ensaio de lise celular de PBMC. Em modalidades, as células receptoras da amostra biológica são coincubadas com células mononucleares de sangue periférico (PBMCs) do indivíduo ou PBMCs de controle de um indivíduo não administrado a composição de fusossoma e, em seguida, avaliadas quanto à lise das células receptoras por PBMCs. Nas modalidades, a imunogenicidade é avaliada por um ensaio de lise de célula exterminadora (NK). Em modalidades, as células receptoras são coincubadas com células NK do indivíduo ou células NK de controle de um indivíduo não administrado a composição de fusossoma e, em seguida, avaliadas quanto à lise das células receptoras pelas células NK. Nas modalidades, a imunogenicidade é avaliada por um ensaio de lise de células T CD8+. Em modalidades, as células receptoras são coincubadas com células T

CD8+ do indivíduo ou células T CD8+ de controle de um indivíduo não administrado a composição de fusossoma e, então, avaliadas quanto à lise das células-alvo pelas células T CD8+. Em algumas modalidades, caso a lise celular ocorra em um nível acima de um nível de referência, o indivíduo é identificado como tendo uma resposta imunológica contra células receptoras.

[00877] Em algumas modalidades, a imunogenicidade é avaliada por fagocitose de células receptoras, por exemplo, por macrófagos. Em modalidades, as células receptoras não são alvejadas por macrófagos para fagocitose. Nas modalidades, a amostra biológica é uma amostra de sangue coletada do indivíduo. Nas modalidades, as células sanguíneas são enriquecidas a partir da amostra de sangue, por exemplo, com o uso de uma solução tamponada de cloreto de amônio. Em modalidades, células sanguíneas enriquecidas são incubadas com um anticorpo anti-CD3 (por exemplo, um anticorpo anti-CD3-FITC de murganho) e as células CD3+ são selecionadas, por exemplo, por classificação de células ativadas por fluorescência. Em modalidades, as células CD3+ identificadas com fluorescência são incubadas com macrófagos e, em seguida, testadas quanto à fluorescência intracelular dentro dos macrófagos, por exemplo, por citometria de fluxo. Em algumas modalidades, caso a fagocitose de macrófagos ocorra em um nível acima de um nível de referência, o indivíduo é identificado como tendo uma resposta imunológica contra células receptoras.

CARACTERÍSTICAS FÍSICAS E FUNCIONAIS DE FUSOSSOMAS

[00878] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um agente, por exemplo, uma proteína, ácido nucleico (por exemplo, mRNA), organela ou metabólito para o citosol de uma célula-alvo. De modo semelhante, em algumas modalidades, um método no presente documento compreende entregar um agente ao citosol de uma célula-alvo. Em algumas modalida-

des, o agente é uma proteína (ou um ácido nucleico que codifica a proteína, por exemplo, um mRNA que codifica a proteína) que está ausente, mutante ou em um nível inferior ao que o tipo selvagem na célula-alvo. Em algumas modalidades, a célula-alvo é de um indivíduo com uma doença genética, por exemplo, uma doença monogênica, por exemplo, uma doença de proteína intracelular monogênica. Em algumas modalidades, o agente compreende um fator de transcrição, por exemplo, um fator de transcrição exógeno ou um fator de transcrição endógeno. Em algumas modalidades, o fusossoma compreende adicionalmente um ou mais (por exemplo, pelo menos 2, 3, 4, 5, 10, 20 ou 50) fatores de transcrição adicionais ou o método compreende adicionalmente entregar os mesmos, por exemplo, fatores de transcrição exógenos, fatores endógenos de transcrição ou uma combinação dos mesmos.

[00879] Em algumas modalidades, o fusossoma compreende (por exemplo, tem capacidade para fornecer à célula-alvo) uma pluralidade de agentes (por exemplo, pelo menos 2, 3, 4, 5, 10, 20 ou 50 agentes), em que cada agente da pluralidade atua em uma etapa de uma via na célula-alvo, por exemplo, em que a via é uma via biossintética, uma via catabólica ou uma cascata de transdução de sinal. Em modalidades, cada agente na pluralidade regula positivamente a via ou regula negativamente a via. Em algumas modalidades, o fusossoma compreende adicionalmente ou o método compreende adicionalmente entregar um ou mais agentes adicionais (por exemplo, compreende uma segunda pluralidade de agentes) que não atuam em uma etapa da via, por exemplo, que atuam em uma etapa de um segundo caminho. Em algumas modalidades, o fusossoma compreende (por exemplo, tem capacidade para entregar à célula-alvo) ou o método compreende adicionalmente entregar uma pluralidade de agentes (por exemplo, pelo menos 2, 3, 4, 5, 10, 20 ou 50 agentes), em que cada agente da plura-

lidade é parte de uma única via, por exemplo, em que a via é uma via biossintética, uma via catabólica ou uma cascata de transdução de sinal. Em algumas modalidades, o fusossoma compreende adicionalmente ou o método compreende adicionalmente entregar um ou mais agentes adicionais (por exemplo, compreende uma segunda pluralidade de agentes) que não fazem parte da via única, por exemplo, fazem parte de uma segunda via.

[00880] Em algumas modalidades, a célula-alvo compreende uma proteína agregada ou mal dobrada. Em algumas modalidades, o fusossoma tem capacidade para reduzir os níveis (por exemplo, reduz os níveis) da proteína agregada ou mal dobrada na célula-alvo ou um método no presente documento compreende reduzir níveis da proteína agregada ou mal dobrada na célula-alvo.

[00881] Em algumas modalidades, o agente é selecionado a partir de um fator de transcrição, enzima (por exemplo, enzima nuclear ou enzima citosólica), reagente que medeia uma modificação específica de sequência de DNA (por exemplo, Cas9, ZFN ou TALEN), mRNA (por exemplo, codificação de mRNA uma proteína intracelular), organela ou metabólito.

[00882] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um agente, por exemplo, uma proteína, para a membrana celular de uma célula-alvo. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar um agente à membrana celular de uma célula-alvo. Em algumas modalidades, a entrega da proteína compreende entregar um ácido nucleico (por exemplo, mRNA) que codifica a proteína para a célula-alvo de modo que a célula-alvo produza a proteína e a localize na membrana. Em algumas modalidades, o fusossoma compreende ou o método compreende adicionalmente entregar a proteína e a fusão do fusossoma com a célula-alvo transfere a proteína

para a membrana celular da célula-alvo. Em algumas modalidades, o agente compreende um ligante da superfície celular ou um anticorpo que se liga a um receptor da superfície celular. Em algumas modalidades, o fusossoma compreende adicionalmente ou o método compreende adicionalmente entregar um segundo agente que compreende ou codifica um segundo ligante de superfície celular ou anticorpo que se liga a um receptor de superfície celular e, opcionalmente, compreende adicionalmente ou codifica um ou mais ligantes de superfície celular adicionais ou anticorpos que se ligam a um receptor de superfície celular (por exemplo, 1, 2, 3, 4, 5, 10, 20, 50 ou mais). Em algumas modalidades, o primeiro agente e o segundo agente formam um complexo, em que, opcionalmente, o complexo compreende adicionalmente um ou mais ligantes adicionais da superfície celular. Em algumas modalidades, o agente compreende ou codifica um receptor de superfície celular, por exemplo, um receptor de superfície celular exógeno. Em algumas modalidades, o fusossoma compreende adicionalmente ou o método compreende adicionalmente entregar de um segundo agente que compreende ou codifica um segundo receptor de superfície celular e, opcionalmente, compreende ou codifica um ou mais receptores de superfície celular adicionais (por exemplo, 1, 2, 3, 4, 5, 10, 20, 50 ou mais receptores de superfície celular).

[00883] Em algumas modalidades, o primeiro agente e o segundo agente formam um complexo, em que, opcionalmente, o complexo compreende adicionalmente um ou mais receptores adicionais da superfície celular. Em algumas modalidades, o agente compreende ou codifica um antígeno ou uma proteína apresentadora de antígeno.

[00884] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um agente secretado, por exemplo, uma proteína secretada para um local-alvo (por exemplo, uma região extracelular), por exemplo, pela entrega de um ácido nucleico (por

exemplo, mRNA) que codifica a proteína para a célula-alvo sob condições que permitam que a célula-alvo produza e segregue a proteína. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar um agente secretado como descrito no presente documento. Em modalidades, a proteína secretada é endógena ou exógena. Em modalidades, a proteína secretada compreende uma proteína terapêutica, por exemplo, uma molécula de anticorpo, uma citocina ou uma enzima. Em modalidades, a proteína secretada compreende uma molécula de sinalização autócrina ou uma molécula de sinalização parácrina. Em modalidades, o agente secretado compreende um grânulo secretor.

[00885] Em algumas modalidades, o fusossoma tem capacidade para reprogramar (por exemplo, reprogramar) uma célula-alvo (por exemplo, uma célula imune), por exemplo, pela entrega de um agente selecionado a partir de um fator de transcrição ou mRNA ou uma pluralidade dos referidos agentes. Da mesma forma, em algumas modalidades, um método no presente documento compreende reprogramar uma célula-alvo. Em modalidades, a reprogramação compreende induzir uma célula endócrina pancreática a assumir uma ou mais características de uma célula beta pancreática, induzindo-se um neurônio não dopaminérgico a assumir uma ou mais características de um neurônio dopaminérgico ou induzindo-se uma célula T exaurida a assumir uma ou mais características de uma célula T não exaurida, por exemplo, uma célula T assassina. Em algumas modalidades, o agente compreende um antígeno. Em algumas modalidades, o fusossoma compreende um primeiro agente que compreende um antígeno e um segundo agente que compreende uma proteína de apresentação de antígeno.

[00886] Em algumas modalidades, o fusossoma tem capacidade para doar (por exemplo, doar) um ou mais receptores de superfície

celular para uma célula-alvo (por exemplo, uma célula imune). Da mesma forma, em algumas modalidades, um método no presente documento compreende doar um ou mais receptores de superfície celular.

[00887] Em algumas modalidades, um fusossoma tem capacidade para modificar, por exemplo, modificar, uma célula tumoral alvo. Da mesma forma, em algumas modalidades, um método no presente documento compreende a modificação de uma célula tumoral alvo. Em modalidades, o fusossoma compreende um mRNA que codifica um ligante imunoestimulador, uma proteína de apresentação de antígeno, uma proteína supressora de tumor ou uma proteína pró-apoptótica. Em algumas modalidades, o fusossoma compreende um miRNA com capacidade para reduzir os níveis em uma célula-alvo de um ligante imunossupressor, um sinal mitogênico ou um fator de crescimento.

[00888] Em algumas modalidades, um fusossoma compreende um agente que é imunomodulador, por exemplo, imunoestimulador.

[00889] Em algumas modalidades, um fusossoma tem capacidade para fazer (por exemplo, fazer com que) a célula-alvo apresentar um antígeno. Da mesma forma, em algumas modalidades, um método no presente documento compreende a apresentação de um antígeno em uma célula-alvo.

[00890] Em algumas modalidades, o fusossoma promove a regeneração em um tecido-alvo. Da mesma forma, em algumas modalidades, um método no presente documento compreende promover a regeneração em um tecido-alvo. Em modalidades, a célula-alvo é uma célula cardíaca, por exemplo, um cardiomiócito (por exemplo, um cardiomiócito quiescente), um hepatoblasto (por exemplo, um hepatoblasto do duto biliar), uma célula epitelial, uma célula T virgem, um macrófago (por exemplo, um tumor macrófago infiltrante) ou um fibroblasto (por exemplo, um fibroblasto cardíaco). Em modalidades, a célula de ori-

gem é uma célula T (por exemplo, um Treg), um macrófago ou um miócito cardíaco.

[00891] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um ácido nucleico a uma célula-alvo, por exemplo, para modificar de forma estável o genoma da célula-alvo, por exemplo, para terapia genética. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar um ácido nucleico a uma célula-alvo. Em algumas modalidades, a célula-alvo tem uma deficiência de enzima, por exemplo, compreende uma mutação em uma enzima que causa a uma atividade reduzida (por exemplo, nenhuma atividade) da enzima.

[00892] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um reagente que medeia uma modificação específica de sequência de DNA (por exemplo, Cas9, ZFN ou TALEN) na célula-alvo. Da mesma forma, em algumas modalidades, um método no presente documento compreende a entrega do reagente à célula-alvo. Nas modalidades, a célula-alvo é uma célula muscular (por exemplo, célula do músculo esquelético), célula do rim ou célula hepática.

[00893] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, distribuir) um ácido nucleico a uma célula-alvo, por exemplo, para modificar transitoriamente a expressão do gene na célula-alvo.

[00894] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) uma proteína a uma célula-alvo, por exemplo, para resgatar transitoriamente uma deficiência de proteína. De modo semelhante, em algumas modalidades, um método no presente documento compreende entregar uma proteína a uma célula-alvo. Em modalidades, a proteína é uma proteína de membrana (por exemplo, uma proteína transportadora de membrana), uma proteína

citoplasmática (por exemplo, uma enzima) ou uma proteína secretada (por exemplo, uma proteína imunossupressora).

[00895] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) uma organela a uma célula-alvo, por exemplo, em que a célula-alvo tem uma rede de organela defeituosa. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar uma organela a uma célula-alvo. Em modalidades, a célula de origem é um hepatócito, célula de músculo esquelético ou neurônio.

[00896] Em algumas modalidades, o fusossoma tem capacidade para entregar (por exemplo, entrega) um núcleo a uma célula-alvo, por exemplo, em que a célula-alvo tem uma mutação genética. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar um núcleo a uma célula-alvo. Em algumas modalidades, o núcleo é autólogo e compreende uma ou mais alterações genéticas em relação à célula-alvo, por exemplo, compreende uma modificação específica da sequência no DNA (por exemplo, Cas9, ZFN ou TALEN) ou um cromossomo artificial, um gene adicional sequência integrada no genoma, uma deleção ou qualquer combinação das mesmas. Em modalidades, a fonte do núcleo autólogo é uma célula-tronco, por exemplo, uma célula-tronco hematopoiética. Em modalidades, a célula-alvo é uma célula muscular (por exemplo, uma célula de músculo esquelético ou cardiomiócito), um hepatócito ou um neurônio.

[00897] Em algumas modalidades, o fusossoma tem capacidade para entrega molecular intracelular, por exemplo, entrega um agente de proteína a uma célula-alvo. Da mesma forma, em algumas modalidades, um método no presente documento compreende entregar uma molécula a uma região intracelular de uma célula-alvo. Nas modalidades, o agente de proteína é um inibidor. Em modalidades, o agente de

proteína compreende um nanocorpo, scFv, anticorpo de camelídeo, peptídeo, macrociclo ou molécula pequena.

[00898] Em algumas modalidades, o fusossoma tem capacidade para fazer com que (por exemplo, faz com que) uma célula-alvo secrete uma proteína, por exemplo, uma proteína terapêutica. Da mesma forma, em algumas modalidades, um método no presente documento compreende fazer com que uma célula-alvo secrete uma proteína.

[00899] Em algumas modalidades, o fusossoma tem capacidade para secretar (por exemplo, secreta) um agente, por exemplo, uma proteína. Em algumas modalidades, o agente, por exemplo, agente secretado, é entregue a um local-alvo em um indivíduo. Em algumas modalidades, o agente é uma proteína que não pode ser produzida de forma recombinante ou é difícil de ser produzida de maneira recombinante. Em algumas modalidades, o fusossoma que secreta uma proteína é de uma célula de origem selecionada a partir de um MSC ou um condrócito.

[00900] Em algumas modalidades, o fusossoma compreende em sua membrana um ou mais ligantes da superfície celular (por exemplo, 1, 2, 3, 4, 5, 10, 20, 50 ou mais ligantes da superfície celular). De modo semelhante, em algumas modalidades, um método no presente documento compreende apresentar um ou mais ligantes da superfície celular a uma célula-alvo. Em algumas modalidades, o fusossoma com um ligante de superfície celular é de uma célula de origem escolhida a partir de um neutrófilo (por exemplo, e a célula-alvo é um linfócito infiltrante de tumor), célula dendrítica (por exemplo, e a célula-alvo é uma célula T virgem) ou neutrófilos (por exemplo, e o alvo é uma célula tumoral ou célula infectada por vírus). Em algumas modalidades, o fusossoma compreende um complexo de membrana, por exemplo, um complexo compreendendo pelo menos 2, 3, 4 ou 5 proteínas, por exemplo, um homodímero, heterodímero, homotrímero, heterotrímero,

homotetrâmero ou heterotetrâmero. Em algumas modalidades, o fusossoma compreende um anticorpo, por exemplo, um anticorpo tóxico, por exemplo, o fusossoma tem capacidade para entregar o anticorpo ao local-alvo, por exemplo, por homing para um local-alvo. Em algumas modalidades, a célula de origem é uma célula NK ou neutrófilo.

[00901] Em algumas modalidades, um método no presente documento compreende causar a secreção de uma proteína de uma célula-alvo ou apresentação de ligante na superfície de uma célula-alvo. Em algumas modalidades, o fusossoma tem capacidade para causar a morte celular da célula-alvo. Em algumas modalidades, o fusossoma é de uma célula NK de origem.

[00902] Em algumas modalidades, um fusossoma ou célula-alvo tem capacidade para fagocitose (por exemplo, de um patógeno). De modo semelhante, em algumas modalidades, um método no presente documento compreende causar fagocitose.

[00903] Em algumas modalidades, um fusossoma detecta e responde ao seu ambiente local. Em algumas modalidades, o fusossoma tem capacidade para detectar o nível de um metabólito, interleucina ou antígeno.

[00904] Em modalidades, um fusossoma tem capacidade para quimiotaxia, extravasamento ou uma ou mais atividades metabólicas. Nas modalidades, a atividade metabólica é selecionada a partir de quineurinina, gliconeogênese, oxidação de ácidos graxos de prostaglandina, metabolismo de adenosina, ciclo de ureia e respiração termogênica. Em algumas modalidades, a célula de origem é um neutrófilo e o fusossoma tem capacidade para homing a um sítio de lesão. Em algumas modalidades, a célula de origem é um macrófago e o fusossoma tem capacidade para fagocitose. Em algumas modalidades, a célula de origem é uma célula de tecido adiposo marrom e o fusossoma tem capacidade para lipólise.

[00905] Em algumas modalidades, o fusossoma compreende (por exemplo, tem capacidade para fornecer à célula-alvo) uma pluralidade de agentes (por exemplo, pelo menos 2, 3, 4, 5, 10, 20 ou 50 agentes). Em modalidades, o fusossoma compreende um ácido nucleico inibidor (por exemplo, siRNA ou miRNA) e um mRNA.

[00906] Em algumas modalidades, o fusossoma compreende (por exemplo, tem capacidade para entregar à célula-alvo) uma proteína de membrana ou um ácido nucleico que codifica a proteína de membrana. Em modalidades, o fusossoma tem capacidade para reprogramar ou transdiferenciar uma célula-alvo, por exemplo, o fusossoma compreende um ou mais agentes que induzem a reprogramação ou transdiferenciação de uma célula-alvo.

[00907] Em algumas modalidades, o indivíduo precisa de regeneração. Em algumas modalidades, o indivíduo sofre de câncer, uma doença autoimune, uma doença infecciosa, uma doença metabólica, uma doença neurodegenerativa ou uma doença genética (por exemplo, deficiência de enzima).

[00908] Em algumas modalidades (por exemplo, modalidades para ensaio de entrega não endocítica de carga), a entrega de carga é avaliada com o uso de uma ou mais dentre (por exemplo, todas) as seguintes etapas: (a) colocar 30.000 células-alvo HEK-293T em um primeiro poço de uma placa de 96 poços que compreende 100 nM de bafilomicina A1 e colocar um número semelhante de células semelhantes em um segundo poço de uma placa de 96 poços sem bafilomicina A1, (b) cultivar as células-alvo durante quatro horas em meio DMEM a 37 °C e 5% De CO2, (c) colocar as células-alvo em contato com 10 μg de fusossomas que compreendem carga, (d) incubar as células-alvo e fusossomas durante 24 horas a 37 °C e 5% de CO2 e (e) determinar a porcentagem de células no primeiro poço e no segundo poço que compõe a carga. A etapa (e) pode compreender a detec-

ção da carga com o uso de microscopia, por exemplo, com o uso de imunofluorescência. A etapa (e) pode compreender a detecção indireta da carga, por exemplo, detectar um efeito a jusante da carga, por exemplo, a presença de uma proteína repórter. Em algumas modalidades, uma ou mais dentre as etapas (a) a (e) acima são realizadas conforme descrito no Exemplo 81.

[00909] Em algumas modalidades, um inibidor de endocitose (por exemplo, cloroquina ou bafilomicina A1) inibe a acidificação endossomal. Em algumas modalidades, a entrega de carga é independente da acidificação lisossomal. Em algumas modalidades, um inibidor de endocitose (por exemplo, Dynasore) inibe a dinamina. Em algumas modalidades, a entrega de carga é independente da atividade da dinamina.

[00910] Em algumas modalidades (por exemplo, modalidades para entrega específica de carga a uma célula-alvo versus uma célula não alvo), a entrega de carga é avaliada com o uso de uma ou mais (por exemplo, todas) dentre as seguintes etapas: (a) colocar 30.000 células-alvo HEK-293T que superexpressam CD8a e CD8b em um primeiro poço de uma placa de 96 poços e colocar 30.000 células não alvo HEK-293T que não superexpressam a CD8a e CD8b em um segundo poço de uma placa com 96 poços (b) cultivar as células durante quatro horas em meios DMEM a 37 °C e 5% de CO2, (c) colocar as células-alvo em contato com 10 μg de fusossomas que compreendem carga, (d) incubar as células-alvo e fusossomas durante 24 horas a 37 °C e 5% de CO2 e (e) determinar a porcentagem de células no primeiro poço e no segundo poço que compreendem a carga. A etapa (e) pode compreender a detecção da carga com o uso de microscopia, por exemplo, com o uso de imunofluorescência. A etapa (e) pode compreender a detecção indireta da carga, por exemplo, detectar um efeito a jusante da carga, por exemplo, a presença de uma proteína repórter.

Em algumas modalidades, uma ou mais dentre as etapas (a) a (e) acima são realizadas conforme descrito no Exemplo 72.

[00911] Em algumas modalidades, o fusossoma é fundido com uma célula-alvo a uma taxa mais alta do que com uma célula não alvo, por exemplo, a pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes, por exemplo, em um ensaio do Exemplo 43. Em algumas modalidades, o fusossoma é fundido com uma célula-alvo a uma taxa mais alta do que com outros fusossomas, por exemplo, por pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, ou 90%, por exemplo, em um ensaio do Exemplo 43. Em algumas modalidades, o fusossoma é fundido com células-alvo a uma taxa de modo que um agente no fusossoma é entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% de células-alvo após 24, 48 ou 72 horas, por exemplo, em um ensaio do Exemplo 54. Em modalidades, a quantidade de fusão direcionada é de aproximadamente 30% a 70%, 35% a 65%, 40% a 60%, 45% a 55% ou 45% a 50%, por exemplo, de aproximadamente 48,8%, por exemplo, em um ensaio do Exemplo 43. Nas modalidades, a quantidade de fusão direcionada é de aproximadamente 20% a 40%, 25% a 35% ou 30% a 35%, por exemplo, de aproximadamente 32,2%, por exemplo, em um ensaio do Exemplo 44.

[00912] Em algumas modalidades, o fusogênio está presente em um número de cópias de pelo menos, ou não mais do que, 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias, por exemplo, conforme medido por um ensaio do Exemplo 27. Em algumas modalidades, pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% ou 99% do fusogênio compreendido pelo fusos-

soma é disposto na membrana celular. Em modalidades, o fusossoma também compreende fusogênio internamente, por exemplo, no citoplasma ou em uma organela. Em algumas modalidades, o fusogênio compreende (ou é identificado como compreendendo) aproximadamente 0,1%, 0,2%, 0,3%, 0,4%, 0,5%, 0,6%, 0,7%, 0,8%, 0,9%, 1%, 5%, 10%, 11%, 12%, 13%, 14%, 15%, 20% ou mais, ou aproximadamente 1 a 30%, 5 a 20%, 10 a 15%, 12 a 15%, 13 a 14% ou 13,6% da proteína total em um fusossoma, por exemplo, conforme determinado de acordo com o método descrito em Exemplo 95 e/ou por um ensaio de espectrometria de massa. Em modalidades, o fusogênio compreende (ou é identificado como compreendendo) aproximadamente 13,6% da proteína total no fusossoma. Em algumas modalidades, o fusogênio é (ou é identificado como sendo) mais ou menos abundante que uma ou mais proteínas adicionais de interesse, por exemplo, conforme determinado de acordo com o método descrito no Exemplo 95. Em uma modalidade, o fusogênio tem (ou é identificado como tendo) uma razão para EGFP de aproximadamente 140, 145, 150, 151, 152, 153, 154, 155, 156, 157 (por exemplo, 156,9), 158, 159, 160, 165 ou 170. Em outra modalidade, o fusogênio tem (ou é identificado como tendo) uma razão entre CD63 de aproximadamente 2.700, 2.800, 2.900, 2.910 (por exemplo, 2.912), 2.920, 2.930, 2.940, 2.950, 2.960, 2.970, 2.980, 2.990 ou 3.000 ou de aproximadamente 1.000 a 5000, 2.000 a 4.000, 2.500 a 3.500, 2.900 a 2.930, 2.910 a 2.915 ou 2.912,0, por exemplo, por um ensaio de espectrometria de massa. Em uma modalidade, o fusogênio tem (ou é identificado como tendo) uma razão entre ARRDC1 de aproximadamente 600, 610, 620, 630, 640, 650, 660 (por exemplo, 664,9), 670, 680, 690 ou 700. Em outra modalidade, o fusogênio tem (ou é identificado como tendo) uma razão para GAPDH de aproximadamente 50, 55, 60, 65, 70 (por exemplo, 69), 75, 80 ou 85, ou de aproximadamente 1 a 30%, 5 a 20%, 10 a 15%, 12 a

15%, 13 a 14% ou 13,6%. Em outra modalidade, o fusogênio tem (ou é identificado como tendo) uma razão para CNX de aproximadamente 500, 510, 520, 530, 540, 550, 560 (por exemplo, 558,4), 570, 580, 590 ou 600 ou aproximadamente 300 a 800, 400 a 700, 500 a 600, 520 a 590, 530 a 580, 540 a 570, 550 a 560 ou 558,4, por exemplo, por um ensaio de espectometria de massa.

[00913] Em algumas modalidades, o fusossoma compreende um agente terapêutico em um número de cópias de pelo menos, ou no máximo, 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 de cópias, por exemplo, conforme medido por um ensaio de Exemplo 89. Em algumas modalidades, o fusossoma compreende um agente terapêutico de proteína em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias, por exemplo, conforme medido por um ensaio do Exemplo 89. Em algumas modalidades, o fusossoma compreende um agente terapêutico de ácido nucleico em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, o fusossoma compreende um agente terapêutico de ácido nucleico em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, o fusossoma compreende um agente terapêutico de RNA em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000,

200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, o fusossoma compreende um agente terapêutico exógeno em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, o fusossoma compreende um agente terapêutico de proteína em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, o fusossoma compreende um agente terapêutico exógeno de ácido nucleico (por exemplo, DNA ou RNA) em um número de cópias de pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias. Em algumas modalidades, a razão entre o número de cópias do fusogênio e o número de cópias do agente terapêutico está entre 1.000.000:1 e 100.000:1, 100.000:1 e 10.000:1, 10.000:1 e 1.000:1, 1.000:1 e 100:1, 100:1 e 50:1, 50:1 e 20:1, 20:1 e 10:1, 10:1 e 5:1, 5:1 e 2:1, 2:1 e 1:1, 1:1 e 1:2, 1:2 e 1:5, 1:5 e 1:10, 1:10 e 1:20, 1:20 e 1:50, 1:50 e 1:100, 1:100 e 1:1.000, 1:1.000 e 1:10.000, 1:10.000 e 1:100.000 ou 1:100.000 e 1:1.000.000.

[00914] Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias de um agente terapêutico. Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10, 50, 100, 500, 1.000,

2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias de um agente terapêutico de proteína. Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias de um agente terapêutico de ácido nucleico. Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias de um agente terapêutico de RNA. Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10, 50, 100, 500, 1.000, 2.000, 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 500.000, 1.000.000, 5.000.000, 10.000.000, 50.000.000, 100.000.000, 500.000.000 ou 1.000.000.000 cópias de um agente terapêutico de DNA.

[00915] Em algumas modalidades, o fusossoma entrega a uma célula-alvo pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% ou 99% da carga (por exemplo, um agente terapêutico, por exemplo, um agente terapêutico endógeno ou um agente terapêutico exógeno) compreendido pelo fusossoma. Em algumas modalidades, os fusossomas que são fundidos à célula-alvo (ou células-alvo) entregam à célula-alvo uma média de pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90 %, 95%, 96%, 97%, 98% ou 99% da carga (por exemplo, um agente terapêutico, por exemplo, um agente terapêutico endógeno ou um agente terapêutico exógeno) composta pelos fusossomas que são fundidos à célula-alvo (ou células-alvo). Em algumas modalidades, a composição do fusossoma fornece a um tecido-alvo pelo menos 10%, 20%, 30%, 40%,

50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% ou 99% da carga (por exemplo, um agente terapêutico, por exemplo, um agente terapêutico endógeno ou um agente terapêutico exógeno) composta pela composição de fusossoma.

[00916] Em algumas modalidades, o fusossoma compreende 0,00000001 mg de fusogênio a 1 mg fusogênio por mg da proteína total em fusossoma, por exemplo, 0,00000001 a 0,0000001, 0,0000001 a 0,000001, 0,000001 a 0,00001, 0,00001 a 0,0001, 0,0001 a 0,001, 0,001 a 0,01, 0,01 a 0,1 ou 0,1 a 1 mg de fusogênio por mg da proteína total no fusossoma. Em algumas modalidades, o fusossoma compreende 0,00000001 mg de fusogênio a 5 mg de fusogênio por mg de lipídio no fusossoma, por exemplo, 0,00000001 a 0,0000001, 0,0000001 a 0,000001, 0,000001 a 0,00001, 0,00001 a 0,0001, 0,0001 a 0,001, 0,001 a 0,01, 0,01 a 0,1, 0,1 a 1 ou 1 a 5 mg de fusogênio por mg de lipídio no fusossoma.

[00917] Em algumas modalidades, a carga é uma carga de proteína. Em modalidades, a carga é uma carga de proteína endógena ou sintética. Em algumas modalidades, os fusossomas têm (ou são identificados como tendo) pelo menos 1, 2, 3, 4, 5, 10, 20, 50, 100 ou mais moléculas de carga de proteína por fusossoma. Em uma modalidade, os fusossomas têm (ou são identificados como tendo) de aproximadamente 100, 110, 120, 130, 140, 150, 160, 166, 170, 180, 190 ou 200 moléculas de agente de proteína por fusossoma, por exemplo, conforme quantificado de acordo com ao método descrito no Exemplo 89. Em algumas modalidades, a carga de proteína endógena ou sintética compreende (ou é identificada como compreendendo) aproximadamente 0,1%, 0,2%, 0,3%, 0,4%, 0,5%, 1%, 5%, 10%, 15%, 20%, 25% ou mais da proteína total em um fusossoma. Em uma modalidade, a carga de proteína sintética compreende (ou é identificada como compreendendo) aproximadamente 13,6% da proteína total em um fusos-

soma. Em algumas modalidades, a carga da proteína sintética tem (ou é identificada como tendo) uma razão de VSV-G de aproximadamente 4 x 10^{-3}, 5 x 10^{-3}, 6 x 10^{-3} (por exemplo, 6,37 x 10^{-3}), 7 x 10^{-3} ou 8 x 10^{-3}. Em modalidades, a carga de proteína sintética tem (ou é identificada como tendo) uma razão para CD63 de aproximadamente 10, 15, 16, 17, 18 (por exemplo, 18,6), 19, 20, 25 ou 30 ou de aproximadamente 10 a 30, 15 a 25, 16 a 19, 18 a 19 ou 18,6. Em modalidades, a carga de proteína sintética tem (ou é identificada como tendo) uma razão para ARRDC1 de aproximadamente 2, 3, 4 (por exemplo, 4,24), 5, 6 ou 7. Em modalidades, a carga de proteína sintética tem (ou é identificada como tendo) uma razão para GAPDH de aproximadamente 0,1, 0,2, 0,3, 0,4 (por exemplo, 0,44), 0,5, 0,6 ou 0,7. Em modalidades, a carga de proteína sintética tem (ou é identificada como tendo) uma razão para CNX de aproximadamente 1, 2, 3 (por exemplo, 3,56), 4, 5 ou 6. Em modalidades, a carga de proteína sintética tem (ou é identificada como tendo) uma razão para TSG101 de aproximadamente 10, 15, 16, 17, 18, 19 (por exemplo, 19,52), 20, 21, 22, 23, 24, 25 ou 30.

[00918] Em algumas modalidades, o fusogênio compreende (ou é identificado como compreendendo) pelo menos 0,5%, 1%, 5%, 10% ou mais da proteína total em um fusossoma, por exemplo, por um ensaio de espectrometria de massa. Em uma modalidade, o fusogênio compreende (ou é identificado como compreendendo) aproximadamente 1 a 30%, 5 a 20%, 10 a 15%, 12 a 15%, 13 a 14% ou 13,6% da proteína total em um fusossoma, por exemplo, por um ensaio de espectrometria de massa. Em algumas modalidades, o fusogênio é mais abundante do que outras proteínas de interesse. Em modalidades, o fusogênio tem (ou é identificado como tendo) uma razão para uma proteína de carga útil, por exemplo, EGFP, de aproximadamente 145 a 170, 150 a 165, 155 a 160, 156,9, por exemplo, por um ensaio de espectrometria de massa. Em modalidades, o fusogênio tem (ou é identificado como

tendo) uma razão para CD63 de aproximadamente 1.000 a 5.000, 2.000 a 4.000, 2.500 a 3.500, 2.900 a 2.930, 2.910 a 2.915 ou 2.912,0, por exemplo, por um ensaio de espectrometria de massa. Em modalidades, o fusogênio tem uma razão para ARRDC1 de aproximadamente 300 a 1.000, 400 a 900, 500 a 800, 600 a 700, 640 a 690, 650 a 680, 660 a 670 ou 664,9, por exemplo, por um ensaio de espectrometria de massa. Em modalidades, o fusogênio tem (ou é identificado como tendo) uma razão para GAPDH de aproximadamente 20 a 120, 40 a 100, 50 a 90, 60 a 80, 65 a 75, 68 a 70 ou 69,0, por exemplo, por um ensaio de espectrometria de massa. Em modalidades, o fusogênio tem uma razão para CNX de aproximadamente 200 a 900, 300 a 800, 400 a 700, 500 a 600, 520 a 590, 530 a 580, 540 a 570, 550 a 560 ou 558,4, por exemplo, por um ensaio de espectrometria de massa. Nas modalidades, o ensaio de espectrometria de massa é um ensaio do Exemplo 95.

[00919] Em algumas modalidades, o número de espécies de lipídios presentes em ambos (por exemplo, compartilhado entre) os fusossomas e células de origem é (ou é identificado como sendo) pelo menos 300, 400, 500, 550, 560 ou 569, ou está entre 500 a 700, 550 a 600 ou 560 a 580, por exemplo, com o uso de um ensaio de espectrometria de massa. Em modalidades, o número de espécies de lipídios presentes em fusossomas a um nível de pelo menos 25% do nível de lipídios correspondente nas células de origem (ambos normalizados para os níveis de lipídios totais em uma amostra) é (ou é identificado como sendo) pelo menos 300, 400, 500, 530, 540 ou 548, ou está entre 400 a 700, 500 a 600, 520 a 570, 530 a 560 ou 540 a 550, por exemplo, com o uso de um ensaio da espectrometria de massa. Em algumas modalidades, a fração de espécies de lipídios presentes em ambos (por exemplo, compartilhada entre) os fusossomas e células de origem para espécies de lipídios totais na célula de origem é (ou é identificada

como sendo) de aproximadamente 0,4 a 1,0, 0,5 a 0,9, 0,6 a 0,8 ou 0,7 ou pelo menos 0,4, 0,5, 0,6 ou 0,7, por exemplo, com o uso de um ensaio de espectrometria de massa. Em algumas modalidades, o ensaio de espectrometria de massa é um ensaio de Exemplo 87.

[00920] Em algumas modalidades, o número de espécies de proteínas presentes em ambos os (por exemplo, compartilhados entre) os fusossomas são células de origem é (ou é identificado como sendo) pelo menos 500, 1.000, 1.100, 1.200, 1.300, 1.400, 1.487, 1.500 ou 1.600 ou é (ou é identificado como sendo) entre 1.200 a 1.700, 1.300 a 1.600, 1.400 a 1.500, 1.450 a 1.500 ou 1.480 a 1.490, por exemplo, com o uso de um ensaio da espectrometria de massa. Em modalidades, o número de espécies de proteínas presentes em fusossomas a um nível de pelo menos 25% do nível de proteína correspondente nas células de origem (ambos normalizados para níveis de proteína total dentro de uma amostra) é (ou é identificado como sendo) pelo menos 500, 600, 700, 800, 900, 950, 957, 1.000 ou 1.200, por exemplo, com o uso de um ensaio da espectrometria de massa. Em algumas modalidades, a fração de espécies de proteínas presentes (por exemplo, compartilhada entre) tanto nos fusossomas quanto nas células de origem para espécies de proteínas totais na célula de origem é (ou é identificada como sendo) de aproximadamente 0,1 a 0,6, 0,2 a 0,5, 0,3 a 0,4 ou 0,333, ou pelo menos de aproximadamente 0,1, 0,2, 0,3, 0,333 ou 0,4, por exemplo, com o uso de um ensaio de espectrometria de massa. Nas modalidades, o ensaio de espectrometria de massa é um ensaio do Exemplo 88.

[00921] Em algumas modalidades, CD63 está (ou é identificado como sendo) presente em menos de 0,048%, 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5% ou 10% da quantidade de proteína total em fusossomas, por exemplo, por um ensaio de espectrometria de massa, por exemplo, um ensaio de Exemplo 90.

[00922] Em algumas modalidades, os fusossomas são produzidos por extrusão através de um filtro, por exemplo, um filtro de aproximadamente 1 a 10, 2 a 8, 3 a 7, 4 a 6 ou 5 um. Em algumas modalidades, os fusossomas têm (ou são identificados como tendo) um diâmetro médio de aproximadamente 1 a 5, 2 a 5, 3 a 5, 4 a 5 ou 5 um. Em algumas modalidades, os fusossomas têm (ou são identificados como tendo) um diâmetro médio de pelo menos 1, 2, 3, 4 ou 5 um.

[00923] Em algumas modalidades, os fusossomas são enriquecidos para (ou são identificados como sendo enriquecidos para) um ou mais (por exemplo, pelo menos 2, 3, 4, 5 ou todos) os seguintes lipídios em comparação às células de origem: éster colesterílico, colesterol livre, liso-fosfatidiletanolamina ligada a éter, liso-fosfatidilserina, fosfatidato, fosfatidiletanolamina ligada a éter, fosfatidilserina e esfingomielina. Em algumas modalidades, os fusossomas são depletados para (ou são identificados como sendo depletados para) um ou mais (por exemplo, pelo menos 2, 3, 4, 5 ou todos) os seguintes lipídios em comparação às células de origem: ceramida, cardiolipina, liso-fosfatidilcolina, liso-fosfatidiletanolamina, liso-fosfatidilglicerol, liso-fosfatidilinositol, fosfatidilcolina ligada a éter, fosfatidiletanolamina, fosfatidilglicerol, fosfatidilinositol, fosfatidiacilglicerol, fosfatidilinositol. Em algumas modalidades, os fusossomas são enriquecidos para (ou são identificados como sendo enriquecidos para) um ou mais dos lipídios enriquecidos mencionados acima e depletados para um ou mais dos lipídios depletados mencionados acima. Em algumas modalidades, os fusossomas compreendem (ou são identificados como compreendendo) o lipídio enriquecido como uma porcentagem do lipídio total que é pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 5 vezes ou 10 vezes maior que o nível correspondente nas células de origem. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) o lipídio depletado como uma porcentagem do lipídio total

a um nível que é inferior a 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% ou 10% do nível correspondente nas células de origem. Nas modalidades, o enriquecimento de lipídios é medido por um ensaio de espectrometria de massa, por exemplo, um ensaio do Exemplo 97.

[00924] Em algumas modalidades, os níveis de lipídios CE são (ou são identificados como sendo) de aproximadamente 2 vezes maiores em fusossomas do que em exossomas e/ou de aproximadamente 5, 6, 7, 8, 9 ou 10 vezes maiores em fusossomas do que em células parentais (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de ceramida são (ou são identificados como sendo) de aproximadamente 2, 3, 4 ou 5 vezes maiores nas células parentais do que nos fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de colesterol são (ou são identificados como sendo) de aproximadamente 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes maior em exossomas do que em fusossomas e/ou de aproximadamente 2 vezes maior em fusossomas do que em células parentais (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de CL são (ou são identificados como sendo) pelo menos de aproximadamente 5, 10, 20, 30 ou 40 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios DAG são (ou são identificados como sendo) aproximadamente 2 ou 3 vezes maiores em exossomas do que em fusossomas e/ou aproximadamente 1,5 ou 2 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PC são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 1,3, 1,4, 1,5, 1,6, 1,7 ou 1,8 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os

níveis de lipídios PC O são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 2 vezes mais elevados nas células parentais do que nos fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PE são (ou são identificados como sendo) de aproximadamente 1,3, 1,4, 1,5, 1,6, 1,7 ou 1,8 vezes maior em fusossomas do que em exossomas e/ou de aproximadamente 1,3, 1,4, 1,5, 1,6, 1,7, ou 1,8 vezes maior em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídio PE O- são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes mais alto em células parentais do que em fusossomas (relativo ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PG são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 2, 3, 4, 5, 6, 7, 8, 9 ou 10 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PI são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 3, 4, 5, 6 ou 7 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PS são (ou são identificados como sendo) (ou são identificados como sendo) aproximadamente iguais entre exossomas e fusossomas e/ou de aproximadamente 2 vezes maiores em fusossomas do que em células parentais (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de SM são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 2, 2,5 ou 3 vezes maiores em fusossomas do que em células parentais (em relação ao lipídio total em

uma amostra). Em algumas modalidades, os níveis de lipídios de TAG são (ou são identificados como sendo) quase iguais entre exossomas e fusossomas e/ou de aproximadamente 10, 20, 30, 40, 50, 60, 70 80, 90, 100 vezes ou mais em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra).

[00925] Em algumas modalidades, os fusossomas são (ou são identificados como sendo) enriquecidos por um ou mais dos (por exemplo, pelo menos 2, 3, 4, 5 ou todos) os seguintes lipídios em comparação a exossomas: éster colesterílico, ceramida, diacilglicerol, liso-fosfatidato e fosfatidiletanolamina e triacilglicerol. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) depletados para um ou mais dentre (por exemplo, pelo menos 2, 3, 4, 5 ou todos) os seguintes lipídios em comparação a exossomas (em relação ao lipídio total em uma amostra): colesterol livre, hexosilceramida, liso-fosfatidilcolina, liso-fosfatidilcolina ligada a éter, liso-fosfatidiletanolamina, liso-fosfatidiletanolamina ligada a éter e liso-fosfatidilserina. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) enriquecidos para um ou mais dentre os lipídios enriquecidos mencionados anteriormente e depletados para um ou mais dos lipídios depletados mencionados anteriormente. Em algumas modalidades, os fusossomas compreendem (ou são identificados como compreendendo) o lipídio enriquecido como uma porcentagem do lipídio total que é pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 5 vezes ou 10 vezes maior do que o nível correspondente em exossomas. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) o lipídio depletado como uma porcentagem do lipídio total a um nível que é inferior a 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% ou 10% do nível correspondente em exossomas. Nas modalidades, o enriquecimento de lipídios é medido por um ensaio de espectrometria de massa, por

exemplo, um ensaio do Exemplo 97.

[00926] Em algumas modalidades, os níveis de lipídeos de ceramida são (ou são identificados como sendo) de aproximadamente 2 vezes mais elevados em fusossomas do que em exossomas e/ou de aproximadamente 2 vezes mais elevados em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios HexCer são (ou são identificados como sendo) de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes maior em exossomas do que em fusossomas e/ou quase igual em células parentais e fusossomas (em relação ao total lipídio em uma amostra). Em algumas modalidades, os níveis de lipídios do LPA são (ou são identificados como sendo) de aproximadamente 3 ou 4 vezes maiores em fusossomas do que em exossomas e/ou de aproximadamente 1,3, 1,4, 1,5, 1,6, 1,7 ou 1,8 vez maiores em fusossomas do que em células parentais (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de LPC são (ou são identificados como sendo) aproximadamente 2 vezes maiores em exossomas do que em fusossomas e/ou de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes maiores em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídio de LPC O são (ou são identificados como sendo) de aproximadamente 3 ou 4 vezes maior em exossomas do que em fusossomas e/ou quase igual entre células parentais e fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios do LPE são (ou são identificados como sendo) de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes maior em exossomas do que em fusossomas e/ou de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9, ou 2 vezes maior em células parentais do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídio de LPE O são (ou são identificados

como sendo) de aproximadamente 2 ou 3 vezes maior em exossomas do que em fusossomas e/ou quase iguais entre células parentais e fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de LPS são (ou são identificados como sendo) de aproximadamente 3 vezes maiores em exossomas do que em fusossomas (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PA são (ou são identificados como sendo) de aproximadamente 1,5, 1,6, 1,7, 1,8, 1,9 ou 2 vezes maior em fusossomas do que em exossomas e/ou de aproximadamente 2 vezes maior em fusossomas do que em células parentais (em relação ao lipídio total em uma amostra). Em algumas modalidades, os níveis de lipídios de PG são (ou são identificados como sendo) quase iguais entre fusossomas e exossomas e/ou de aproximadamente 10, 11, 12, 13, 14 ou 15 vezes maiores em células parentais do que em fusossomas (em relação ao total lipídio em uma amostra).

[00927] Em algumas modalidades, o fusossoma compreende uma composição lipídica substancialmente semelhante à da célula de origem ou em que um ou mais dentre CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM e TAG está dentro de 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% ou 50% do nível de lipídio correspondente na célula de origem. Em modalidades, a composição lipídica de fusossomas é semelhante às células das quais eles são derivados. Em modalidades, fusossomas e células parentais têm (ou são identificados como tendo) uma composição lipídica semelhante, caso aproximadamente 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85% ou 90% ou menos das espécies de lipídios identificadas em qualquer amostra replicada das células parentais estejam presentes (ou sejam identificados como estando presentes) em qualquer amostra replicada dos fusossomas, por exemplo, conforme determinado de acordo com o Exemplo 87. Em modalidades, de lipídios identificados,

o nível médio no fusossoma é maior do que de aproximadamente 10%, 15%, 20%, 25%, 30%, 35% ou 40% do nível médio de espécies de lipídios correspondente na célula parental (em relação ao lipídio total em uma amostra). Em uma modalidade, a composição lipídica do fusossoma é enriquecida e/ou depletada para lipídios específicos em relação à célula parental (em relação ao lipídio total em uma amostra).

[00928] Em algumas modalidades, a composição lipídica do fusossoma é (ou é identificado como trazer) enriquecido e/ou depletado para lipídios específicos em relação à célula parental, por exemplo, conforme determinado de acordo com o método descrito no Exemplo 97.

[00929] Em algumas modalidades, o fusossoma tem (ou é identificado como tendo) uma razão entre fosfatidilserina e o total de lipídios que é maior que a da célula parental. Em modalidades, o fusossoma tem (ou é identificado como tendo) uma razão entre fosfatidilserina e o total de lipídios de aproximadamente 110%, 115%, 120%, 121%, 122%, 123%, 124%, 125%, 130%, 135%, 140% ou mais em relação à célula parental. Em algumas modalidades, o fusossoma é (ou é identificado como sendo) enriquecido com éster de colesterol, colesterol livre, liso-fosfatidiletanolamina ligada a éter, liso-fosfatidilserina, fosfatidato, fosfatidiletanolamina ligada a éter, fosfatidilserina e/ou esfingomielina em relação à célula parental. Em algumas modalidades, os fusossomas são (ou é identificado como sendo) depletado para ceramida, cardiolipina, liso-fosfatidilcolina, liso-fosfatidiletanolamina, liso-fosfatidilglicerol, liso-fosfatidilinositol, fosfatidilcolina ligada a éter, fosfatidiletanolamina, fosfatidilglicerol/fosfatidilglicerol, em relação à célula parentalglicerol/fosfatidilglicerol ou fosfatidilglicerol/trental. Em algumas modalidades, o fusossoma é (ou é identificado como sendo) enriquecido com éster colesterílico, ceramida, diacilglicerol, liso-fosfatidato, fosfatidiletanolamina e/ou triacilglicerol em relação a um exossoma. Em algumas modalidades, o fusossoma é (ou é identificado como sendo)

depletado para colesterol livre, hexosil ceramida, liso-fosfatidilcolina, liso-fosfatidilcolina ligada a éter, liso-fosfatidiletanolamina, liso-fosfatidiletanolamina ligada a éter e/ou liso-fosfatidilserina em relação a um exossoma.

[00930] Em algumas modalidades, o fusossoma tem uma razão entre cardiolipina:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:ceramida na célula de origem; ou tem uma razão entre cardiolipina:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:diacilglicerol na célula de origem; ou tem uma razão entre cardiolipina:hexosilceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:hexosilceramida na célula de origem; ou tem uma razão entre cardiolipina:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:lisofosfatidato na célula de origem; ou tem uma razão entre cardiolipina:liso-fosfatidilcolina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:liso-fosfatidilcolina na célula de origem; ou tem uma razão entre cardiolipina:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre cardiolipina:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre cardiolipina:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre cardiolipina:liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:liso-fosfatidilserina na célula de origem; ou tem uma razão entre cardiolipina:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidato na célula de origem; ou tem uma razão entre cardiolipina:fosfatidilcolina que está

dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidilcolina na célula de origem; ou tem uma razão entre cardiolipina:fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidiletanolamina na célula de origem; ou tem uma razão entre cardiolipina:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidilglicerol na célula de origem; ou tem uma razão entre cardiolipina:fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidilinositol na célula de origem; ou tem uma razão entre cardiolipina:fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidilserina na célula de origem; ou tem uma razão entre cardiolipina:éster colesterílico que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:éster colesterílico na célula de origem; ou tem uma razão entre cardiolipina:esfingomielina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:esfingomielina na célula de origem; ou tem uma razão entre cardiolipina:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:triacilglicerol na célula de origem; ou tem uma razão entre fosfatidilcolina:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:ceramida na célula de origem; ou tem uma razão entre fosfatidilcolina:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:diacilglicerol na célula de origem; ou tem uma razão entre fosfatidilcolina:hexosilceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina: hexosilceramida na célula de origem; ou tem uma razão entre fosfatidilcolina:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:lisofosfatidato na célula de origem; ou tem uma razão entre fosfatidilcolina: liso-fosfatidilcolina que está dentro de 10%, 20%, 30%,

40% ou 50% da razão entre fosfatidilcolina: liso-fosfatidilcolina na célula de origem; ou tem uma razão entre fosfatidilcolina:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina: liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre fosfatidilcolina:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidilcolina:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidilcolina:liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:liso-fosfatidilserina na célula de origem; ou tem uma razão entre fosfatidilcolina:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre cardiolipina:fosfatidato na célula de origem; ou tem uma razão entre fosfatidilcolina: fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:fosfatidiletanolamina na célula de origem; ou tem uma razão entre cardiolipina:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidilcolina:fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidilcolina:fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:fosfatidilserina na célula de origem; ou tem uma razão entre fosfatidilcolina:éster colesterílico que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:éster colesterílico na célula de origem; ou tem uma razão entre fosfatidilcolina:esfingomielina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:esfingomielina na célula de origem; ou tem uma razão entre fosfati-

dilcolina:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilcolina:triacilglicerol na célula de origem; ou tem uma razão entre fosfatidiletanolamina:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:ceramida na célula de origem; ou tem uma razão entre fosfatidiletanolamina:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:diacilglicerol na célula de origem; ou tem uma razão entre fosfatidiletanolamina:hexosilceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:hexosilceramida na célula de origem; ou tem uma razão entre fosfatidiletanolamina:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:lisofosfatidato na célula de origem; ou tem uma razão entre fosfatidiletanolamina:liso-fosfatidilcolina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:liso-fosfatidilcolina na célula de origem; ou tem uma razão entre fosfatidiletanolamina:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre fosfatidiletanolamina:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidiletanolamina:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidiletanolamina: liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:liso-fosfatidilserina na célula de origem; ou tem uma razão entre fosfatidiletanolamina:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:fosfatidato na célula de origem; ou tem uma razão entre fosfatidiletanolami-

na:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidiletanolamina: fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidiletanolamina:fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:fosfatidilserina na célula de origem; ou tem uma razão entre fosfatidiletanolamina:éster colesterílico que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:éster colesterílico na célula de origem; ou tem uma razão entre fosfatidiletanolamina:esfingomielina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:esfingomielina na célula de origem; ou tem uma razão entre fosfatidiletanolamina:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidiletanolamina:triacilglicerol na célula de origem; ou tem uma razão entre fosfatidilserina:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:ceramida na célula de origem; ou tem uma razão entre fosfatidilserina:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:diacilglicerol na célula de origem; ou tem uma razão entre fosfatidilserina:hexosilceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:hexosilceramida na célula de origem; ou tem uma razão entre fosfatidilserina:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:lisofosfatidato na célula de origem; ou tem uma razão entre fosfatidilserina:liso-fosfatidilcolina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:liso-fosfatidilcolina na célula de origem; ou tem uma razão entre fosfatidilserina:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50%

da razão entre fosfatidilserina:liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre fosfatidilserina:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidilserina:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidilserina:liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:liso-fosfatidilserina na célula de origem; ou tem uma razão entre fosfatidilserina:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:fosfatidato na célula de origem; ou tem uma razão entre fosfatidilserina:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:fosfatidilglicerol na célula de origem; ou tem uma razão entre fosfatidilserina:fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:fosfatidilinositol na célula de origem; ou tem uma razão entre fosfatidilserina:éster colesterílico que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:éster colesterílico na célula de origem; ou tem uma razão entre fosfatidilserina:esfingomielina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:esfingomielina na célula de origem; ou tem uma razão entre fosfatidilserina:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre fosfatidilserina:triacilglicerol na célula de origem; ou tem uma razão entre esfingomielina:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:ceramida na célula de origem; ou tem uma razão entre esfingomielina:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:diacilglicerol na célula de origem; ou tem uma razão entre esfingomielina:hexosilceramida que está dentro de

10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:hexosilceramida na célula de origem; ou tem uma razão entre esfingomielina:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:lisofosfatidato na célula de origem; ou tem uma razão entre esfingomielina:liso-fosfatidilcolina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:liso-fosfatidilcolina na célula de origem; ou tem uma razão entre esfingomielina:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre esfingomielina:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre esfingomielina:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre esfingomielina:liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:liso-fosfatidilserina na célula de origem; ou tem uma razão entre esfingomielina:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:fosfatidato na célula de origem; ou tem uma razão entre esfingomielina:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:fosfatidilglicerol na célula de origem; ou tem uma razão entre esfingomielina:fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:fosfatidilinositol na célula de origem; ou tem uma razão entre esfingomielina:éster colesterílico que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:éster colesterílico na célula de origem; ou tem uma razão entre esfingomielina:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre esfingomielina:triacilglicerol na célula de ori-

gem; ou tem uma razão entre éster colesterílico:ceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:ceramida na célula de origem; ou tem uma razão entre éster colesterílico:diacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:diacilglicerol na célula de origem; ou tem uma razão entre éster colesterílico:hexosilceramida que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:hexosilceramida na célula de origem; ou tem uma razão entre éster colesterílico:lisofosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:lisofosfatidato na célula de origem; ou tem uma razão entre éster colesterílico:liso-fosfatidilcolina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:liso-fosfatidilcolina na célula de origem; ou tem uma razão entre éster colesterílico:liso-fosfatidiletanolamina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:liso-fosfatidiletanolamina na célula de origem; ou tem uma razão entre éster colesterílico:liso-fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:liso-fosfatidilglicerol na célula de origem; ou tem uma razão entre éster colesterílico:liso-fosfatidilinositol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:liso-fosfatidilinositol na célula de origem; ou tem uma razão entre éster colesterílico:liso-fosfatidilserina que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:liso-fosfatidilserina na célula de origem; ou tem uma razão entre éster colesterílico:fosfatidato que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:fosfatidato na célula de origem; ou tem uma razão entre éster colesterílico:fosfatidilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:fosfatidilglicerol na célula de origem; ou tem uma razão entre éster colesterílico:fosfatidilinositol que

está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:fosfatidilinositol na célula de origem; ou tem uma razão entre éster colesterílico:triacilglicerol que está dentro de 10%, 20%, 30%, 40% ou 50% da razão entre éster colesterílico:triacilglicerol na célula de origem.

[00931] Em algumas modalidades, o fusossoma compreende uma composição proteômica semelhante à da célula de origem, por exemplo, com o uso de um ensaio do Exemplo 88. Em algumas modalidades, a composição de proteína dos fusossomas é semelhante às células parentais das quais são derivados. Em algumas modalidades, o teor fracionário de cada uma dentre uma pluralidade de categorias de proteínas é determinado como a soma dos sinais de intensidade de cada categoria dividida pela soma dos sinais de intensidade de todas as proteínas identificadas na amostra, por exemplo, conforme descrito no Exemplo 88. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) quantidades variáveis de proteínas específicas do compartimento em relação às células parentais e/ou exossomas, por exemplo, conforme determinado de acordo com o método descrito no Exemplo 98. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) depletados com a proteína do retículo endoplasmático em comparação às células parentais e exossomas. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) depletados para proteína exossômica em comparação aos exossomas. Em algumas modalidades, os fusossomas têm (ou são identificados como tendo) menos de 15%, 20% ou 25% da proteína no fusossoma como sendo uma proteína exossômica. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) depletados para a proteína mitocondrial em comparação às células parentais. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) enriquecidos para proteína nuclear

em comparação às células parentais. Em algumas modalidades, os fusossomas são (ou são identificados como sendo) enriquecidos para proteínas ribossômicas em comparação a células parentais e exossomas. Em algumas modalidades, pelo menos 0,025%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%, 0,08%, 0,09%, 0,1%, 1%, 2%, 3%, 4%, 5%, 6%, 7% 8%, 9% ou 10% da proteína no fusossoma é proteína ribossomal ou de aproximadamente 0,025 a 0,2%, 0,05 a 0,15%, 0,06 a 1,4%, 0,07% a 1,3%, 0,08% a 1,2%, 0,09% a 1,1 %, 1% a 20%, 3% a 15%, 5% a 12,5%, 7,5% a 11%, ou 8,5% a 10,5%, ou 9% a 10% da proteína no fusossoma é proteína ribossomal.

[00932] Em algumas modalidades, o fusossoma compreende uma razão entre lipídios e proteínas que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um ensaio do Exemplo 41. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão entre massa lipídica e proteínas aproximadamente igual à razão entre massa lipídica e proteína para células nucleadas. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão lipídio:proteína maior que a célula parental. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão lipídio:proteína de aproximadamente 110%, 115%, 120%, 125%, 130%, 131%, 132%, 132,5%, 133%, 134%, 135%, 140%, 145% ou 150% da razão lípido:proteína da célula parental. Em algumas modalidades, o fusossoma ou composição de fusossoma tem (ou é identificado como tendo) uma razão fosfolipídio:proteína de aproximadamente 100 a 180, 110 a 170, 120 a 160, 130 a 150, 135 a 145, 140 a 142 ou 141 μmol/g, por exemplo, em um ensaio do Exemplo 84. Em algumas modalidades, o fusossoma ou composição de fusossoma tem (ou é identificado como tendo) uma razão fosfolipídeo:proteína que é de aproximadamente 60 a 90%, 70 a

80% ou 75% da razão correspondente nas células de origem, por exemplo, em um ensaio do Exemplo 84.

[00933] Em algumas modalidades, o fusossoma compreende uma razão entre proteínas e ácidos nucleicos (por exemplo, DNA ou RNA) que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um ensaio do Exemplo 42. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão entre massa de proteína e massa de DNA semelhante à de uma célula parental. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão proteína:DNA que é aproximadamente 85%, 90%, 95%, 96%, 97%, 98%, 98,2%, 99%, 100%, 101%, 102%, 103%, 104%, 105% ou 110% da célula parental. Em algumas modalidades, o fusossoma compreende uma razão entre proteínas para DNA que é maior que a razão correspondente na célula de origem, por exemplo, pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% maior, por exemplo, conforme medido com o uso de um ensaio do Exemplo 42. Em algumas modalidades, o fusossoma ou composição de fusossoma compreende (ou é identificado como compreendendo) uma razão de proteína:DNA que é aproximadamente 20 a 35, 25 a 30, 26 a 29, 27 a 28 ou 27,8 g/g, por exemplo, por um ensaio do Exemplo 85. Em algumas modalidades, o fusossoma ou composição de fusossoma compreende (ou é identificado como compreendendo) uma razão proteína:DNA que está dentro de aproximadamente 1%, 2%, 5%, 10% ou 20% da razão correspondente nas células de origem, por exemplo, por um ensaio do Exemplo 85.

[00934] Em algumas modalidades, o fusossoma compreende uma razão entre lipídios para ácidos nucleicos (por exemplo, DNA) que está dentro de 10%, 20%, 30%, 40% ou 50% da razão correspondente na célula de origem, por exemplo, conforme medido com o uso de um en-

saio do Exemplo 92. Em algumas modalidades, o fusossoma ou composição de fusossoma compreende (ou é identificado como compreendendo) uma razão lipídios:DNA que é de aproximadamente 2,0 a 6,0, 3,0 a 5,0, 3,5 a 4,5, 3,8 a 4,0 ou 3,92 µmol/mg, por exemplo, por um ensaio do Exemplo 86. Em algumas modalidades, o fusossoma compreende uma razão entre lipídios e ácidos nucleicos (por exemplo, DNA) que é maior que a razão correspondente na célula de origem, por exemplo, pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% maior, por exemplo, conforme medido com o uso de um ensaio do Exemplo 92. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma razão lipídio:DNA maior que a célula parental. Em modalidades, o fusossoma compreende de aproximadamente 105%, 110%, 115%, 120%, 125%, 130%, 135%, 140%, 145%, 150% ou maior razão lipídio:DNA em comparação à célula parental.

[00935] Em algumas modalidades, a composição do fusossoma tem meia-vida em um indivíduo, por exemplo, em um camundongo, que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% da meia-vida de uma composição de célula de referência, por exemplo, a célula de origem, por exemplo, por um ensaio do Exemplo 61. Em algumas modalidades, a composição de fusossoma tem meia-vida em um indivíduo, por exemplo, em um camundongo, que é de pelo menos 1 hora, 2 horas, 3 horas, 4 horas, 5 horas, 6 horas, 12 horas ou 24 horas, por exemplo, em um indivíduo humano ou em um camundongo, por exemplo, por um ensaio do Exemplo 61. Em modalidades, a composição de fusossoma tem uma meia-vida de pelo menos 1, 2, 4, 6, 12 ou 24 horas em um indivíduo, por exemplo, em um ensaio do Exemplo 80. Em algumas modalidades, o agente terapêutico tem uma meia-vida em um indivíduo que é maior que a meia-vida da composição de fusossoma, por exemplo, em

pelo menos 10%, 20%, 50%, 2 vezes, 5 vezes ou 10 vezes. Por exemplo, o fusossoma pode entregar o agente terapêutico à célula-alvo, e o agente terapêutico pode estar presente após o fusossoma não estar mais presente ou detectável.

[00936] Em algumas modalidades, o fusossoma transporta glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) através de uma membrana, por exemplo, em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% a mais que um controle negativo, por exemplo, um fusossoma de outra forma similar na ausência de glicose, por exemplo, conforme medido com o uso de um ensaio do Exemplo 51. Em algumas modalidades, o fusossoma transporta (ou é identificado como transportando) glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) através de uma membrana em um nível maior que fusossomas semelhantes tratados com floretina, por exemplo, em um ensaio de Exemplo 73. Em modalidades, um fusossoma não tratado com floretina transporta (ou é identificado como não transportando) glicose em um nível pelo menos 1%, 2%, 3%, 5% ou 10% maior (e opcionalmente até 15% maior) que um fusossoma de outra forma similar tratado com floretina, por exemplo, em um ensaio do Exemplo 73. Em algumas modalidades, o fusossoma compreende atividade esterase no lúmen que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% da atividade da esterase em uma célula de referência, por exemplo, a célula de origem ou um fibroblasto embrionário de camundongo, por exemplo, com o uso de um ensaio do Exemplo 52. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) atividade de esterase no lúmen que é pelo menos 10 vezes, 20 vezes, 50 vezes, 100 vezes, 200 vezes, 500 vezes, 1.000 vezes, 2.000 vezes ou 5.000 vezes maior que um controle não marcado, por exemplo, por um ensaio de Exemplo 74. Em algumas

modalidades, o fusossoma compreende (ou é identificado como compreendendo) atividade de esterase no lúmen que é de aproximadamente 10 a 100 vezes menor que a das células de origem, por exemplo, por um ensaio do Exemplo 74. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma atividade de acetilcolinesterase de aproximadamente 1E5 a 1E6, 6E5 a 8E5, 6,5E5 a 7E5, ou equivalentes de exossoma 6,83E5, por exemplo, por um ensaio do Exemplo 75. Em algumas modalidades, o fusossoma compreende um nível de atividade metabólica (por exemplo, atividade de citrato sintase) que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% do nível de atividade metabólica em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito no Exemplo 54. Em algumas modalidades, o fusossoma compreende um nível de atividade metabólica (por exemplo, atividade de citrato sintase) que é pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% do nível de atividade metabólica em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito no Exemplo 54. Em algumas modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma atividade de citrato sintase que é de aproximadamente 1E-2 a 2 E-2, 1,3E-2 a 1,8E-2, 1,4E-2 a 1,7E-2, 1,5E-2 a 1,6E-2 ou 1,57E a 2 μmol/μg de fusossoma/min, por exemplo, por um ensaio do Exemplo 76. Em algumas modalidades, o fusossoma compreende um nível de respiração (por exemplo, taxa de consumo de oxigênio), por exemplo, nível de respiração basal, desacoplado ou máximo, que está dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% do nível de respiração em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito no Exemplo 55. Em algumas modalidades, o fusossoma compreende um nível de respira-

ção (por exemplo, taxa de consumo de oxigênio), por exemplo, nível de respiração basal, desacoplado ou máximo, que é pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% do nível de respiração em uma célula de referência, por exemplo, a célula de origem, por exemplo, conforme descrito no Exemplo 55. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma taxa de respiração basal de aproximadamente 8 a 15, 9 a 14, 10 a 13, 11 a 12 ou 11,3 pmol/min/20 µg de fusossoma, por exemplo, por um ensaio do Exemplo 77. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma taxa de respiração desacoplada de aproximadamente 8 a 13, 9 a 12, 10 a 11, 10 a 10,2 ou 10,1 pmol/min/20 µg de fusossoma, por exemplo, por um ensaio de Exemplo 77. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) uma taxa de respiração máxima de aproximadamente 15 a 25, 16 a 24, 17 a 23, 18 a 22, 19 a 21 ou 20 pmol/min/20 µg de fusossoma, por exemplo, por um ensaio do Exemplo 77. Nas modalidades, o fusossoma tem (ou é identificado como tendo) uma taxa de respiração basal mais alta que a taxa de respiração não acoplada, por exemplo, por aproximadamente 1%, 2%, 5% ou 10%, por exemplo, até aproximadamente 15%, por exemplo, por um ensaio do Exemplo 77. Em modalidades, o fusossoma tem (ou é identificado como tendo) uma taxa de respiração máxima mais alta que a taxa de respiração basal, por exemplo, em aproximadamente 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, por exemplo, por um ensaio do Exemplo 77. O fusossoma compreende um nível de marcação de Anexina-V de no máximo 18.000, 17.000, 16.000, 15.000, 14.000, 13.000, 12.000, 11.000 ou 10.000 MFI, por exemplo, com o uso de um ensaio do Exemplo 56, ou em que o fusossoma compreende um nível de marcação de Anexina-V pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou

90% mais baixo que o nível de marcação de Anexina-V de um fusossoma semelhante tratado com menadiona no ensaio do Exemplo 56, ou em que o fusossoma compreende um nível de marcação de Anexina-V pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% menor que o nível de marcação de Anexina-V de um macrófago tratado com menadiona no ensaio do Exemplo 56. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) um nível de marcação de Anexina V que é pelo menos de aproximadamente 1%, 2%, 5% ou 10% menor que o nível de marcação de Anexina V de um fusossoma tratado de outra forma similar com antimicina A, por exemplo, em um ensaio do Exemplo 78. Em modalidades, o fusossoma compreende (ou é identificado como compreendendo) um nível de marcação de Anexina V que está dentro de aproximadamente 1%, 2%, 5% ou 10% do nível de marcação de Anexina V de um fusossoma, de outro modo semelhante tratado com Antimicina A, por exemplo, em um ensaio do Exemplo 78.

[00937] Em algumas modalidades, o fusossoma tem um nível de conteúdo de miRNA de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou maior que a célula de origem, por exemplo, por um ensaio do Exemplo 34. Em algumas modalidades, o fusossoma tem um nível de teor de miRNA de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do nível do teor de miRNA da célula de origem (por exemplo, até 100% do nível de conteúdo de miRNA da célula de origem), por exemplo, por um ensaio do Exemplo 34. Em algumas modalidades, o fusossoma tem um nível do teor total de RNA de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do nível do teor total de RNA da célula de origem (por exemplo, até 100% do nível do teor total da RNA da célula de origem), por exemplo, conforme medido por um ensaio do Exemplo 67.

[00938] Em algumas modalidades, o fusossoma tem uma razão proteína solúvel:não solúvel dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que a célula de origem, por exemplo, dentro de 1% a 2%, 2% a 3%, 3% a 4%, 4% a 5%, 5% a 10%, 10% a 20%, 20% a 30%, 30% a 40%, 40% a 50%, 50% a 60%, 60% a 70%, 70% a 80% ou 80% a 90% daquela da célula de origem, por exemplo, por um ensaio do Exemplo 39. Em modalidades, o fusossoma tem uma razão proteína solúvel:não solúvel de aproximadamente 0,3 a 0,8, 0,4 a 0,7 ou 0,5 a 0,6, por exemplo, de aproximadamente 0,563, por exemplo, por um ensaio do Exemplo 39. Em algumas modalidades, a população de fusossomas tem (ou é identificada como tendo) uma razão de massa entre proteína solúvel:insolúvel de aproximadamente 0,3 a 0,8, 0,4 a 0,7, 0,5 a 0,6 ou 0,563 ou maior que aproximadamente 0,1, 0,2, 0,3, 0,4 ou 0,5. Em algumas modalidades, a população de fusossomas tem (ou é identificada como tendo) uma razão de massa entre proteína solúvel:insolúvel que é maior que a das células de origem, por exemplo, pelo menos 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes ou 20 vezes maior. Nas modalidades, a razão em massa de proteína solúvel: insolúvel é determinada por um ensaio do Exemplo 71. Em modalidades, a razão de massa entre proteína solúvel:insolúvel é (ou é identificada como sendo) menor na população de fusossomas do que nas células parentais. Em modalidades, quando a razão de fusossomas para células parentais é (ou é identificada como sendo) de aproximadamente 3%, 4%, 5%, 6%, 7% ou 8%, a razão solúvel:insolúvel da população de fusossomas é (ou é identificado como sendo) aproximadamente igual à razão solúvel:insolúvel das células parentais.

[00939] Em algumas modalidades, o fusossoma tem um nível de LPS inferior a 5%, 1%, 0,5%, 0,01%, 0,005%, 0,0001%, 0,00001% ou menos do teor de LPS da célula de origem, por exemplo, conforme

medido por espectrometria de massa, por exemplo, em um ensaio do Exemplo 40. Em algumas modalidades, o fusossoma tem capacidade para transdução de sinal, por exemplo, transmitir um sinal extracelular, por exemplo, fosforilação AKT em resposta à insulina ou absorção de glicose (por exemplo, glicose identificada, por exemplo, 2-NBDG) em resposta à insulina, por exemplo, por pelo menos 1 %, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% a mais que um controle negativo, por exemplo, um fusossoma de outra forma similar na ausência de insulina, por exemplo, com o uso de um ensaio do Exemplo 50. Em algumas modalidades, o fusossoma tem como alvo um tecido, por exemplo, fígado, pulmões, coração, baço, pâncreas, trato gastrointestinal, rim, testículos, ovários, cérebro, órgãos reprodutivos, sistema nervoso central, sistema nervoso periférico, músculo esquelético, endotélio, orelha ou olho, quando administrado a um indivíduo, por exemplo, um camundongo, por exemplo, em que pelo menos 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50%, 60%, 70%, 80% ou 90% dos fusossomas em uma população de fusossomas administrados estão presentes no tecido-alvo após 24, 48 ou 72 horas, por exemplo, por um ensaio do Exemplo 65. Em algumas modalidades, o fusossoma tem um nível de sinalização de justácrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% a mais que o nível de sinalização de justacrina induzida por uma célula de referência, por exemplo, a célula de origem ou uma célula estromal da medula óssea (BMSC), por exemplo, por um ensaio do Exemplo 57. Em algumas modalidades, o fusossoma tem um nível de sinalização de justacrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% (por exemplo, até 100%) do nível de sinalização justácrina induzida por uma célula de referência, por exemplo, a célula de origem ou uma célula estromal da medula óssea

(BMSC), por exemplo, por um ensaio do Exemplo 57. Em algumas modalidades, o fusossoma tem um nível de sinalização de parácrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% a mais que o nível de sinalização de parácrina induzida por uma célula de referência, por exemplo, a célula de origem ou um macrófago, por exemplo, por um ensaio de Exemplo 58. Em algumas modalidades, o fusossoma tem um nível de sinalização de parácrina de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% (por exemplo, até 100%) do nível de sinalização de parácrina induzida por uma célula de referência, por exemplo, a célula de origem ou um macrófago, por exemplo, por um ensaio do Exemplo 58. Em algumas modalidades, o fusossoma polimeriza a actina em um nível dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% em comparação ao nível de actina polimerizada em uma célula de referência, por exemplo, a célula de origem ou uma célula C2C12, por exemplo, pelo ensaio do Exemplo 59. Em algumas modalidades, o fusossoma polimeriza actina (ou é identificado como polimerizando a actina) em um nível que é constante ao longo do tempo, por exemplo, ao longo de pelo menos 3, 5 ou 24 horas, por exemplo, por um ensaio do Exemplo 82. Nas modalidades, o nível de polimerização de actina muda em menos de 1%, 2%, 5%, 10% ou 20% durante um período de 5 horas, por exemplo, pelo ensaio do Exemplo 82. Em algumas modalidades, o fusossoma tem um potencial de membrana dentro de aproximadamente 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% do potencial de membrana de uma célula de referência, por exemplo, a célula de origem ou uma célula C2C12, por exemplo, por um ensaio do Exemplo 60 ou em que o fusossoma tem um potencial de membrana de aproximadamente -20 a -150 mV, -20 a -50 mV, -50 a -100 mV ou -100 a -150 mV ou em que o fusossoma tem um po-

tencial de membrana inferior a -1 mv, -5 mv, -10 mv, -20 mv, -30 mv, -40 mv, -50 mv, -60 mv, -70 mv, -80 mv, -90 mv, -100 mv. Em algumas modalidades, o fusossoma tem (ou é identificado como tendo) um potencial de membrana de aproximadamente -25 a -35, -27 a -32, -28 a -31, -29 a -30 ou -29,6 milivolts, por exemplo, em um ensaio do Exemplo 79. Em algumas modalidades, o fusossoma tem capacidade para extravasamento dos vasos sanguíneos, por exemplo, a uma taxa de pelo menos 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% a taxa de extravasamento da célula de origem, por exemplo, com o uso de um ensaio do Exemplo 45, por exemplo, em que a célula de origem é um neutrófilo, linfócito, célula B, macrófago ou célula NK. Em algumas modalidades, o fusossoma tem capacidade para quimiotaxia, por exemplo, de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% (por exemplo, até 100%) em comparação a uma célula de referência, por exemplo, um macrófago, por exemplo, com o uso de um ensaio do Exemplo 46. Em algumas modalidades, o fusossoma tem capacidade para fagocitose, por exemplo, pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% (por exemplo, até 100%) em comparação a uma célula de referência, por exemplo, um macrófago, por exemplo, com o uso de um ensaio do Exemplo 48. Em algumas modalidades, o fusossoma tem capacidade para atravessar uma membrana celular, por exemplo, uma membrana celular endotelial ou a barreira hematoencefálica. Em algumas modalidades, o fusossoma tem capacidade para secretar uma proteína, por exemplo, a uma taxa de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou 100% a mais que uma célula de referência, por exemplo, um fibroblasto embrionário de camundongo, por exemplo, com o uso de um ensaio do Exemplo 49. Em algumas modalidades, o fusossoma tem capacidade para secretar uma proteína, por exemplo, a uma taxa

de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% (por exemplo, até 100%) em comparação a uma célula de referência, por exemplo, um fibroblasto embrionário de camundongo, por exemplo, com o uso de um ensaio do Exemplo 49.

[00940] Em algumas modalidades, o fusossoma não tem capacidade para transcrição ou tem atividade transcricional inferior a 1%, 2,5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% daquela da atividade transcricional de uma célula de referência, por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 25. Em algumas modalidades, o fusossoma não tem capacidade para replicação de DNA nuclear ou tem replicação de DNA nuclear inferior a 1%, 2,5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% da replicação do DNA nuclear de uma célula de referência, por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 26. Em algumas modalidades, o fusossoma não tem cromatina ou tem um teor de cromatina inferior a 1%, 2,5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% do teor de cromatina de uma célula de referência, por exemplo, a célula de origem, por exemplo, com o uso de um ensaio do Exemplo 33.

[00941] Em algumas modalidades, uma característica de um fusossoma é descrita em comparação a uma célula de referência. Nas modalidades, a célula de referência é a célula de origem. Nas modalidades, a célula de referência é uma célula HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080 ou BJ. Em algumas modalidades, uma característica de uma população de fusossomas é descrita em comparação a uma população de células de referência, por exemplo, uma população de células de origem ou uma população de células HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080 ou BJ.

[00942] Em algumas modalidades, o fusossoma atende a um pa-

drão de práticas farmacêuticas ou de boas práticas de fabricação (GMP). Em algumas modalidades, o fusossoma foi produzido de acordo com boas práticas de fabricação (GMP). Em algumas modalidades, o fusossoma tem um nível de patógeno abaixo de um valor de referência predeterminado, por exemplo, está substancialmente livre de patógenos. Em algumas modalidades, o fusossoma tem um nível de contaminante abaixo de um valor de referência predeterminado, por exemplo, é substancialmente livre de contaminantes. Em algumas modalidades, o fusossoma tem baixa imunogenicidade, por exemplo, conforme descrito no presente documento.

[00943] Em algumas modalidades, a imunogenicidade de uma composição de fusossoma é avaliada por um ensaio de inativação do soro (por exemplo, um ensaio que detecta neutralização mediada por anticorpo ou degradação mediada por complemento). Em algumas modalidades, os fusossomas não são inativados pelo soro ou são inativados em um nível abaixo de um valor predeterminado. Em algumas modalidades, o soro de um indivíduo naïve de fusossoma (por exemplo, humano ou camundongo) é colocado em contato com uma composição de fusossoma de teste. Em algumas modalidades, o soro de um indivíduo que recebeu uma ou mais administrações de fusossomas, por exemplo, recebeu pelo menos duas administrações de fusossomas, é colocado em contato com a composição de fusossomas de teste. Em modalidades, os fusossomas expostos ao soro são, então, testados quanto à capacidade de entregar uma carga às células-alvo. Em algumas modalidades, a porcentagem de células que compreende de forma detectável a carga após o tratamento com fusossomas incubados com soro é de pelo menos 50%, 60%, 70%, 80%, 90% ou 95% da porcentagem de células que de forma detectável compreendem a carga após o tratamento com fusossomas de controle positivo não contatado com soro. Em algumas modalidades, a inativação do soro é

medida com o uso de um ensaio do Exemplo 100.

[00944] Em algumas modalidades, a imunogenicidade de uma composição de fusossomas é avaliada detectando-se a ativação do complemento em resposta aos fusossomas. Em algumas modalidades, os fusossomas não ativam o complemento ou ativam o complemento em um nível abaixo de um valor predeterminado. Em algumas modalidades, o soro de um indivíduo naïve de fusossoma (por exemplo, humano ou camundongo) é colocado em contato com uma composição de fusossoma de teste. Em algumas modalidades, o soro de um indivíduo que recebeu uma ou mais administrações de fusossomas, por exemplo, recebeu pelo menos duas administrações de fusossomas, é colocado em contato com a composição de fusossomas de teste. Em modalidades, a composição que compreende soro e fusossomas é, então, testada quanto a um fator de complemento ativado (por exemplo, C3a), por exemplo, por ELISA. Em algumas modalidades, um fusossoma que compreende uma modificação descrita no presente documento (por exemplo, níveis elevados de uma proteína regulatória do complemento em comparação com uma célula de referência) sofre ativação do complemento reduzida em comparação com um fusossoma semelhante que carece da modificação, por exemplo, reduzido em pelo menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% ou 99%. Em algumas modalidades, a ativação do complemento é medida com o uso de um ensaio de Exemplo 101.

[00945] Em algumas modalidades, um fusossoma ou população de fusossomas não será substancialmente inativado pelo soro. Em algumas modalidades, um fusossoma ou população de fusossomas é resistente à inativação do soro, por exemplo, conforme quantificado de acordo com o método descrito no Exemplo 100. Em modalidades, o fusossoma ou população de fusossomas não é substancialmente inativado por soro ou é resistente à inativação do soro após múltiplas ad-

ministrações do fusossoma ou população de fusossomas a um indivíduo, por exemplo, de acordo com os métodos descritos no presente documento. Em algumas modalidades, um fusossoma é modificado para ter uma inativação do soro reduzida, por exemplo, em comparação a um fusossoma não modificado correspondente, por exemplo, após múltiplas administrações do fusossoma modificado, por exemplo, conforme quantificado de acordo com o método descrito no Exemplo 100.

[00946] Em algumas modalidades, um fusossoma não induz substancialmente a atividade do complemento, por exemplo, conforme medido de acordo com o método descrito no Exemplo 101. Em algumas modalidades, um fusossoma é modificado para induzir a atividade do complemento reduzida em comparação a um fusossoma não modificado correspondente. Nas modalidades, a atividade do complemento é medida determinando-se a expressão ou atividade de uma proteína do complemento (por exemplo, DAF, proteínas que se ligam ao fator de aceleração do decaimento (DAF, CD55), por exemplo, proteína-1 semelhante ao fator H (FH) (FHL-1), proteína de ligação C4b (C4BP), receptor de complemento 1 (CD35), proteína cofator de membrana (MCP, CD46), Profectina (CD59), proteínas que inibem as enzimas CD/C5 convertase da via do complemento clássica e alternativa ou proteínas que regulam a montagem do MAC) em uma célula.

[00947] Em algumas modalidades, a célula de origem é uma célula endotelial, um fibroblasto, uma célula sanguínea (por exemplo, um macrófago, um neutrófilo, um granulócito, um leucócito), uma célula-tronco (por exemplo, uma célula-tronco mensequimal, uma célula-tronco do cordão umbilical, célula-tronco da medula óssea, uma célula-tronco hematopoiética, uma célula-tronco pluripotente induzida por exemplo, uma célula-tronco pluripotente induzida derivada de uma célula do indivíduo), uma célula-tronco embrionária (por exemplo, uma

célula-tronco de vesícula vitelina embrionária, placenta, cordão umbilical, pele fetal, pele adolescente, sangue, medula óssea, tecido de adipose, tecido eritropoiético, tecido hematopoiético), um mioblasto, uma célula parenquimal (por exemplo, hepatócito), uma célula alveolar, um neurônio (por exemplo, uma célula neuronal retiniana) uma célula precurosora (por exemplo, uma célula precursora retiniana, um mieloblasto, células precursoras mieloides, um timócito, um meiócito, um megacarioblasto, um promegacarioblasto, um melanoblasto, um linfoblasto, uma célula precursora da medula óssea, um normoblasto ou um angioblasto), uma célula progenitora (por exemplo, uma célula progenitora pancreática, uma célula satélite, uma célula da glia radial, uma célula estromal da medula óssea, uma célula progtenitora pancreática, uma célula progenitora endotelial, uma célula do blasto) ou uma célula imortalizada (por exemplo, célula HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080 ou BJ). Em algumas modalidades, a célula de origem é outra que não uma célula 293, célula HEK, célula endotelial humana ou uma célula epitelial humana, monócito, macrófago, célula dendrítica ou célula-tronco.

[00948] Em algumas modalidades, a célula de origem expressa (por exemplo, superexpressa) ARRDC1 ou um fragmento ativo ou variante deste. Em algumas modalidades, o fusossoma ou composição de fusossoma tem uma razão de fusogênio para ARRDC1 de aproximadamente 1 a 3, 1 a 10, 1 a 100, 3 a 10, 4 a 9, 5 a 8, 6 a 7, 15 a 100, 60 a 200, 80 a 180, 100 a 160, 120 a 140, 3 a 100, 4 a 100, 5 a 100, 6 a 100, 15 a 100, 80 a 100, 3 a 200, 4 a 200, 5 a 200, 6 a 200, 15 a 200, 80 a 200, 100 a 200, 120 a 200, 300 a 1000, 400 a 900, 500 a 800, 600 a 700, 640 a 690, 650 a 680, 660 a 670, 100 a 10.000 ou de aproximadamente 6.64,9, por exemplo, por um ensaio de espectrometria de massa. Em algumas modalidades, o nível de ARRDC1 como uma porcentagem do teor total da proteína é de pelo menos de aproxima-

damente 0,01%, 0,02%, 0,03%, 0,04%, 0,05%; 0,1%, 0,15%, 0,2%, 0,25%; 0,5%, 1%, 2%, 3%, 4%, 5%; ou o nível de ARRDC1 como uma porcentagem do teor total da proteína é de aproximadamente 0,05 a 1,5%, 0,1% a 0,3%, 0,05 a 0,2%, 0,1 a 0,2%, 0,25 a 7,5%, 0,5% a 1,5%, 0,25 a 1%, 0,5 a 1%, 0,05 a 1,5%, 10% a 30%, 5 a 20% ou 10 a 20%, por exemplo, por espectrometria de massa, por exemplo, conforme medido de acordo com o método descrito em Exemplo 99. Em algumas modalidades, o fusossoma ou composição de fusossoma tem uma razão de fusogênio para TSG101 de aproximadamente 100 a 1.000, 100 a 400, 100 a 500, 200 a 400, 200 a 500, 200 a 1.000, 300 a 400, 1.000 a 10.000, 2.000 a 5.000, 3.000 a 4.000, 3.050 a 3.100, 3.060 a 3.070 ou de aproximadamente 3.064, 10.000 a 100.000, 10.000 a 200.000, 10.000 a 500.000, 20.000 a 500.000, 30.000 a 400.000, por exemplo, com o uso de um ensaio de espectrometria de massa, por exemplo, um ensaio do Exemplo 95. Em algumas modalidades, o fusossoma ou composição de fusossoma tem uma razão de carga para tsg101 de aproximadamente 1 a 3, 1 a 30, 1 a 20, 1 a 25, 1,5 a 30, 10 a 30, 15 a 25, 18 a 21, 19 a 20, 10 a 300, 10 a 200, 15 a 300, 15 a 200, 100 a 300, 100 a 200, 150 a 300 ou de aproximadamente 19,5, por exemplo, com o uso de um ensaio da espectrometria de massa, por exemplo, um ensaio de Exemplo 96. Em algumas modalidades, o nível de TSG101 como uma porcentagem do teor total da proteína é de pelo menos de aproximadamente 0,0001%, 0,0002%, 0,0003%, 0,0004%, 0,0005%, 0,0006%, 0,0007%, 0,001%, 0,002%, 0,003%, 0,004 %, 0,005%, 0,006%, 0,007%; 0,01%, 0,02%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%; ou o nível de TSG101 como uma porcentagem do teor total da proteína é de aproximadamente 0,0001 a 0,001, 0,0001 a 0,002, 0,0001 a 0,01, 0,0001 a 0,1, 0,001 a 0,01, 0,002 a 0,006, 0,003 a 0,005, 0,001 a 0,1, 0,01 a 0,1, 0,02 a 0,06, 0,03 a 0,05 ou 0,004, por exemplo, por espectrometria de massa, por

exemplo, conforme medido de acordo com o método descrito no Exemplo 99.

[00949] Em algumas modalidades, o fusossoma compreende uma carga, por exemplo, um agente terapêutico, por exemplo, um agente terapêutico endógeno ou um agente terapêutico exógeno. Em algumas modalidades, o agente terapêutico é escolhido de uma ou mais de uma proteína, por exemplo, uma enzima, uma proteína transmembranar, um receptor, um anticorpo; um ácido nucléico, por exemplo, DNA, um cromossomo (por exemplo, um cromossomo artificial humano), RNA, mRNA, siRNA, miRNA ou uma pequena molécula. Em algumas modalidades, o agente terapêutico é uma organela diferente de uma mitocôndria, por exemplo, uma organela selecionada a partir de: núcleo, aparelho de Golgi, lisossoma, retículo endoplasmático, vacúolo, endossoma, acrossoma, autofagossomo, centríolo, glicossomo, glioxissoma, hidrogenossoma, melanossoma, mitossoma, cnidocisto, peroxissomo, proteassoma, vesícula e grânulo de tensão. Em algumas modalidades, a organela é uma mitocôndria.

[00950] Em algumas modalidades, o fusossoma entra na célula-alvo por endocitose, por exemplo, em que o nível de agente terapêutico administrado por meio de uma via endocítica é 0,01 a 0,6, 0,01 a 0,1, 0,1 a 0,3 ou 0,3 a 0,6 ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais que uma célula de referência tratada com cloroquina em contato com fusossomas semelhantes, por exemplo, com o uso de um ensaio do Exemplo 63. Em algumas modalidades, pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% dos fusossomas em uma composição de fusossoma que entra em uma célula-alvo entra por meio de uma via não endocítica, por exemplo, os fusossomas entram na célula-alvo por meio de fusão com a superfície celular. Em algumas modalidades, o nível de um agente terapêutico é administrado por meio de

uma via não endocítica

[00951] para um determinado fusossoma é 0,1 a 0,95, 0,1 a 0,2, 0,2 a 0,3, 0,3 a 0,4, 0,4 a 0,5, 0,5 a 0,6, 0,6 a 0,7, 0,7 a 0,8, 0,8 a 0,9, 0,9 a

[00952] 0,95, ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou

[00953] maior que uma célula de referência tratada com cloroquina, por exemplo, com o uso de um ensaio do Exemplo 62. Em algumas modalidades, pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% dos fusossomas em uma composição de fusossoma que entra em uma célula-alvo entra no citoplasma (por exemplo, não entra em um endossomo ou lisossoma). Em algumas modalidades, menos de 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2% ou 1% de fusossomas em uma composição de fusossoma que entra em uma célula-alvo entra em um endossoma ou lisossoma. Em algumas modalidades, o fusossoma entra na célula-alvo por uma via não endocítica, por exemplo, em que o nível de agente terapêutico administrado é de pelo menos 90%, 95%, 98% ou 99% de uma célula de referência tratada com cloroquina, por exemplo, com o uso de um ensaio do Exemplo 63. Em uma modalidade, um fusossoma entrega um agente a uma célula-alvo por meio de uma via mediada por dinamina. Em uma modalidade, o nível do agente entregue por meio de uma via mediada por dinamina está na faixa de 0,01 a 0,6 ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou maior que células-alvo tratadas com Dynasore em contato com fusossomas semelhantes, por exemplo, conforme medido em um ensaio do Exemplo 64. Em uma modalidade, um fusossoma entrega um agente a uma célula-alvo por meio de macropinocitose. Em uma modalidade, o nível de agente distribuído por meio de macropinocitose está na faixa de 0,01 a 0,6 ou pelo me-

nos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que as células-alvo tratadas com EIPA em contato com fusossomas semelhantes, por exemplo, conforme medido em um ensaio do Exemplo 64. Em uma modalidade, um fusossoma entrega um agente a uma célula-alvo por meio de uma via mediada por actina. Em uma modalidade, o nível do agente distribuído por meio de uma via mediada por actina estará na faixa de 0,01 a 0,6 ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que as células-alvo tratadas com Latrunculina B em contato com fusossomas semelhantes, por exemplo, conforme medido em um ensaio do Exemplo 64.

[00954] Em algumas modalidades, o fusossoma tem uma densidade de <1, 1 a 1,1, 1,05 a 1,15, 1,1 a 1,2, 1,15 a 1,25, 1,2 a 1,3, 1,25 a 1,35 ou >1,35, por exemplo, por um ensaio do Exemplo 31.

[00955] Em algumas modalidades, a composição de fusossoma compreende menos de 0,01%, 0,05%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 4%, 5% ou 10% de células de origem por massa de proteína ou menos de 0,01%, 0,05%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 4%, 5% ou 10% das células têm um núcleo funcional. Em algumas modalidades, pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% ou 99% dos fusossomas na composição de fusossomas compreendem uma organela, por exemplo, uma mitocôndria.

[00956] Em algumas modalidades, o fusossoma compreende adicionalmente um agente terapêutico exógeno. Em algumas modalidades, o agente terapêutico exógeno é escolhido a partir de uma ou mais de uma proteína, por exemplo, uma enzima, uma proteína transmembranar, um receptor, um anticorpo; um ácido nucléico, por exemplo, DNA, um cromossomo (por exemplo, um cromossomo artificial humano), RNA, mRNA, siRNA, miRNA ou uma pequena molécula.

[00957] Em modalidades, o fusossoma entra na célula por endoci-

tose ou por uma via não endocítica.

[00958] Em modalidades, a composição de fusossoma é estável a uma temperatura inferior a 4 °C por pelo menos 1, 2, 3, 6 ou 12 horas; 1, 2, 3, 4, 5 ou 6 dias; 1, 2, 3 ou 4 semanas; 1, 2, 3 ou 6 meses; ou 1, 2, 3, 4 ou 5 anos. Em modalidades, a composição de fusossoma é estável a uma temperatura inferior a -20 °C por pelo menos 1, 2, 3, 6 ou 12 horas; 1, 2, 3, 4, 5 ou 6 dias; 1, 2, 3 ou 4 semanas; 1, 2, 3 ou 6 meses; ou 1, 2, 3, 4 ou 5 anos. Em modalidades, a composição de fusossoma é estável a uma temperatura inferior a -80 °C por pelo menos 1, 2, 3, 6 ou 12 horas; 1, 2, 3, 4, 5 ou 6 dias; 1, 2, 3 ou 4 semanas; 1, 2, 3 ou 6 meses; ou 1, 2, 3, 4 ou 5 anos.

[00959] Em modalidades, o fusossoma tem um tamanho ou a população de fusossomas tem um tamanho médio, dentro de aproximadamente 0,01%, 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% da célula de origem, por exemplo, conforme medido por um ensaio do Exemplo 28. Em modalidades, o fusossoma tem um tamanho ou a população de fusossomas tem um tamanho médio que é inferior a aproximadamente 0,01%, 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% da célula de origem, por exemplo, conforme medido por um ensaio do Exemplo 28. Em modalidades, os fusossomas têm (ou são identificados como tendo) um tamanho menor que as células parentais. Em modalidades, os fusossomas têm (ou são identificados como tendo) um tamanho dentro de aproximadamente 50%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 80% ou 90% de células parentais. Em modalidades, os fusossomas têm (ou são identificados como tendo) menos que de aproximadamente 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 1% ou menos da variabilidade da célula parental na distribuição de tamanho, por exemplo, dentro de aproximadamente 90% da amostra. Em modalidades, os fusossomas têm (ou são identi-

ficados como tendo) aproximadamente 40%, 45%, 50%, 55%, 56%, 57%, 58%, 59%, 60%, 65% ou 70% menos da variabilidade da célula parental na distribuição de tamanho, por exemplo, dentro de aproximadamente 90% da amostra. Em algumas modalidades, os fusossomas têm (ou são identificados como tendo) um tamanho médio maior que 30, 35, 40, 45, 50, 55, 60, 65 ou 70 nm de diâmetro. Em modalidades, os fusossomas têm um tamanho médio de aproximadamente 100, 110, 120, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 140 ou 150 nm de diâmetro. Nas modalidades, o fusossoma tem um tamanho, ou a população de fusossomas tem um tamanho médio, dentro de aproximadamente 0,01% -0,05%, 0,05% -0,1%, 0,1% -0,5%, 0,5% - 1%, 1% -2%, 2% -3%, 3% -4%, 4% -5%, 5% -10%, 10% -20%, 20% -30%, 30% -40%, 40% -50%, 50% -60%, 60% -70%, 70% -80% ou 80% -90% do tamanho da célula de origem, por exemplo, conforme medido por um ensaio do Exemplo 28. Nas modalidades, o fusossoma tem um tamanho, ou a população de fusossomas tem um tamanho médio, que é inferior a de aproximadamente 0,01% -0,05%, 0,05% - 0,1%, 0,1% -0,5%, 0,5% - 1%, 1% -2%, 2% -3%, 3% -4%, 4% -5%, 5% -10%, 10% -20%, 20% -30%, 30% -40%, 40% -50%, 50% -60%, 60% - 70%, 70% -80% ou 80% -90% do tamanho da célula de origem, por exemplo, conforme medido por um ensaio do Exemplo 28. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio inferior a de aproximadamente 500 nm (por exemplo, inferior a de aproximadamente 10, 50, 100, 150, 200, 250, 300, 350, 400 ou 450 nm), por exemplo, conforme medido por um ensaio do Exemplo 70. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio de aproximadamente 80 a 180, 90 a 170, 100 a 160, 110 a 150, 120 a 140 ou 130 nm, por exemplo, conforme medido por um ensaio do Exemplo 70. Nas modalidades, o fusossoma tem um diâmetro ou a população de

fusossomas tem um diâmetro médio, de aproximadamente 11.000 nm e 21.000 nm, por exemplo, conforme medido por um ensaio do Exemplo 70. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio, entre de aproximadamente 10 a 22.000, 12 a 20.000, 14 a 18.720 nm, 20 a 16.000 nm, por exemplo, conforme medido por um ensaio do Exemplo 70. Nas modalidades, o fusossoma tem um volume ou a população de fusossomas tem um volume médio, de aproximadamente 0,01 a 0,1 μm^3, 0,02 a 1 μm^3, 0,03 a 1 μm^3, 0,04 a 1 μm^3, 0,05 a 0,09 μm^3, 0,06 a 0,08 μm^3, 0,07 μm^3, por exemplo, conforme medido por um ensaio do Exemplo 70. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio de pelo menos de aproximadamente 10 nm, 20 nm, 30 nm, 40 nm, 50 nm, 60 nm, 70 nm, 80 nm, 90 nm, 100 nm, 150 nm, 200 nm ou 250 nm, por exemplo, conforme medido por um ensaio do Exemplo 30. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio, de aproximadamente 10 nm, 20 nm, 30 nm, 40 nm, 50 nm, 60 nm, 70 nm, 80 nm, 90 nm, 100 nm, 150 nm, 200 nm ou 250 nm (por exemplo, ± 20%), por exemplo, conforme medido por um ensaio do Exemplo 30. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio de pelo menos de aproximadamente 500 nm, 750 nm, 1.000 nm, 1.500 nm, 2.000 nm, 2.500 nm, 3.000 nm, 5.000 nm, 10.000 nm, ou 20.000 nm, por exemplo, conforme medido por um ensaio do Exemplo 30. Nas modalidades, o fusossoma tem um diâmetro ou a população de fusossomas tem um diâmetro médio, de aproximadamente 500 nm, 750 nm, 1.000 nm, 1.500 nm, 2.000 nm, 2.500 nm, 3.000 nm, 5.000 nm, 10.000 nm ou 20.000 nm (por exemplo, ± 20%), por exemplo, conforme medido por um ensaio do Exemplo 30. Nas modalidades, a população de fusossomas tem (ou é identificada como tendo) um ou mais dentre: um

diâmetro de quantil de 10% de aproximadamente 40 a 90 nm, 45 a 60 nm, 50 a 55 nm ou 53 nm; um diâmetro 25% quantil de aproximadamente 70 a 100 nm, 80 a 95 nm, 85 a 90 nm ou 88 nm; um diâmetro 75% quantil de aproximadamente 200 a 250 nm, 210 a 240 nm, 220 a 230 nm ou 226 nm; ou um quantil de 90% de aproximadamente 4.000 a 5.000 nm, 4.300 a 4.600 nm, 4.400 a 4.500 nm, 4.450 nm, por exemplo, por um ensaio do Exemplo 69.

[00960] Nas modalidades, a composição de fusossoma compreende (ou é identificada como compreendendo) uma concentração de GAPDH de aproximadamente 35 a 40, 36 a 39, 37 a 38 ou 37,2 ng/ml, por exemplo, em um ensaio do Exemplo 83. Nas modalidades, a concentração de GAPDH da composição de fusossoma está (ou é identificada como sendo) dentro de aproximadamente 1%, 2%, 5%, 10% ou 20% da concentração de GAPDH das células de origem, por exemplo, em um ensaio do Exemplo 83. Nas modalidades, a concentração de GAPDH da composição de fusossoma é (ou é identificada como sendo) pelo menos 1%, 2%, 5%, 10% ou 20% menor do que a concentração de GAPDH das células de origem, por exemplo, em um ensaio do Exemplo 83. Nas modalidades, a composição de fusossoma compreende (ou é identificada como compreendendo) menos que aproximadamente 30, 35, 40, 45, 46, 47, 48, 49, 50, 55, 60, 65 ou 70 µg de GAPDH por grama de proteína total. Nas modalidades, a composição de fusossoma compreende (ou é identificada como compreendendo) menos de aproximadamente 500, 250, 100 ou 50 µg de GAPDH por grama de proteína total. Nas modalidades, a célula parental compreende (ou é identificada como compreendendo) pelo menos 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 30%, 50% ou mais GAPDH por proteína total do que a composição fusossoma.

[00961] Nas modalidades, o teor fracionário médio de calnexina no fusossoma é (ou é identificando como sendo) inferior a aproximada-

mente $1x10^{-4}$, $1,5x10^{-4}$, $2x10^{-4}$, $2,1x10^{-4}$, $2,2x10^{-4}$, $2,3x10^{-4}$, $2,4x10^{-4}$, $2,43x10^{-4}$, $2,5x10^{-4}$, $2,6x10^{-4}$, $2,7x10^{-4}$, $2,8x10^{-4}$, $2,9x10^{-4}$, $3x10^{-4}$, $3,5x10^{-4}$ ou $4x10^{-4}$. Nas modalidades, o fusossoma compreende uma quantidade de calnexina por proteína total que é inferior à da célula parental em aproximadamente 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99% ou mais.

[00962] Em algumas modalidades, os fusossomas compreendem ou são enriquecidos para lipídios que afetam a curvatura da membrana (consultar, por exemplo, Thiam et al., Nature Reviews Molecular Cell Biology, 14 (12): 775 a 785, 2013). Alguns lipídios têm um pequeno grupo de cabeça hidrofílica e grandes caudas hidrofóbicas, que facilitam a formação de um poro de fusão por concentração em uma região local. Em algumas modalidades, os fusossomas compreendem ou são enriquecidos em lipídios de curvatura negativa, como colesterol, fosfatidiletanolamina (PE), diglicerídeo (DAG), ácido fosfatídico (PA), ácido graxo (FA). Em algumas modalidades, os fusossomas não compreendem, são depletados de ou têm poucos lipídios de curvatura positiva, tais como lisofosfatidilcolina (LPC), fosfatidilinositol (PtdIns), ácido lisofosfatídico (LPA), lisofosfatidiletanolamina (LPE), monoacilglicerol (MAG).

[00963] Em algumas modalidades, os lipídios são adicionados a um fusossoma. Em algumas modalidades, os lipídios são adicionados às células de origem em cultura que incorporam os lipídios nas membranas das mesmas antes ou durante a formação de um fusossoma. Em algumas modalidades, os lipídios são adicionados às células ou fusossomas na forma de um lipossoma. Em algumas modalidades, metilbetaciclodextrano (mβ-CD) é usado para enriquecer ou depletar lipídios (consultar, por exemplo, Kainu et al, Journal of Lipid Research, 51(12): 3.533 a 3.541, 2010).

COMPOSIÇÕES FARMACÊUTICAS E MÉTODOS PARA PREPA-

RAR AS MESMAS

[00964] Em algumas modalidades, uma ou mais unidades de transdução de vetor retroviral são administradas ao indivíduo. Em algumas modalidades, pelo menos 1, 10, 100, 1.000, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13} ou 10^{14} unidades de transdução por kg são administradas ao indivíduo. Em algumas modalidades, pelo menos 1, 10, 100, 1.000, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13} ou 10^{14} unidades de transdução por células-alvo por ml de sangue são administradas ao indivíduo.

CONCENTRAÇÃO E PURIFICAÇÃO DE LENTIVÍRUS

[00965] Em algumas modalidades, uma formulação de vetor retroviral descrita no presente documento pode ser produzida por um processo que compreende um ou mais dentre, por exemplo, todas as seguintes etapas (i) a (vi), por exemplo, em ordem cronológica:

[00966] (i) cultivar células que produzem vetor retroviral;

[00967] (ii) colher o vetor retroviral que contém o sobrenadante;

[00968] (iii) opcionalmente clarificar o sobrenadante;

[00969] (iv) purificar o vetor retroviral para gerar uma preparação de vetor retroviral;

[00970] (v) opcionalmente, esterilizar por filtração a preparação do vetor retroviral; e

[00971] (vi) concentrar a preparação do vetor retroviral para produzir o produto final volumoso.

[00972] Em algumas modalidades, o processo não compreende a etapa de clarificação (iii). Em outras modalidades, o processo inclui a etapa de clarificação (iii). Em algumas modalidades, a etapa (vi) é realizada com o uso de ultrafiltração ou filtração de fluxo tangencial, com mais preferência, ultrafiltração de fibra oca. Em algumas modalidades, o método de purificação na etapa (iv) é cromatografia de troca iônica, com mais preferência, cromatografia de troca aniônica. Em algumas

modalidades, a esterilização por filtro na etapa (v) é realizada com o uso de um filtro de esterilização de 0,22 μm ou 0,2 μm. Em algumas modalidades, a etapa (iii) é realizada por clarificação de filtro. Em algumas modalidades, a etapa (iv) é realizada com o uso de um método ou uma combinação de métodos selecionados a partir de cromatografia, ultrafiltração/diafiltração ou centrifugação. Em algumas modalidades, o método de cromatografia ou uma combinação de métodos é selecionada a partir de cromatografia de troca iônica, cromatografia por interação hidrofóbica, cromatografia por exclusão de tamanhos, cromatografia de afinidade, cromatografia de fase reversa e cromatografia de afinidade de íon metálico imobilizado. Em algumas modalidades, o método de centrifugação é selecionado a partir de centrifugação zonal, centrifugação isopícnica e centrifugação de peletização. Em algumas modalidades, o método de ultrafiltração/diafiltração é selecionado a partir de diafiltração de fluxo tangencial, diafiltração de células agitadas e diálise. Em algumas modalidades, pelo menos uma etapa é incluída no processo para degradar o ácido nucleico para melhorar a purificação. Em algumas modalidades, a dita etapa é o tratamento com nuclease.

[00973] Em algumas modalidades, a concentração dos vetores é feita antes da filtração. Em algumas modalidades, a concentração dos vetores é feita após a filtração. Em algumas modalidades, as etapas de concentração e filtração são repetidas.

[00974] Em algumas modalidades, a etapa final de concentração é realizada após a etapa de esterilização por filtro. Em algumas modalidades, o processo é um processo em grande escala para a produção de formulações de grau clínico que são adequadas para administração a seres humanos como agentes terapêuticos. Em algumas modalidades, a etapa de esterilização por filtro ocorre antes de uma etapa de concentração. Em algumas modalidades, a etapa de concentração é a

etapa final do processo, e a etapa de esterilização por filtro é a penúltima etapa do processo. Em algumas modalidades, a etapa de concentração é realizada com o uso de ultrafiltração, de preferência, filtração de fluxo tangencial, com mais preferência, ultrafiltração de fibra oca. Em algumas modalidades, a etapa de esterilização por filtro é realizada com o uso de um filtro de esterilização com um tamanho de poro máximo de aproximadamente 0,22 μm. Em outra modalidade preferencial, o tamanho máximo do poro é de 0,2 μm.

[00975] Em algumas modalidades, a concentração do vetor é inferior ou igual a aproximadamente 4,6 x 10^{11} cópias de genoma de RNA de por ml de preparação antes do filtro de esterilização. O nível de concentração apropriado pode ser alcançado controlando a concentração do vetor usando, por exemplo, uma etapa de diluição, se apropriado. Desse modo, em algumas modalidades, uma preparação de vetor retroviral é diluída antes da esterilização por filtro.

[00976] A clarificação pode ser feita por uma etapa de filtração, removendo detritos celulares e outras impurezas. Filtros adequados podem utilizar filtros de celulose, fibras de celulose regeneradas, fibras de celulose combinadas com auxiliares de filtro inorgânicos (por exemplo, terra diatomácea, perlita, sílica pirogênica), filtros de celulose combinados com auxiliares de filtro inorgânicos e resinas orgânicas, ou qualquer combinação dos mesmos, e filtros poliméricos (exemplos incluem, porém sem limitação, náilon, polipropileno, polietersulfona) para obter uma remoção eficaz e recuperações aceitáveis. Um processo de múltiplos estágios pode ser usado. Um exemplo de processo de dois ou três estágios consistiria em um filtro (ou filtros) grosso para remover precipitado (ou precipitados) grande e detritos celulares, seguido por filtro (ou filtros) de segundo estágio de polimento com tamanhos de poros nominais maiores que 0,2 mícron, porém menores que 1 mícron. A combinação ideal pode ser uma função da distribuição do

tamanho do precipitado, bem como de outras variáveis. Além disso, as operações de estágio único que usam um filtro de tamanho de poro relativamente pequeno ou centrifugação também podem ser usadas para clarificação. De forma mais geral, qualquer abordagem de clarificação incluindo, porém sem limitação, filtração sem saída, microfiltração, centrifugação ou alimentação corporal de auxiliares de filtro (por exemplo, terra diatomácea) em combinação com filtração sem saída ou de profundidade, que fornece um filtrado de clareza adequada para não sujar a membrana e/ou resinas nas etapas subsequentes, será aceitável para uso na etapa de clarificação da presente invenção.

[00977] Em algumas modalidades, são usadas a filtração profunda e a filtração por membrana. Produtos comercialmente disponíveis úteis para tal propósito são mencionados, por exemplo, no documento WO 03/097797, página 20 a 21. As membranas que podem ser usadas podem ser compostas de diferentes materiais podem ser diferentes no tamanho dos poros e podem ser usadas em combinações. As mesmas podem ser obtidas comercialmente junto a vários fornecedores. Em algumas modalidades, o filtro usado para clarificação está na faixa de 1,2 a 0,22 µm. Em algumas modalidades, o filtro usado para clarificação é um filtro de 1,2/0,45 µm ou um filtro assimétrico com um tamanho de poro nominal mínimo de 0,22 µm.

[00978] Em algumas modalidades, o método usa nuclease para degradar DNA/RNA contaminante, ou seja, principalmente ácidos nucleicos de células hospedeiras. Nucleases exemplificativas adequadas para uso na presente invenção incluem Benzonase® Nuclease (EP 0229866) que ataca e degrada todas as formas de DNA e RNA (fita simples, fita dupla linear ou circular) ou qualquer outra DNase e/ou RNase comumente usada na técnica com a finalidade de eliminar DNA e/ou RNA indesejável ou contaminante de uma preparação. Nas modalidades preferenciais, a nuclease é Benzonase® Nuclease que hi-

drolisa rapidamente ácidos nucleicos por hidrólise de ligações fosfodiéster internas entre nucleotídeos específicos, reduzindo assim o tamanho dos polinucleotídeos no vetor que contém o sobrenadante. A Benzonase® Nuclease pode ser obtida comercialmente junto à Merck KGaA (código W214950). A concentração na qual a nuclease é usada está, de preferência, na faixa de 1 a 100 unidades/ml.

[00979] Em algumas modalidades, a suspensão de vetor é submetida à ultrafiltração (também denominada de diafiltração quando usada para troca de tampão) pelo menos uma vez durante o processo, por exemplo, para concentrar a troca de vetor e/ou de tampão. O processo usado para concentrar o vetor pode incluir qualquer processo de filtração (por exemplo, ultrafiltração (UF)) em que a concentração do vetor é aumentada forçando-se o diluente a passar por um filtro de tal maneira que o diluente seja removido da preparação do vetor, enquanto o vetor não consegue passar pelo filtro e, assim, permanece, na forma concentrada, na preparação do vetor. A UF é descrita detalhadamente, por exemplo, Microfiltration and Ultrafiltration: Principles and Applications, L. Zeman e A. Zydney (Marcel Dekker, Inc., Nova York, NY, 1996); e em: Ultrafiltration Handbook, Munir Cheryan (Technomic Publishing, 1986; nº ISBN 87762-456-9). Um processo de filtração adequado é a Filtração de Fluxo Tangencial (“TFF”) conforme descrito, por exemplo, no catálogo MILLIPORE intitulado “Pharmaceutical Process Filtration Catalog” páginas 177 a 202 (Bedford, Mass., 1995/96). A TFF é amplamente usada na indústria de bioprocessamento para coleta, clarificação, purificação e concentração de células de produtos, incluindo vírus. O sistema é composto por três fluxos de processo distintos: a solução de alimentação, o permeado e o retentado. Dependendo da aplicação, filtros com diferentes tamanhos de poros podem ser usados. Em algumas modalidades, o retentado contém o produto (vetor lentiviral). A membrana de ultrafiltração particular selecionada pode ter

um tamanho de poro suficientemente pequeno para reter o vetor, porém grande o suficiente para limpar com eficácia as impurezas. Dependendo do fabricante e do tipo de membrana, para vetores retrovirais, pontos de corte de peso molecular nominal (NMWC) entre 100 e 1.000 kDa podem ser apropriados, por exemplo, membranas com NMWC de 300 kDa ou 500 kDa. A composição da membrana pode ser, porém sem limitação, celulose regenerada, polietersulfona, polissulfona ou derivados das mesmas. As membranas podem ser folhas planas (também chamadas de telas planas) ou fibras ocas. Uma UF adequada é a UF de fibra oca, por exemplo, filtração usando filtros com um tamanho de poro menor que 0,1 μm. Os produtos são geralmente retidos, enquanto o volume pode ser reduzido através da permeação (ou ser mantido constante durante a diafiltração pela adição de tampão com a mesma velocidade com a qual o permeado, que contém tampão e impurezas, é removido no lado do permeado).

[00980] As duas geometrias usadas mais amplamente para TFF na indústria biofarmacêutica são placas e estruturas (telas planas) e módulos de fibra oca. As unidades de fibra oca para ultrafiltração e microfiltração foram desenvolvidas pela Amicon e Ramicon no início dos anos 1970 (Cheryan, M. Ultrafiltration Handbook), embora agora haja vários fornecedores, incluindo Spectrum e GE Healthcare. Os módulos de fibra oca consistem em uma série de fibras autossustentáveis com uma densa camada de revestimento. Os diâmetros das fibras variam de 0,5 mm a 3 mm. Em certas modalidades, fibras ocas são usadas para TFF. Em certas modalidades, são utilizadas fibras ocas de tamanho de poro de 500 kDa (0,05 μm). A ultrafiltração pode compreender a diafiltração (DF). Os microssolutos podem ser removidos adicionando-se solvente à solução que está sendo ultrafiltrada a uma taxa igual à taxa de UF. Isso lava as microespécies da solução em um volume constante, purificando o vetor retido.

[00981] UF/DF podem ser usadas para concentrar e/ou trocar o tampão das suspensões de vetor em diferentes estágios do processo de purificação. O método pode utilizar uma etapa DF para trocar o tampão do sobrenadante após cromatografia ou outras etapas de purificação, porém a mesma também pode ser usada antes da cromatografia.

[00982] Em algumas modalidades, o eluato da etapa de cromatografia é concentrado e posteriormente purificado por ultrafiltração-diafiltração. Durante esse processo, o vetor é trocado por tampão de formulação. A concentração até a concentração final desejada pode ocorrer após a etapa de esterilização por filtro. Após a dita filtração estéril, a substância esterilizada por filtro é concentrada por UF asséptico para produzir o produto vetorial a granel.

[00983] Nas modalidades, a ultrafiltração/diafiltração pode ser diafiltração de fluxo tangencial, diafiltração de células agitadas e diálise.

[00984] As técnicas de purificação tendem a envolver a separação das partículas do vetor do meio celular e, se necessário, a purificação adicional das partículas do vetor. Um ou mais dentre uma variedade de métodos cromatográficos podem ser usados para essa purificação. Cromatografia de troca iônica, e mais particularmente troca aniônica, é um método adequado e outros métodos podem ser usados. Uma descrição de algumas técnicas cromatográficas é fornecida abaixo.

[00985] A cromatografia de troca iônica se baseia no fato de que espécies carregadas, como biomoléculas e vetores virais, podem se ligar reversivelmente a uma fase estacionária (como uma membrana, ou então o empacotamento em uma coluna) que tem, fixados em sua superfície, grupos que possuem uma carga oposta. Há dois tipos de trocadores de íons. Trocadores de ânions são fases estacionárias que carregam grupos com carga positiva e, portanto, podem ligar espécies com carga negativa. Os trocadores de cátions carregam grupos com

carga negativa e, portanto, podem se ligar a espécies com carga positiva. O pH do meio tem influência sobre isso, pois pode alterar a carga de uma espécie. Assim, para uma espécie como uma proteína, caso o pH esteja acima do pI, a carga líquida será negativa, enquanto abaixo do pI, a carga líquida será positiva.

[00986] O deslocamento (eluição) das espécies ligadas pode ser efetuado pelo uso de tampões adequados. Assim, comumente, a concentração iônica do tampão é aumentada até que a espécie seja deslocada por meio da competição de íons tampão pelos locais iônicos na fase estacionária. Um método alternativo de eluição envolve a mudança do pH do tampão até que a carga líquida da espécie não favoreça mais o desvio para a fase estacionária. Um exemplo seria reduzir o pH até que a espécie assuma uma carga positiva líquida e não se ligue mais a um trocador de ânions.

[00987] Alguma purificação pode ser alcançada se as impurezas não estiverem carregadas, ou então se as mesmas carregarem uma carga de sinal oposto ao da espécie desejada, porém o mesmo sinal daquele do trocador de íons. Isso ocorre porque as espécies sem carga e aquelas que têm uma carga do mesmo sinal daquele trocador de íons normalmente não se ligam. Para diferentes espécies ligadas, a força da ligação varia com fatores como a densidade de carga e a distribuição de cargas nas várias espécies. Assim, aplicando um gradiente iônico ou de pH (como um gradiente contínuo, ou como uma série de etapas), a espécie desejada pode ser eluída separadamente das impurezas.

[00988] A cromatografia de exclusão de tamanho é uma técnica que separa as espécies de acordo com seu tamanho. Normalmente, isso é realizado pelo uso de uma coluna cheia de partículas com poros de um tamanho bem definido. Para a separação cromatográfica, são escolhidas partículas que tenham tamanhos de poros adequados em re-

lação aos tamanhos das espécies na mistura a ser separada. Quando a mistura é aplicada, como uma solução (ou suspensão, no caso de um vírus), à coluna e então eluída com tampão, as partículas maiores irão eluir primeiro, pois têm acesso limitado (ou nenhum acesso) aos poros. Partículas menores eluirão mais tarde, pois podem entrar nos poros e, portanto, percorrer um caminho mais longo através da coluna. Assim, ao considerar o uso de cromatografia de exclusão de tamanho para a purificação de vetores virais, seria de se esperar que o vetor fosse eluído antes de impurezas menores, como proteínas.

[00989] As espécies, tais como as proteínas, têm em suas superfícies regiões hidrofóbicas que podem se ligar de maneira reversível a locais de baixa hidrofobicidade em uma fase estacionária. Em meios com uma concentração de sal relativamente alta, essa ligação é promovida. Normalmente, em HIC, a amostra a ser purificada é ligada à fase estacionária em um ambiente com alto teor de sal. Em seguida, a eluição é obtida pela aplicação de um gradiente (contínuo ou como uma série de etapas) de concentração decrescente de sal. Um sal comumente usado é o sulfato de amônio. As espécies com diferentes níveis de hidrofobicidade tenderão a ser eluídas em diferentes concentrações de sal e, portanto, a espécie alvo pode ser purificada de impurezas. Outros fatores, tais como pH, temperatura e aditivos ao meio de eluição, como detergentes, sais caotrópicos e orgânicos também podem influenciar a força de ligação das espécies às fases estacionárias de HIC. Um ou mais desses fatores podem ser ajustados ou utilizados para otimizar a eluição e purificação do produto.

[00990] Os vetores virais têm em sua superfície, porções químicas hidrofóbicas, tais como proteínas, e, portanto, o HIC pode ser potencialmente usado como meio de purificação.

[00991] Como o HIC, o RPC separa as espécies de acordo com as diferenças em suas hidrofobicidades. É usada uma fase estacionária

de maior hidrofobicidade do que aquela usada em HIC. A fase estacionária geralmente consiste em um material, tipicamente sílica, ao qual estão ligadas porções químicas hidrofóbicas, tais como grupos alquila ou grupos fenila. Alternativamente, a fase estacionária pode ser um polímero orgânico, sem grupos anexados. A amostra que contém a mistura de espécies a ser resolvida é aplicada à fase estacionária em um meio aquoso de polaridade relativamente alta que promove a ligação. A eluição é então alcançada reduzindo-se a polaridade do meio aquoso pela adição de um solvente orgânico, como isopropanol ou acetonitrila. Normalmente, um gradiente (contínuo ou como uma série de etapas) de concentração crescente de solvente orgânico é usado e as espécies são eluídas na ordem de suas respectivas hidrofobicidades.

[00992] Outros fatores, como o pH do meio de eluição e o uso de aditivos, também podem influenciar a força de ligação das espécies às fases estacionárias RPC. Um ou mais desses fatores podem ser ajustados ou utilizados para otimizar a eluição e purificação do produto. Um aditivo comum é o ácido trifluoroacético (TFA). Isso suprime a ionização de grupos ácidos, tais como porções químicas carboxila na amostra. Isso também reduz o pH do meio de eluição e suprime a ionização de grupos silanol livres que podem estar presentes na superfície de fases estacionárias que têm uma matriz de sílica. O TFA faz parte de uma classe de aditivos conhecidos, tais como agentes de emparelhamento iônico. Esses interagem com grupos iônicos, presentes nas espécies da amostra, que carregam uma carga oposta. A interação tende a mascarar a carga, aumentando a hidrofobicidade da espécie. Os agentes de emparelhamento de íons aniônicos, tais como TFA e ácido pentafluoropropiônico interagem com grupos carregados positivamente em uma espécie. Os agentes de emparelhamento de íons catiônicos, como a trietilamina, interagem com grupos carregados ne-

gativamente.

[00993] Os vetores virais têm em sua superfície porções químicas hidrofóbicas, como proteínas e, portanto, RPC, potencialmente, pode ser usado como um meio de purificação.

[00994] A cromatografia de afinidade se baseia no fato de que determinados ligantes que se ligam especificamente a biomoléculas, tais como proteínas ou nucleotídeos, podem ser imobilizados em uma fase estacionária. A fase estacionária modificada pode então ser usada para separar a biomolécula relevante de uma mistura. Exemplos de ligantes altamente específicos são anticorpos, para a purificação de antígenos alvo e inibidores de enzimas para a purificação de enzimas. Interações mais gerais também podem ser utilizadas, tais como o uso do ligante de proteína A para o isolamento de uma ampla gama de anticorpos.

[00995] Normalmente, a cromatografia de afinidade é realizada pela aplicação de uma mistura que contém as espécies de interesse, à fase estacionária que tem o ligante relevante anexado. Sob condições apropriadas, isso causará a ligação da espécie à fase estacionária. Os componentes não ligados são, então, lavados antes da aplicação de um meio de eluição. O meio de eluição é escolhido para interromper a ligação do ligante à espécie-alvo. Isso é comumente alcançado pela escolha de uma força iônica apropriada, pH ou pelo uso de substâncias que irão competir com as espécies-alvo por locais ligantes. Para algumas espécies ligadas, um agente caotrópico, como a ureia, é usado para efetuar o deslocamento do ligante. Isso, no entanto, pode resultar em desnaturação irreversível da espécie.

[00996] Os vetores virais têm em sua superfície porções químicas como proteínas, que podem ter capacidade para se ligar especificamente a ligantes apropriados. Isso significa que, potencialmente, a cromatografia de afinidade pode ser usada em seu isolamento.

[00997] Biomoléculas, tais como proteínas, podem ter em sua superfície porções químicas doadoras de elétrons que podem formar ligações coordenadas com íons metálicos. Isso pode facilitar sua ligação a fases estacionárias que transportam íons metálicos imobilizados, como Ni^{2+}, Cu^{2+}, Zn^{2+} ou Fe^{3+}. As fases estacionárias usadas no IMAC têm agentes quelantes, tipicamente ácido nitriloacético ou ácido iminodiacético covalentemente ligado à sua superfície, e é o agente quelante que retém o íon metálico. É necessário que o íon metálico quelado tenha pelo menos um sítio de coordenação disponível para formar uma ligação coordenada a uma biomolécula. Potencialmente, existem várias metades na superfície das biomoléculas que podem ter capacidade para se ligarem ao íon metálico imobilizado. Estes incluem resíduos de histidina, triptofano e cisteína, bem como grupos fosfato. Para proteínas, no entanto, o doador predominante parece ser o grupo imidazol do resíduo de histidina. As proteínas nativas podem ser separadas com o uso de IMAC caso exibam porções químicas doadoras adequadas em sua superfície. Caso contrário, o IMAC pode ser usado para a separação de proteínas recombinantes que contêm uma cadeia de vários resíduos de histidina ligados.

[00998] Normalmente, o IMAC é realizado pela aplicação de uma mistura que contêm as espécies de interesse, à fase estacionária. Sob condições apropriadas, causará ligação coordenada das espécies à fase estacionária. Os componentes não ligados são, então, lavados antes da aplicação de um meio de eluição. Para a eluição, podem ser usados gradientes (contínuos ou como uma série de etapas) de aumento da concentração de sal ou diminuição do pH. Um procedimento também muito usado é a aplicação de um gradiente de concentração crescente de imidazol. Biomoléculas com diferentes propriedades de doador, por exemplo, que têm resíduos de histidina em ambientes diferentes, podem ser separadas pelo uso de eluição em gradiente.

[00999] Os vetores virais têm em sua superfície porções químicas, tais como proteínas, que podem ter capacidade para se ligar a fases estacionárias de IMAC. Isso significa que, potencialmente, o IMAC pode ser usado isoladamente.

[001000] As técnicas de centrifugação adequadas incluem centrifugação zonal, ultra isopicnica e centrifugação por peletização.

[001001] A esterilização em filtro é adequada para processos de materiais de grau farmacêutico. A esterilização em filtro torna a formulação resultante substancialmente livre de contaminantes. O nível de contaminantes após a esterilização em filtro é de modo que a formulação seja adequada para uso clínico. Concentração adicional (por exemplo, por ultrafiltração) após a etapa de esterilização por filtro pode ser realizada em condições assépticas. Em algumas modalidades, o filtro de esterilização tem um tamanho de poro máximo de 0,22 μm.

[001002] Os vetores retrovirais no presente documento também podem ser submetidos a métodos para concentrar e purificar um vetor lentiviral usando ultracentrifugação de fluxo direto e centrifugação de alta velocidade e filtração de fluxo tangencial. O fluxo através da ultracentrifugação pode ser usado para a purificação de vírus de tumor de RNA (Toplin et al, Applied Microbiology 15: 582 a 589, 1967; Burger et al., Journal of the National Cancer Institute 45: 499 a 503, 1970). Ultracentrifugação de fluxo direto pode ser usada para a purificação de vetores lentivirais. Esse método pode compreender uma ou mais das seguintes etapas. Por exemplo, um vetor lentiviral pode ser produzido a partir de células usando uma fábrica de células ou sistema de biorreator. Um sistema de transfecção transiente pode ser usado ou linhagens de células produtoras ou de empacotamento também podem ser usadas de forma similar. Uma etapa de pré-clarificação antes de carregar o material na ultracentrífuga pode ser usada se desejado. A ultracentrifugação de fluxo direto pode ser realizada usando fluxo contí-

nuo ou sedimentação em lote. Os materiais usados para a sedimentação são, por exemplo: cloreto de césio, tartarato de potássio e brometo de potássio, que criam altas densidades com baixa viscosidade, embora sejam todos corrosivos. O CsCl é frequentemente usado para o desenvolvimento do processo, pois um alto grau de pureza pode ser alcançado devido ao amplo gradiente de densidade que pode ser criado (1,0 a 1,9 g/cm^3). O brometo de potássio pode ser usado em altas densidades, por exemplo, em temperaturas elevadas, como 25 °C, o que pode ser incompatível com a estabilidade de algumas proteínas. A sacarose é amplamente utilizada por ser barata, não tóxica e pode formar um gradiente adequado para a separação da maioria das proteínas, frações subcelulares e células inteiras. Tipicamente, a densidade máxima é de aproximadamente 1,3 g/cm^3. O potencial osmótico da sacarose pode ser tóxico para as células, caso em que um material de gradiente complexo pode ser usado, por exemplo, Nycodenz. Um gradiente pode ser usado com 1 ou mais etapas no gradiente. Uma modalidade é usar um gradiente de sacarose em etapas. O volume do material pode ser de 0,5 litros a mais de 200 litros por execução. A velocidade da vazão pode ser de 5 a mais de 25 litros por hora. Uma velocidade operacional adequada está entre 25.000 e 40.500 rpm, produzindo uma força de até 122.000 × g. O rotor pode ser descarregado estaticamente nas frações de volume desejadas. Uma modalidade é descarregar o material centrifugado em frações de 100 ml. A fração isolada que contém o vetor Lentiviral purificado e concentrado pode então ser trocada em um tampão desejado usando filtração em gel ou cromatografia de exclusão por tamanho. A cromatografia de troca aniônica ou catiônica também pode ser usada como um método alternativo ou adicional para troca de tampão ou purificação adicional. Além disso, a Filtração de Fluxo Tangencial também pode ser usada para troca de tampão e formulação final, se necessário. A Filtração de fluxo tangen-

cial (TFF) também pode ser usada como uma etapa alternativa à centrifugação de ultra ou alta velocidade, em que um procedimento de TFF de duas etapas seria implementado. A primeira etapa reduz o volume do sobrenadante do vetor, ao passo que a segunda etapa é usada para a troca do tampão, formulação final e alguma concentração adicional do material. A membrana de TFF pode ter um tamanho de membrana entre 100 e 500 kilodaltons, em que a primeira etapa de TFF pode ter um tamanho de membrana de 500 quilodaltons, enquanto o segundo TFF pode ter um tamanho de membrana entre 300 a 500 kilodaltons. O tampão final deve conter materiais que permitem que o vetor seja armazenado para armazenamento de longo prazo.

[001003] Nas modalidades, o método usa fábricas de células que contêm células aderentes ou um biorreator que contém células em suspensão que são submetidas à transfecção ou submetidas à transdução com o vetor e construtos auxiliares para produzir o vetor lentiviral. Exemplos não limitantes ou biorreatores incluem o sistema de biorreator Wave e os biorreatores Xcellerex. Ambos são sistemas descartáveis. No entanto, sistemas não descartáveis também podem ser usados. Os construtos podem ser aqueles descritos no presente documento, assim como outros vetores de transdução lentiviral. Alternativamente, a linhagem celular pode ser projetada para produzir o vetor lentiviral sem a necessidade de transdução ou transfecção. Após a transfecção, o vetor lentiviral pode ser colhido e filtrado para remover as partículas e, em seguida, é centrifugado usando fluxo contínuo de alta velocidade ou ultracentrifugação. Uma modalidade preferencial é usar um dispositivo de fluxo contínuo de alta velocidade como o rotor de fluxo zonal e contínuo JCF-A com uma centrífuga de alta velocidade. É preferencial usar, também, a centrífuga Contifuge Stratus para produção de vetor lentiviral em média escala. Além disso, qualquer centrífuga de fluxo contínuo em que a velocidade de centrifugação se-

ja superior a 5.000 × g RCF e inferior a 26.000 × g RCF. De preferência, a força centrífuga de fluxo contínuo é aproximadamente 10.500 x g a 23.500 x g RCF com um tempo de rotação entre 20 horas e 4 horas, com tempos de centrifugação mais longos sendo usados com força centrífuga mais lenta. O vetor lentiviral pode ser centrifugado em uma almofada de material mais denso (um exemplo não limitante é a sacarose, porém outros reagentes podem ser usados para formar a almofada e estes são bem conhecidos na técnica) de modo que o vetor lentiviral não forme agregados que são não filtráveis, como às vezes ocorre com a centrifugação direta do vetor que resulta em um pélete de vetor viral. A centrifugação de fluxo contínuo em uma almofada permite que o vetor evite a formação de grande agregado, porém permite que o vetor seja concentrado em altos níveis de grandes volumes de material transfectado que produz o vetor lentiviral. Além disso, uma segunda camada menos densa de sacarose pode ser usada para unir a preparação do vetor lentiviral. A taxa de fluxo para a centrífuga de fluxo contínuo pode estar entre 1 e 100 ml por minuto, porém taxas de fluxo maiores e menores também podem ser usadas. A taxa de fluxo é ajustada para fornecer tempo suficiente para o vetor entrar no núcleo da centrífuga sem que quantidades significativas de vetor sejam perdidas devido à alta taxa de fluxo. Caso uma taxa de fluxo mais alta seja desejada, o material que sai da centrífuga de fluxo contínuo pode ser recirculado e pode passar pela centrífuga uma segunda vez. Depois que o vírus é concentrado usando centrifugação de fluxo contínuo, o vetor pode ser ainda mais concentrado usando Filtração de Fluxo Tangencial (TFF), ou o sistema TFF pode ser usado simplesmente para troca de tampão. Um exemplo não limitante de um sistema TFF é o sistema de cartucho Xampler produzido pela GB-Healthcare. Os cartuchos preferenciais são aqueles com um corte de MW de 500.000 MW ou menos. De preferência, um cartucho é usado com um corte de MW

de 300.000 MW. Um cartucho de 100.000 MW de corte também pode ser usado. Para volumes maiores, cartuchos maiores podem ser usados e será fácil para aqueles na técnica encontrar o sistema TFF certo para essa troca de tampão final e/ou etapa de concentração antes do preenchimento final da preparação do vetor. A preparação de enchimento final pode conter fatores que estabilizam o vetor - açúcares são geralmente usados e são conhecidos na técnica.

TEOR DE PROTEÍNA

[001004] Em algumas modalidades, a partícula retroviral inclui várias proteínas derivadas do genoma da célula de origem, proteínas exógenas e proteínas derivadas do genoma viral. Em algumas modalidades, a partícula retroviral contém várias razões entre proteínas derivadas do genoma da célula de origem e proteínas derivadas do genoma viral, entre proteínas derivadas do genoma da célula de origem e proteínas exógenas e entre proteínas exógenas e proteínas derivadas do genoma viral.

[001005] Em algumas modalidades, as proteínas derivadas do genoma viral são precursores de poliproteína GAG, HIV-1 Integrase, precursor de poliproteína POL, Capsídeo, Nucleocapsídeo, matriz p17, p6, p2, VPR, Vif.

[001006] Em algumas modalidades, as proteínas derivadas de células de origem são Ciclofilina A, Choque Térmico 70kD, Fator de Elongação Humana-1 Alfa (EF-1R), Histonas H1, H2A, H3, H4, beta-globina, Precursor de Tripsina, Parvulina, Gliceraldeído-3-fosfato desidrogenase, Lck, Ubiquitina, SUMO-1, CD48, Sintenina-1, Nucleofosmina, ribonucleoproteínas nucleares heterogêneas C1/C2, Nucleolina, provável helicase dependente de ATP DDX48, Matrina-3, ERTPase de transição, proteína nuclear de ligação a GTP Ran, ribonucleoproteína U nuclear heterogênea, fator de ligação de intensificador de interleucina 2, domínio não POU contendo proteína de ligação de octâmero,

RuvB like 2, HSP 90-b, HSP 90-a, fator de alongamento 2, D-3-fosfoglicerato desidrogenase, a-enolase, C-1-tetrahidrofolato sintase, citoplasmática, Piruvato quinase, isoenzimas M1/M2, Enzima de ativação de Ubiquitina E1, subunidade reguladora de protease 26S S10B, proteína ribossômica ácida P2 60S, proteína ribossômica ácida 60S P0, proteína ribossômica 40S S ribossomal SA, proteína ribossomal 40S S2, Proteína ribossomal S3 40S, ribossoma 60S l proteína L4, 60S proteína ribossômica L3, 40S proteína ribossômica S3a, 40S proteína ribossômica S7, 60S proteína ribossômica L7a, 60S proteína ribossômica ácida L31, 60S proteína ribossômica L10a, 60S proteína ribossômica L6, 26S proteassoma não-ATP subunidade regulatória de tubulina 1 cadeia b-2, Actina, citoplasmática 1, Actina, músculo liso aórtico, Tubulina a-cadeia ubíqua, Clatrina cadeia pesada 1, Histona H2B.b, Histona H4, Histona H3.1, Histona H3.3, Histona H2A tipo 8, Subunidade reguladora de protease 26S 6A, Ubiquitina-4, RuvB como 1, subunidade reguladora de protease 26S 7, Leucil-tRNA sintetase, citoplasmática, proteína ribossômica 60S L19, subunidade reguladora não-ATPase de proteassoma 26S subunidade 13 reguladora, Histona H2B.F, ribonucleoproteína nuclear pequena U5 200 kDa helicase, poli[ADP-ribose]polimerase-1, DNA helicase II dependente de ATP, fator de licenciamento de replicação de DNA MCM5, proteína de ligação de elemento sensível à nuclease 1, RNA helicase A dependente de ATP, fator de ligação de intensificador de interleucina 3, fator de alongamento da transcrição Polipeptídeo B 1, Pre-m Fator de processamento de RNA 8, domínio de nuclease estafilocócica contendo proteína 1, proteína de interação 6 de morte celular programada, Mediador de RNA polimerase II transcrição subunidade 8 homólogo, Nucleolar RNA helicase II, Endoplasmina, DnaJ homólogo subfamília A membro 1, Choque térmico 70 kDa proteína 1L, proteína do complexo T 1 e subunidade, proteína semelhante a GCN1 1, Serotransferrina, Frutose bifosfato al-

dolase A, Inosina-5'monofosfato desidrogenase 2, 26S subunidade reguladora de protease 6B, ácido graxo sintase, subunidade catalítica de proteína quinase dependente de DNA, 40S proteína ribossômica S17, proteína ribossômica L7 60S, proteína ribossômica L12 60S, proteína ribossômica L9 60S, proteína ribossômica S8 40S, proteína ribossômica S4 40S isoforma X, proteína ribossômica L11 60S, subunidade regulatória 2 de proteassoma não-ATPase 26S, subunidade reguladora da histona de Coatase H2A.z, Histona H1.2, Dineína citosólica da cadeia pesada. Consultar: Saphire et al., *Journal of Proteome Research,* 2005, e Wheeler et al., *Proteomics Clinical Applications*, 2007.

[001007] Em algumas modalidades, o vetor retroviral é peguilado.

TAMANHO DE PARTÍCULA

[001008] Em algumas modalidades, o diâmetro médio do vetor retroviral está entre 10 e 1.000 nM, 25 e 500 nm, 40 e 300 nm, 50 e 250 nm, 60 e 225 nm, 70 e 200 nm, 80 e 175 nm, ou 90 e 150 nm.

[001009] Em algumas modalidades, 90% dos vetores retrovirais estão dentro de 50% do diâmetro médio do retrovírus. Em algumas modalidades, 90% dos vetores retrovirais estão dentro de 25% do diâmetro médio do retrovírus. Em algumas modalidades, 90% dos vetores retrovirais estão dentro de 20% do diâmetro médio do retrovírus. Em algumas modalidades, 90% dos vetores retrovirais estão dentro de 15% do diâmetro médio do retrovírus. Em algumas modalidades, 90% dos vetores retrovirais estão dentro de 10% do diâmetro médio do retrovírus.

INDICAÇÕES E USOS

[001010] Os fusossomas, vetores retrovirais, VLPs ou composições farmacêuticas descritas no presente documento podem ser administradas a um indivíduo, por exemplo, um mamífero, por exemplo, um ser humano. Em tais modalidades, o indivíduo pode estar correndo o

risco, pode ter um sintoma ou pode ter sido diagnosticado ou identificado como tendo uma doença ou afecção específica (por exemplo, uma doença ou afecção descrita no presente documento). Em algumas modalidades, a doença é uma deficiência genética, por exemplo, uma deficiência genética listada na Tabela 5. Em algumas modalidades, o fusossoma, por exemplo, vetores retrovirais ou partículas, contém sequências de ácidos nucleicos que codificam um agente exógeno para o tratamento da doença ou afecção no indivíduo. Por exemplo, o agente exógeno é aquele que é específico ou pode ser usado para tratar uma deficiência genética, por exemplo, uma deficiência genética listada na Tabela 5, e o fusossoma é administrado a um indivíduo para o tratamento da deficiência genética no indivíduo.

[001011] Desse modo, também são fornecidos métodos de administração e usos, tais como usos profiláticos, dos fusossomas fornecidos, incluindo os vetores e partículas retrovirais fornecidos, tais como vetores lentivirais e partículas e/ou composições que compreendem os mesmos. Tais métodos e usos incluem métodos e usos terapêuticos que envolvem, por exemplo, a administração dos fusossomas, por exemplo, vetores ou partículas retrovirais, tais como vetores ou partículas lentivirais ou composições que contêm os mesmos a um indivíduo que tem uma doença, afecção ou distúrbio para entrega de um agente exógeno para tratamento da doença, afecção ou distúrbio. Em algumas modalidades, o fusossoma (por exemplo, vetor ou partícula retroviral, tal como um vetor ou partícula lentiviral) é administrado em uma quantidade ou dose eficaz para efetuar o tratamento da doença, afecção ou distúrbio. No presente documento são fornecidos usos de qualquer um dos fusossomas fornecidos (por exemplo, vetor ou partícula retroviral, tais como vetor ou partícula lentiviral) em tais métodos e tratamentos e na preparação de um medicamento a fim de realizar tais métodos terapêuticos. Em algumas modalidades, os métodos são rea-

lizados administrando-se os fusossomas (por exemplo, vetor ou partícula retroviral, como vetor ou partícula lentiviral), ou composições que compreendem os mesmos, ao indivíduo que tem, teve ou que suspeita ter a doença ou afecção ou distúrbio. Em algumas modalidades, os métodos tratam, então, a doença ou afecção ou distúrbio no indivíduo. Também são fornecidos no presente documento o uso de qualquer uma das composições, tais como composições farmacêuticas fornecidas no presente documento, para o tratamento de uma doença, afecção ou distúrbio associado a um determinado gene ou proteína alvejado por ou fornecido pelo agente exógeno.

[001012] Células-alvo de tecido de mamífero (por exemplo, ser humano) incluem células de tecido ou células epiteliais, conjuntivas, musculares ou nervosas e combinações dos mesmos. Células-alvo de mamífero (por exemplo, de seres humanos) e sistemas de órgãos incluem o sistema cardiovascular (coração, vasculatura); sistema digestivo (esôfago, estômago, fígado, vesícula biliar, pâncreas, intestinos, cólon, reto e ânus); sistema endócrino (hipotálamo, glândula pituitária, corpo pineal ou glândula pineal, tireoide, paratireoides, glândulas suprarrenais); sistema excretor (rins, ureteres, bexiga); sistema linfático (linfa, nódulos linfáticos, vasos linfáticos, amígdalas, adenoides, timo, baço); sistema tegumentar (pele, cabelo, unhas); sistema muscular (por exemplo, músculo esquelético); sistema nervoso (cérebro, medula espinhal, nervos)'; sistema reprodutivo (ovários, útero, glândulas mamárias, testículos, vasos deferentes, vesículas seminais, próstata); sistema respiratório (faringe, laringe, traqueia, brônquios, pulmões, diafragma); sistema esquelético (osso, cartilagem) e combinações dos mesmos. Em algumas modalidades, um sistema de células ou de órgãos não alvo é escolhido a partir do sistema cardiovascular (coração, vasculatura); sistema digestivo (esôfago, estômago, fígado, vesícula biliar, pâncreas, intestinos, cólon, reto e ânus); sistema endócrino (hi-

potálamo, glândula pituitária, corpo pineal ou glândula pineal, tireoide, paratireoides, glândulas suprarrenais); sistema excretor (rins, ureteres, bexiga); sistema linfático (linfa, nódulos linfáticos, vasos linfáticos, amígdalas, adenoides, timo, baço); sistema tegumentar (pele, cabelo, unhas); sistema muscular (por exemplo, músculo esquelético); sistema nervoso (cérebro, medula espinhal, nervos)'; sistema reprodutivo (ovários, útero, glândulas mamárias, testículos, vasos deferentes, vesículas seminais, próstata); sistema respiratório (faringe, laringe, traqueia, brônquios, pulmões, diafragma); sistema esquelético (osso, cartilagem) e combinações dos mesmos.

[001013] A administração de uma composição farmacêutica descritas no presente documento pode ocorrer por meio de uma administração oral, inalação, transdérmica ou parenteral (incluindo intravenosa, intratumoral, intraperitoneal, intramuscular, intracavidade e subcutânea). Os fusossomas podem ser administrados isoladamente ou formulados como uma composição farmacêutica.

[001014] Nas modalidades, a composição de fusossoma medeia um efeito em uma célula-alvo e o efeito dura pelo menos 1, 2, 3, 4, 5, 6 ou 7 dias, 2, 3 ou 4 semanas, ou 1, 2, 3, 6 ou 12 meses. Em algumas modalidades (por exemplo, em que a composição de fusossoma compreende uma proteína exógena), o efeito dura menos de 1, 2, 3, 4, 5, 6 ou 7 dias, 2, 3 ou 4 semanas, ou 1, 2, 3, 6 ou 12 meses.

[001015] Nas modalidades, a composição de fusossoma descrita no presente documento é entregue *ex vivo* a uma célula ou tecido, por exemplo, uma célula ou tecido humano.

[001016] As composições de fusossoma descritas no presente documento podem ser administradas a um indivíduo, por exemplo, um mamífero, por exemplo, um ser humano. Em tais modalidades, o indivíduo pode estar correndo o risco, pode ter um sintoma ou pode ter sido diagnosticado ou identificado como tendo uma doença ou afecção

específica (por exemplo, uma doença ou afecção descrita no presente documento).

[001017] Em algumas modalidades, a fonte de fusossomas é proveniente do mesmo indivíduo que recebe uma composição de fusossomas. Em outras modalidades, os mesmos são diferentes. Por exemplo, a fonte de fusossomas e tecido receptor pode ser autóloga (do mesmo indivíduo) ou heteróloga (de diferentes indivíduos). De todo modo, o tecido doador para as composições de fusossoma descritas no presente documento pode ser um tipo de tecido diferente do tecido receptor. Por exemplo, o tecido do doador pode ser tecido muscular e o tecido receptor pode ser tecido conjuntivo (por exemplo, tecido adiposo). Em outras modalidades, o tecido doador e o tecido receptor podem ser do mesmo tipo ou de tipo diferente, porém de sistemas de órgãos diferentes.

[001018] Em algumas modalidades, o fusossoma é coadministrado com um inibidor de uma proteína que inibe a fusão da membrana. Por exemplo, Suppressyn é uma proteína humana que inibe a fusão célula-célula (Sugimoto et al., “A novel human endogenous retroviral protein inhibits cell-cell fusion” Scientific Reports 3: 1.462 DOI: 10.1038/srep01462). Assim, em algumas modalidades, o fusossoma é coadministrado com um inibidor de sypressyn, por exemplo, um siRNA ou anticorpo inibidor.

[001019] As composições descritas no presente documento podem ser usadas, em algumas modalidades, para modular a função ou fisiologia celular ou tecidual de vários outros organismos, incluindo, porém sem limitação: animais de fazenda ou de trabalho (cavalos, vacas, porcos, galinhas, etc.), animais de estimação ou de zoológico (gatos, cães, lagartos, pássaros, leões, tigres e ursos, etc.), animais de aquicultura (peixes, caranguejos, camarões, ostras etc.), espécies de plantas (árvores, plantações, flores ornamentais etc.), espécies de fermen-

tação (saccharomyces, etc.). As composições de fusossoma descritas no presente documento podem ser produzidas a partir de tais origens não humanas e administradas a uma célula-alvo ou tecido não humano ou a um indivíduo.

[001020] As composições de fusossoma podem ser autólogas, alogênicas ou xenogênicas para o alvo.

AGENTES TERAPÊUTICOS ADICIONAIS

[001021] Em algumas modalidades, a composição de fusossoma é coadministrada com um agente adicional, por exemplo, um agente terapêutico, a um indivíduo, por exemplo, um receptor, por exemplo, um receptor descrito no presente documento. Em algumas modalidades, o agente terapêutico coadministrado é um agente imunossupressor, por exemplo, um glicocorticoide (por exemplo, dexametasona), citostático (por exemplo, metotrexato), anticorpo (por exemplo, Muromonab-CD3) ou modulador de imunofilina (por exemplo, Ciclosporina ou rapamicina). Nas modalidades, o agente imunossupressor diminui a depuração mediada por imunidade de fusossomas. Em algumas modalidades, a composição de fusossoma é coadministrada com um agente imunoestimulador, por exemplo, um adjuvante, uma interleucina, uma citocina ou uma quimiocina.

[001022] Em algumas modalidades, a composição de fusossoma e o agente imunossupressor são administrados ao mesmo tempo, por exemplo, administrados simultaneamente. Em algumas modalidades, a composição de fusossoma é administrada antes da administração do agente imunossupressor. Em algumas modalidades, a composição de fusossoma é administrada após a administração do agente imunossupressor.

[001023] Em algumas modalidades, o agente imunossupressor é uma molécula pequena, como ibuprofeno, acetaminofeno, ciclosporina, tacrolimus, rapamicina, micofenolato, ciclofosfamida, glicocorticoi-

des, sirolimus, azatriopina ou metotrexato.

[001024] Em algumas modalidades, o agente imunossupressor é uma molécula de anticorpo, incluindo, porém sem limitação: muronomabe (anti-CD3), Daclizumabe (anti-IL12), Basiliximabe, Infliximabe (Anti-TNFa) ou rituximabe (Anti-CD20).

[001025] Em algumas modalidades, a coadministração da composição de fusossoma com o agente imunossupressor resulta em persistência aumentada da composição de fusossoma no indivíduo em comparação com a administração da composição de fusossoma sozinha. Em algumas modalidades, a persistência aprimorada da composição de fusossoma na coadministração é de pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais, em comparação com persistência da composição fusossoma quando administrado sozinho. Em algumas modalidades, a persistência aumentada da composição de fusossoma na coadministração é de pelo menos 1, 2, 3, 4, 5, 6, 7, 10, 15, 20, 25 ou 30 dias ou mais, em comparação com a sobrevivência da composição de fusossoma quando administrado sozinho.

[001026] Em algumas modalidades, a composição de fusossoma é usada em combinação com uma terapia gênica, por exemplo, uma terapia gênica que entrega um transgene de interesse a um tecido de interesse. Em algumas modalidades, o tecido de interesse é o fígado. Em algumas modalidades, o tecido de interesse não é o fígado. Em algumas modalidades, a composição de fusossoma tem como alvo o mesmo transgene de interesse para expressão no fígado. Sem se ater nenhuma teoria, a expressão do transgene de interesse em hepatócitos pode favorecer a indução sistêmica de células T regulatórias que são específicas para o transgene, alcançando assim tolerância imunológica ao transgene.

EXEMPLOS

[001027] Os Exemplos a seguir são apresentados para auxiliar na

compreensão das invenções, porém não se destinam tampouco devem ser interpretados de modo algum como limitativos de seu escopo.

EXEMPLO 1. ENSAIO DE CÉLULAS FORA DO ALVO PARA DETECTAR A ESPECIFICIDADE DA ENTREGA DE ÁCIDO NUCLEICO RETROVIRAL

[001028] Esse Exemplo descreve a quantificação de um ácido nucleico em células receptoras fora do alvo medindo-se o número de cópias do vetor em células individuais.

[001029] Em uma modalidade, os camundongos tratados têm um número de cópias de vetor semelhante em células fora do alvo como aqueles de camundongos não tratados, por exemplo, nenhum vetor ou um número de vetor semelhante aos níveis de controle negativo. Em uma modalidade, os camundongos tratados têm uma porcentagem semelhante de células fora do alvo que contêm o vetor igual àquelas de camundongos não tratados, por exemplo, nenhuma célula ou um número de células semelhante aos níveis de controle negativo.

[001030] Nesse exemplo, a célula receptora fora do alvo é uma célula CD11c+. No entanto, esse protocolo pode ser adaptado a qualquer tipo de célula para o qual existem marcadores de superfície adequados e que podem ser isolados do indivíduo. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo.

[001031] Os camundongos são tratados com vetor retroviral produzido conforme descrito no presente documento ou com PBS (controle negativo). 28 dias após o tratamento, o sangue periférico é coletado de camundongos que receberam vetor retroviral e de camundongos que receberam tratamento com PBS. O sangue é coletado em 1 ml de PBS que contém EDTA a 5 μM e misturado imediatamente para evitar a coagulação. Os tubos são mantidos em gelo, e os glóbulos vermelhos são removidos com uma solução tamponada de cloreto de amônio

(ACK). As células são marcadas com um anticorpo CD11c:APC-Cy7 de murganho (n° de Catálogo da Biolegend: 117323) ou um anticorpo de controle de isótipo APC-Cy7 (n° de Catálogo da Biolegend: 400230) a 4 °C durante 30 minutos no escuro, após ser bloqueado por Fc (n° de Catálogo da Biolegend: 101319) em tampão de marcação de células (n° de Catálogo da Biolegend: 420201) durante 10 minutos. Após serem lavadas duas vezes com PBS, as células são analisadas em um FACS Aria (BD Biosciences, San José, CA.) com excitação a laser de 640 nm e emissão coletada a 780 -/+ 60 nm executando o software FACSDiva™ (BD Biosciences, San José, CA) para definir portas negativas com o uso das células identificadas com anticorpo APC-Cy7 de controle de isótipo. As células positivas APC-Cy7 são classificadas em poços únicos da placa para análise do número de cópias do vetor.

[001032] O número de cópias do vetor é avaliado com o uso de PCR aninhada de célula única. A PCR é realizada com qPCR usando iniciadores e sondas específicas para o vetor e um gene de controle endógeno. O número de cópias do vetor é determinado dividindo-se a quantidade de sinal qPCR do vetor pela quantidade do sinal qPCR do gene de controle endógeno. Uma célula que recebeu o vetor terá um número de cópias do vetor de pelo menos 1,0. O número de cópias do vetor é avaliado em toda a população calculando-se a média do número de cópias do vetor da pluralidade de células.

[001033] Em algumas modalidades, os camundongos tratados com vetores retrovirais têm um número médio de cópias de vetor em células fora do alvo semelhante ao dos camundongos tratados com o veículo. Em algumas modalidades, os camundongos tratados com vetores retrovirais têm uma porcentagem semelhante de células fora do alvo que receberam o vetor como aquelas de camundongos tratados com veículo.

EXEMPLO 2. ENSAIO DE CÉLULAS FORA DO ALVO PARA DE-

TECTAR A ESPECIFICIDADE DA ENTREGA DE UM AGENTE DE PROTEÍNA EXÓGENA

[001034] Esse Exemplo descreve a quantificação da expressão de um agente exógeno em células receptoras fora do alvo pela expressão de agente exógeno em células individuais.

[001035] Em uma modalidade, os camundongos tratados têm expressão de agente exógeno semelhante em células fora do alvo igualmente àquelas de camundongos não tratados. Em uma modalidade, os camundongos tratados têm uma porcentagem semelhante de células fora do alvo que expressam o agente exógeno igualmente àquelas de camundongos não tratados.

[001036] Nesse exemplo, a célula receptora fora do alvo é uma célula CD11c+. No entanto, esse protocolo pode ser adaptado a qualquer tipo de célula para o qual existem marcadores de superfície adequados e que podem ser isolados do indivíduo. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo. Nesse exemplo, o agente exógeno é uma proteína fluorescente, e a expressão é medida por meio de citometria de fluxo. Em outras modalidades, a expressão de um agente de proteína exógena pode ser medida com imunomarcação para a proteína. Em outras modalidades, a expressão do agente de proteína exógena pode ser medida por meio de microscopia ou Western blot.

[001037] Os camundongos são tratados com vetor retroviral com um agente de proteína tdtomato fluorescente produzido por meio de qualquer um dos métodos descritos no presente pedido ou com PBS (controle negativo). 28 dias após o tratamento, o sangue periférico é coletado de camundongos que receberam vetor retroviral e de camundongos que receberam tratamento com PBS. O sangue é coletado em 1 ml de PBS que contém EDTA a 5 μM e misturado imediatamente para

evitar a coagulação. Os tubos são mantidos em gelo, e os glóbulos vermelhos são removidos com uma solução tamponada de cloreto de amônio (ACK). As células são marcadas com um anticorpo CD11c:APC-Cy7 de murganho (n° de Catálogo da Biolegend: 117323), ou o isótipo controla o anticorpo APC-Cy7 (n° de Catálogo da Biolegend: 400230) a 4 °C durante 30 minutos no escuro, após ser bloqueado Fc (n° de Catálogo da Biolegend: 101319) em tampão de marcação de células (n° de Catálogo da Biolegend: 420201) durante 10 minutos. Após serem lavadas duas vezes com PBS, as células são analisadas em um FACS Aria (BD Biosciences, San José, CA.) executando o software FACSDiva™ (BD Biosciences, San José, CA). Uma porta negativa para CD11c é definida com o uso das células identificadas com anticorpo APC-Cy7 de controle de isótipo e com uma excitação a laser de 640 nm e emissão coletada em 780 -/+ 60. Uma porta negativa para a expressão de tdtomato é definida com células isoladas de camundongos tratados com veículo e com uma excitação de laser de 552 nm e uma emissão coletada em 585 -/+ 42 nm.

[001038] A porcentagem de células CD11c+ que são tdtomato-positivas é medida. Em algumas modalidades, a porcentagem de células CD11c+ que são tdtomato-positivas é semelhante em células de camundongos tratados e não tratados. O nível mediano de fluorescência de tdtomato é medido em células CD11c+. Em algumas modalidades, o nível de fluorescência de tdtomato mediano em células CD11c+ é semelhante em células de camundongos tratados e não tratados.

EXEMPLO 3. ENSAIO DE CÉLULAS-ALVO PARA DETECTAR A ESPECIFICIDADE DA ENTREGA DE ÁCIDO NUCLEICO RETROVIRAL

[001039] Esse Exemplo descreve a quantificação de um ácido nucleico em células-alvo receptoras medindo-se o número de cópias do vetor em células individuais.

[001040] Em uma modalidade, os camundongos tratados têm um número de cópias do vetor em células-alvo maior que aqueles de camundongos não tratados. Em uma modalidade, os camundongos tratados têm uma porcentagem de células-alvo maior que aquelas de camundongos não tratados.

[001041] Nesse exemplo, a célula receptora de destino é uma célula CD3+. No entanto, esse protocolo pode ser adaptado a qualquer tipo de célula para o qual existem marcadores de superfície adequados e que podem ser isolados do indivíduo. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo.

[001042] Os camundongos são tratados com vetor retroviral, e uma amostra de sangue é coletada conforme descrito acima no Exemplo 1. As células são marcadas com um anticorpo CD3:APC-Cy7 de murganho (n° de Catálogo da Biolegend: 100330) ou um controle de isótipo usando do protocolo descrito acima no Exemplo 1. O número de cópias do vetor é avaliado com o uso de PCR aninhada de célula única, conforme descrito no Exemplo 1.

[001043] Em algumas modalidades, camundongos tratados com vetores retrovirais têm um número médio de cópias de vetor em células-alvo maior que aqueles de camundongos tratados com veículo. Em algumas modalidades, camundongos tratados com vetores retrovirais têm uma porcentagem de células-alvo que receberam o vetor maior que aquelas de camundongos tratados com veículo.

EXEMPLO 4. ENSAIO DE CÉLULAS-ALVO PARA DETECTAR A ESPECIFICIDADE DA ENTREGA DE UM AGENTE DE PROTEÍNA EXÓGENA

[001044] Esse Exemplo descreve a quantificação da expressão de um agente de proteína exógena em células-alvo receptoras por expressão do agente de proteína exógena em células individuais.

[001045] Em uma modalidade, os camundongos tratados têm expressão de agente de proteína exógena em células-alvo maior que aquelas de camundongos não tratados. Em uma modalidade, os camundongos tratados têm uma porcentagem de células-alvo que expressam o agente de proteína exógena maior que aquela de camundongos não tratados.

[001046] Nesse exemplo, a célula receptora de destino é uma célula CD3+. No entanto, esse protocolo pode ser adaptado a qualquer tipo de célula para o qual existem marcadores de superfície adequados e que podem ser isolados do indivíduo. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo. Nesse exemplo, o agente de proteína exógena é uma proteína fluorescente, e a expressão é medida por meio de citometria de fluxo. Em outras modalidades, a expressão de um agente de proteína exógena pode ser medida com imunomarcação para a proteína. Em outras modalidades, a expressão do agente de proteína exógena pode ser medida por meio de microscopia ou Western blot.

[001047] Os camundongos são tratados com vetor retroviral, e uma amostra de sangue é coletada conforme descrito acima no Exemplo 2. As células são marcadas com um anticorpo CD3:APC-Cy7 de murganho (Catálogo Biolegend nº: 100330) ou controles de isótipo e analisados por citometria de fluxo usando o protocolo descrito no Exemplo 2.

[001048] A porcentagem de células CD3+ que são tdtomato-positivas é medida. Em algumas modalidades, a porcentagem de células CD3+ que são tdtomato-positivas será maior nas células em camundongos tratados do que em camundongos não tratados. O nível mediano de fluorescência de tdtomato é medido em células CD3+. Em algumas modalidades, o nível de fluorescência de tdtomato mediano em células CD3+ é maior em células de camundongos tratados do que não trata-

dos.

EXEMPLO 5. MODIFICAÇÃO DO VETOR RETROVIRAL COM HLA-G OU HLA-E PARA DIMINUIÇÃO DA CITOTOXICIDADE MEDIADA PELA LISE DE CÉLULAS PBMC

[001049] Esse Exemplo descreve vetores retrovirais derivados de células modificadas para ter citotoxicidade diminuída devido à lise celular por células mononucleares de sangue periférico (PBMCs).

[001050] Em uma modalidade, a lise celular de vetores retrovirais mediada por citotoxicidade por PBMCs é uma medida de imunogenicidade de vetores retrovirais, uma vez que a lise reduzirá, por exemplo, inibir ou interromper, a atividade de um vetor retroviral.

[001051] Os vetores retrovirais são criados a partir de: células não modificadas (doravante NMCs, controle positivo), células que são submetidas à transfecção com cDNA HLA-G ou HLA-E (doravante NMC-HLA-G) e células submetidas à transfecção com um controle de vetor vazio (doravante vetor sem NMC, controle negativo).

[001052] A lise mediada por PBMC de um vetor retroviral é determinada por ensaios de liberação de európio, conforme descrito em Bouma, et al. Murmurar. Immunol. 35(2):85 a 92; 1992 & van Besouw et al. Transplantation 70(1):136 a 143; 2000. PBMCs (doravante células efetoras) são isoladas de um doador apropriado e estimuladas com PMBCs irradiadas com gama alogênica e IL-2 a 200 UI/ml (proleucina, Chiron BV Amsterdam, Holanda) em uma placa com 96 poços de fundo redondo durante 7 dias a 37 °C. Os vetores retrovirais são identificados com európio-dietilenotriaminopenta-acetato (DTPA) (sigma, St. Louis, MO, EUA).

[001053] No dia 7, os ensaios de lise mediada por citotoxicidade são realizados incubando-se o vetor retroviral identificado por ^{63}Eu com células em uma placa com 96 poços para 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 24 ou 48 horas após a colocação em placa a razões entre efeto-

ras/alvo em uma faixa de 1.000:1 a 1:1 e 1:1,25 a 1:1.000. Após a incubação, as placas são centrifugadas e uma amostra do sobrenadante é transferida a placas com 96 poços com baixa fluorescência em segundo plano (fluoroimunoplacas, Nunc, Roskilde, Dinamarca).

[001054] Posteriormente, a solução de intensificação (PerkinElmer, Groningen, Holanda) é adicionada a cada poço. O európio liberado é medido em um fluorômetro resolvido no tempo (contator de múltiplas identificações Victor 1420, LKB-Wallac, Finlândia). A fluorescência é expressa em contagens por segundo (CPS). A liberação percentual máxima de európio por um vetor retroviral alvo é determinada incubando-se um número apropriado (1×10^2 a 1×10^8) de vetores retrovirais com triton a 1% (sigma-aldrich) durante um período de tempo apropriado. A liberação espontânea de európio pelo vetor retroviral alvo é medida por incubação do vetor retroviral alvo identificado sem células efetoras. Em seguida, a porcentagem de vazamento é calculada como: (liberação espontânea/liberação máxima) x 100%. A percentagem de lise mediada por citotoxicidade é calculada como % de lise = [(lise medida-lise espontânea-liberação espontânea)/(liberação máxima-liberação espontânea)] x 100%. Os dados são analisados observando-se a porcentagem de lise como uma função de diferentes razões de alvos efetores.

[001055] Em uma modalidade, os vetores retrovirais gerados a partir de células NMC-HLA-G terão uma porcentagem diminuída de lise por células-alvo em instantes do tempo específicos, em comparação a vetores retrovirais gerados a partir de NMCs ou vetor sem NMC.

EXEMPLO 6. MODIFICAÇÃO DO VETOR RETROVIRAL COM HLA-G OU HLA-E PARA DIMINUIÇÃO DA ATIVIDADE DE LISE DE NK

[001056] Esse Exemplo descreve a geração de uma composição de vetor retroviral derivada de uma fonte de células que foi modificada para diminuir a lise celular mediada por citotoxicidade por células NK.

Em uma modalidade, a lise celular de vetores retrovirais por células NK mediada por citotoxicidade é uma medida de imunogenicidade para vetores retrovirais.

[001057] Os vetores retrovirais são criados a partir de: células não modificadas (doravante NMCs, controle positivo), células que são submetidas à transfecção com cDNA HLA-G ou HLA-E (doravante NMC-HLA-G) e células submetidas à transfecção com um controle de vetor vazio (doravante vetor sem NMC, controle negativo).

[001058] A lise mediada por células NK de um vetor retroviral é determinada por ensaios de liberação de európio, conforme descrito em Bouma, et al. Murmurar. Immunol. 35(2):85 a 92; 1992 & van Besouw et al. Transplantation 70(1):136 a 143; 2000. As células NK (doravante células efetoras) são isoladas de um doador apropriado de acordo com os métodos em Crop et al. Cell transplantation (20):1.547 a 1.559; 2011 e estimuladas com PMBCs alogênicas irradiadas com gama e 200IU/ml de IL-2 (proleucina, Chiron BV Amsterdã, Países baixos) em uma placa com 96 poços de fundo redondo durante 7 dias a 37 °C. Os vetores retrovirais são identificados com európio-dietilenotriaminopenta-acetato (DTPA) (sigma, St. Louis, MO, EUA). Os ensaios de lise mediada por citotoxicidade e a análise de dados são realizados, conforme descrito acima no Exemplo 5.

[001059] Em uma modalidade, os vetores retrovirais gerados a partir de células NMC-HLA-G terão uma porcentagem diminuída de lise por células-alvo em instantes do tempo específicos, em comparação a vetores retrovirais gerados a partir de NMCs ou vetor sem NMC.

EXEMPLO 7. MODIFICAÇÃO DO VETOR RETROVIRAL COM HLA-G OU HLA-E PARA DIMINUIÇÃO DA LISE DE CÉLULAS T EXTERMINADORAS CD8

[001060] Esse Exemplo descreve a geração de uma composição de vetor retroviral derivada de uma fonte de células que foi modificada

para diminuir a lise celular mediada por citotoxicidade por células CD8+ T. Em uma modalidade, a lise celular mediada por citotoxicidade do vetor retroviral por células T CD8+ é uma medida de imunogenicidade para vetores retrovirais.

[001061] Os vetores retrovirais são criados a partir de: células não modificadas (doravante NMCs, controle positivo), células que são submetidas à transfecção com cDNA HLA-G ou HLA-E (doravante NMC-HLA-G) e células submetidas à transfecção com um controle de vetor vazio (doravante vetor sem NMC, controle negativo).

[001062] A lise mediada por células T CD8+ de um vetor retroviral é determinada por ensaios de liberação de európio, conforme descrito em Bouma, et al. Murmurar. Immunol. 35(2):85 a 92; 1992 & van Besouw et al. Transplantation 70(1):136 a 143; 2000. As células T CD8+ (doravante células efetoras) são isoladas de um doador apropriado de acordo com os métodos em Crop et al. Cell transplantation (20):1.547 a 1.559; 2011 e estimuladas com PMBCs alogênicas irradiadas com gama e 200IU/ml de IL-2 (proleucina, Chiron BV Amsterdã, Países baixos) em uma placa com 96 poços de fundo redondo durante 7 dias a 37 °C. Os vetores retrovirais são identificados com európio-dietilenotriaminopenta-acetato (DTPA) (sigma, St. Louis, MO, EUA). Os ensaios de lise mediada por citotoxicidade e a análise de dados são realizados, conforme descrito acima no Exemplo 5.

[001063] Em uma modalidade, os vetores retrovirais gerados a partir de células NMC-HLA-G terão uma porcentagem diminuída de lise por células-alvo em instantes do tempo específicos, em comparação a vetores retrovirais gerados a partir de NMCs ou vetor sem NMC.

EXEMPLO 8: MODIFICAÇÃO DO VETOR RETROVIRAL COM CD47 PARA EVITAR A FAGOCITOSE DE MACRÓFAGOS

[001064] Esse Exemplo descreve a quantificação da evasão da fagocitose pelo vetor retroviral modificado. Em uma modalidade, o vetor

retroviral modificado evitará a fagocitose por macrófagos.

[001065] As células iniciam a fagocitose, engolfando partículas, o que permite o sequestro e destruição de invasores estranhos, tais como bactérias ou células mortas. Em algumas modalidades, a fagocitose de vetores lentivirais por macrófagos reduziria sua atividade. Em algumas modalidades, a fagocitose de vetores lentivirais é uma medida de imunogenicidade de vetores retrovirais.

[001066] Os vetores retrovirais são produzidos a partir de células que carecem de CD47 (doravante NMC, controle positivo), células que são submetidas à transfecção com cDNA de CD47 (a seguir NMC-CD47) e células submetidas à transfecção com um controle de vetor vazio (doravante vetor sem NMC, controle negativo). Antes da produção do vetor retroviral, as células são identificadas com CSFE.

[001067] A redução da depuração imune mediada por macrófagos é determinada com um ensaio de fagocitose de acordo com o protocolo a seguir. Os macrófagos são semeados imediatamente após a colheita em pratos confocais com fundo de vidro. Os macrófagos são incubados em DMEM+10%FBS+1%P/S durante 1 horas para fixação. Um número apropriado de vetores retrovirais produzidos a partir de NMC, NMC-CD47, vetor sem NMC são adicionados aos macrófagos, conforme indicado no protocolo e são incubados durante 2 horas, tools.thermofisher.com/content/sfs/manuals/mp06694.pdf.

[001068] Após 2 horas, a prato é cuidadosamente lavada, e a fluorescência intracelular é examinada. A fluorescência intracelular emitida por partículas retrovirais engolfadas é visualizada por microscopia confocal a 488 de excitação. O número de macrófagos fagocitóticos positivos é quantificado usando software de imagem. Os dados são expressos como o índice fagocítico = (número total de células engolfadas/número total de macrófagos contados) × (número de macrófagos contendo células engolfadas/número total de macrófagos contados) ×

100.

[001069] Em uma modalidade, o índice fagocítico será reduzido quando os macrófagos forem incubados com vetores retrovirais derivados de NMC-CD47, versus aqueles derivados de NMC ou vetor sem NMC.

EXEMPLO 9: MODIFICAÇÃO DO VETOR RETROVIRAL COM PROTEÍNAS REGULATÓRIAS DO COMPLEMENTO PARA EVITAR O COMPLEMENTO

[001070] Este Exemplo descreve a quantificação da atividade do complemento contra um vetor retroviral com o uso de um ensaio *in vitro*. Em algumas modalidades, um vetor retroviral modificado descrito no presente documento terá atividade de complemento reduzida em comparação a um vetor retroviral não modificado.

[001071] Nesse Exemplo, o soro de um camundongo é avaliado quanto à atividade do complemento contra um vetor retroviral. O exemplo mede o nível de complemento C3a, que é um nó central em todas as vias do complemento. Os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo.

[001072] Nesse exemplo, os vetores retrovirais são gerados a partir de células HEK293 submetidas à transfecção com um cDNA que codifica para a proteína regulatória do complemento DAF (vetor retroviral HEK293-DAF) ou células HEK 293 que não expressam uma proteína regulatória do complemento (vetor retroviral HEK293). Em outras modalidades, podem ser usadas outras proteínas regulatórias do complemento, tais como proteínas que se ligam ao fator acelerador de decaimento (DAF, CD55), por exemplo, proteína-1 semelhante ao fator H (FH) (FHL-1), por exemplo, proteína de ligação a C4b (C4BP), por exemplo, receptor de complemento 1 (CD35), por exemplo, proteína cofator de membrana (MCP, CD46), por exemplo, Profectina (CD59),

por exemplo, proteínas que inibem as enzimas CD/C5 convertase da via clássica e alternativa do complemento, por exemplo, proteínas que regulam a montagem do MAC.

[001073] O soro é recuperado de camundongos naïve, camundongos que recebem o vetor retroviral HEK293-DAF ou camundongos que recebem o vetor retroviral HEK293. Os soros são coletados de camundongos coletando-se sangue total fresco e permitindo-se que coagule completamente por várias horas. Os coágulos são peletizados por centrifugação, e os sobrenadantes do soro são removidos. Um controle negativo é soro de camundongo inativado por calor. As amostras de controle negativo são aquecidas a 56 graus Celsius durante 1 hora. O soro pode ser congelado em alíquotas.

[001074] Os diferentes vetores retrovirais são testados para a dose em que 50% das células em uma população de células-alvo recebem o agente exógeno no vetor retroviral. O vetor retroviral pode conter qualquer um dos agentes exógenos descritos no presente documento. Muitos métodos para testar a entrega retroviral de um agente exógeno às células receptoras também são descritos no presente documento. Nesse exemplo particular, o agente exógeno é a proteína Cre (codificada pelo ácido nucleico retroviral) e as células-alvo são células RPMI8226 que expressam de maneira estável um cassete “LoxP-GFP-stop-LoxP-RFP” sob um promotor CMV que, mediante a recombinação por Cre, muda de GFP para a expressão RFP, como um marcador de entrega. A dose identificada na qual 50% das células receptoras são RFP-positivas é usada para outros experimentos. Em algumas modalidades, a dose identificada na qual 50% das células receptoras recebem o agente exógeno será semelhante entre os vetores retrovirais.

[001075] As diluições duplas em solução salina tamponada com fosfato (PBS, pH 7,4) dos vetores retrovirais, começando na dose de vetores retrovirais em que 50% das células-alvo recebem o agente exó-

geno, são misturadas com uma diluição de 1:10 dos soros de camundongos tratados com os mesmos vetores retrovirais ou camundongos naïve (volume de ensaio, 20 µl) e incubados por 1 hora a 37 °C. As amostras são posteriormente diluídas 1:500 e utilizadas em um ensaio de ensaio de imunoadsorção ligado à enzima (ELISA) específico para C3a. O ELISA é um produto do kit LS-F4210 de complemento de camundongo ELISA para C3a vendido pela LifeSpan BioSciences Inc, que mede a concentração de C3a em uma amostra. A dose de vetor retroviral na qual 200 pg/ml de C3a está presente é comparada entre soros isolados de camundongos.

[001076] Em algumas modalidades, a dose de vetor retroviral em que 200 pg/ml de C3a está presente para o vetor retroviral HEK293-DAF incubado com soro de camundongo HEK-293 DAF será maior do que para o vetor retroviral HEK293 incubado com soro de camundongo HEK293, indicando que a atividade de complemento direcionadas ao vetor retroviral é maior em camundongos tratados com vetor retroviral HEK293 do que com o vetor retroviral HEK293-DAF. Em algumas modalidades, a dose de vetor retroviral em que 200 pg/ml de C3a está presente para o vetor retroviral HEK293-DAF incubado com soro de camundongo naïve será maior que para o vetor retroviral HEK293 incubado com soro de camundongo naïve, indicando que a atividade de complemento direcionadas ao vetor retroviral em camundongos tratados com vetor retroviral HEK293 é maior do que com o vetor retroviral HEK293-DAF.

EXEMPLO 10: MODIFICAÇÃO DO VETOR RETROVIRAL PARA REALIZAR KNOCKDOWN DA PROTEÍNA IMUNOGÊNICA PARA REDUZIR A IMUNOGENICIDADE

[001077] Esse Exemplo descreve a geração de uma composição de vetor retroviral derivada de uma fonte de células que foi modificada para reduzir a expressão de uma molécula que é imunogênica e a

quantificação da expressão reduzida. Em uma modalidade, um vetor retroviral pode ser derivado de uma fonte de células, que foi modificada para reduzir a expressão de uma molécula que é imunogênica.

[001078] As terapias que estimulam uma resposta imunológica podem reduzir a eficácia terapêutica ou causar toxicidade ao receptor. Assim, a imunogenicidade é uma propriedade importante para vetores retrovirais terapêuticos seguros e eficazes. A expressão de determinados agentes de ativação imunológica pode criar uma resposta imunológica. A classe I de MHC representa um exemplo de um agente de ativação imunológica.

[001079] Os vetores retrovirais são produzidos a partir de células não modificadas que normalmente expressam MHC-1 (doravante, NMC, controle positivo), células que são submetidas à transfecção com um DNA que codifica para um shRNA que alveja a classe I de MHC (doravante, classe I de NMC-shMHC) e células submetidas à transfecção com um DNA que codifica para controle de vetor shRNA embaralhado não alvejado (doravante controle de vetor de NMC, controle negativo). Antes da produção retroviral, as células são identificadas com CSFE.

[001080] Os vetores retrovirais são testados quanto à expressão de classe I de MHC com o uso de citometria de fluxo. Uma quantidade apropriada de vetores retrovirais é lavada e suspensa novamente em PBS, mantida em gelo durante 30 minutos com diluição 1:10 a 1:4.000 de anticorpos monoclonais conjugados por fluorescência contra classe I de MHC (Harlan Sera-Lab, Belton, UK). Os vetores retrovirais são lavados três vezes em PBS e suspensos novamente em PBS. A fluorescência não específica é determinada com o uso de alíquotas iguais de preparação de vetor retroviral incubadas com anticorpo de controle de isótipo, e conjugadas de maneira fluorescentemente apropriada com o mesmo, em diluições equivalentes. Os vetores retrovirais são testados um citômetro de fluxo (FACSort, Becton-Dickinson), e os da-

dos são analisados com software de análise de fluxo (Becton-Dickinson).

[001081] Os dados médios de fluorescência dos vetores retrovirais derivados de NMCs, classe I de NMC-shMHC e controle do vetor de NMC são comparados. Em uma modalidade, os vetores retrovirais derivados de classe I de NMC-shMHC terão menor expressão de MHC de classe I em comparação a NMCs e o controle de vetor de NMC.

EXEMPLO 11: MEDIÇÃO DA INATIVAÇÃO SÉRICA PRÉ-EXISTENTE DE VETORES RETROVIRAIS

[001082] Esse Exemplo descreve a quantificação da inativação do soro pré-existente de vetores retrovirais com o uso de um ensaio de entrega *in vitro*.

[001083] Em algumas modalidades, uma medida de imunogenicidade para vetores retrovirais é a inativação sérica. A inativação sérica de vetores retrovirais pode ser devido à neutralização mediada por anticorpos ou degradação mediada por complemento. Em uma modalidade, alguns receptores de vetores retrovirais descritos no presente documento terão fatores em seu soro que se ligam e inativam os vetores retrovirais.

[001084] Nesse exemplo, um camundongo naïve de vetor retroviral é avaliado quanto à presença de fatores que inativam vetores retrovirais no soro. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo.

[001085] O controle negativo é o soro de camundongo inativado por calor e o controle positivo é soro derivado de um camundongo que recebeu múltiplas injeções de vetor retroviral gerado a partir de uma célula de origem xenogênica. Os soros são coletados de camundongos coletando-se sangue total fresco e permitindo-se que coagule completamente por várias horas. Os coágulos são peletizados por centrifuga-

ção, e os sobrenadantes do soro são removidos. As amostras de controle negativo são aquecidas a 56 graus Celsius durante 1 hora. O soro pode ser congelado em alíquotas.

[001086] Os vetores retrovirais são testados para a dose na qual 50% das células em uma população de células-alvo recebem o agente exógeno no vetor retroviral, conforme descrito acima no Exemplo 9.

[001087] A fim de avaliar a inativação sérica de vetores retrovirais, os vetores retrovirais são diluídos 1:5 em soro normal ou inativado por calor (ou meio que contém FBS inativado por calor a 10% como o controle sem soro), e a mistura é incubada a 37 °C durante 1 hora. Após a incubação, o meio é adicionado à reação para uma diluição adicional de 1:5 e, em seguida, diluído em série duas vezes na razão de 1:10. Após essa etapa, os vetores retrovirais devem estar presentes na dose previamente identificada em que 50% das células receptoras receberam o agente exógeno (por exemplo, são positivas para RFP).

[001088] Os vetores retrovirais que foram expostos ao soro são, em seguida, incubados com células-alvo. É calculada a porcentagem de células que recebem o agente exógeno e, portanto, são RFP-positivas. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno não será diferente entre as amostras de vetor retroviral que foram incubadas com soro e soro inativado por calor de camundongos naïve de vetor retroviral, o que indica que não há inativação do soro do vetor retroviral. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno não será diferente entre as amostras de vetor retroviral que foram incubadas com soro de camundongos naïve de vetor retroviral e incubações de controle sem soro, o que indica que não há inativação do soro de vetores retrovirais. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno em amostras de vetor retroviral que foram incubadas com soro de controle positivo será menor do que em amostras de ve-

tor retroviral que foram incubadas com soro de camundongos naïve de vetor retroviral, o que indica que não há inativação sérica de vetores retrovirais.

EXEMPLO 12: MEDIÇÃO DA INATIVAÇÃO SÉRICA DE VETORES RETROVIRAIS APÓS MÚLTIPLAS ADMINISTRAÇÕES

[001089] Esse Exemplo descreve a quantificação da inativação do soro de vetores retrovirais com o uso de um ensaio de entrega *in vitro* após múltiplas administrações dos vetores retrovirais. Em uma modalidade, um vetor retroviral modificado, por exemplo, modificado por um método descrito no presente documento, terá uma inativação sérica reduzida (por exemplo, reduzida em comparação à administração de um vetor retroviral não modificado) após múltiplas (por exemplo, mais de uma, por exemplo, 2 ou mais) administrações do vetor retroviral modificado. Em uma modalidade, um vetor retroviral descrito no presente documento não será inativado pelo soro após múltiplas administrações.

[001090] Em algumas modalidades, uma medida de imunogenicidade para o vetor retroviral é a inativação do soro. Em uma modalidade, as injeções repetidas de um vetor retroviral podem levar ao desenvolvimento de anticorpos de vetor antirretroviral, por exemplo, anticorpos que reconhecem vetores retrovirais. Em uma modalidade, os anticorpos que reconhecem vetores retrovirais podem se ligar de maneira que possam limitar a atividade do vetor retroviral ou a longevidade e mediar a degradação do complemento.

[001091] Nesse Exemplo, a inativação do soro é examinada após uma ou mais administrações de vetores retrovirais. Os vetores retrovirais são produzidos por qualquer um dos Exemplos anteriores. Nesse exemplo, retrovirais são criados a partir de: células que são submetidas à transfecção com cDNA de HLA-G ou HLA-E (doravante NMC-HLA-G) e células submetidas à transfecção com um controle de vetor

vazio (doravante vetor sem NMC, controle negativo). Em algumas modalidades, os vetores retrovirais são derivados de células que expressam outras proteínas imunorregulatórias.

[001092] O soro é retirado de coortes diferentes: camundongos injetados sistemicamente e/ou localmente com 1, 2, 3, 5 ou 10 injeções de veículo (grupo virgem de vetor retroviral), vetor retroviral HEK293-HLA-G ou vetor retroviral HEK293. Os soros são coletados de camundongos coletando-se sangue total fresco e permitindo-se que coagule completamente por várias horas. Os coágulos são peletizados por centrifugação, e os sobrenadantes do soro são removidos. Um controle negativo é soro de camundongo inativado por calor. As amostras de controle negativo são aquecidas a 56 graus Celsius durante 1 hora. O soro pode ser congelado em alíquotas.

[001093] Os vetores retrovirais são testados para a dose na qual 50% das células em uma população de células-alvo recebem o agente exógeno no vetor retroviral, conforme descrito acima no Exemplo 9.

[001094] A fim de avaliar a inativação do soro de vetores retrovirais, os vetores retrovirais são expostos ao soro e incubados com células-alvo, conforme descrito no Exemplo 11 acima.

[001095] É calculada a porcentagem de células que recebem o agente exógeno e, portanto, são RFP-positivas. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno não será diferente entre as amostras de vetor retroviral que foram incubadas com soro e soro inativado por calor de camundongos tratados com vetores retrovirais HEK293-HLA-G, o que indica que não há inativação sérica de vetores retrovirais ou uma resposta imunológica adaptativa. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno não será diferente entre as amostras de vetor retroviral que foram incubadas de camundongos tratados 1, 2, 3, 5 ou 10 vezes com vetores retrovirais HEK293-HLA-G, indicando que não é a

inativação sérica de vetores retrovirais ou uma resposta imunológica adaptativa. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno não será diferente entre as amostras de vetor retroviral que foram incubadas com soro de camundongos tratados com veículo e de camundongos tratados com vetores retrovirais HEK293-HLA-G, indicando que não há inativação sérica de vetores retrovirais ou uma resposta imunológica adaptativa. Em algumas modalidades, a porcentagem de células que recebem o agente exógeno será menor para vetores retrovirais derivados de HEK293 do que para vetores retrovirais HEK293-HLA-G, indicando que não há inativação sérica de vetores retrovirais HEK293-HLA-G ou uma resposta imunológica adaptativa.

EXEMPLO 13: MEDIÇÃO DE ANTICORPOS IGG E IGM PRÉ-EXISTENTES REATIVOS CONTRA VETORES RETROVIRAIS

[001096] Esse Exemplo descreve a quantificação de titulações de anticorpos de vetor antirretroviral pré-existentes medidos com o uso de citometria de fluxo.

[001097] Em algumas modalidades, uma medida de imunogenicidade para um vetor retroviral são as respostas de anticorpos. Os anticorpos que reconhecem o vetor retroviral podem se ligar de modo que possa limitar a atividade ou longevidade do vetor retroviral. Em uma modalidade, alguns receptores de um vetor retroviral descrito no presente documento terão anticorpos pré-existentes que se ligam ao vetor retroviral e o reconhecem.

[001098] Nesse Exemplo, as titulações de anticorpos do vetor antirretroviral são testadas com o uso do vetor retroviral produzido usando uma célula de origem xenogênica. Nesse Exemplo, um camundongo naïve de vetor retroviral é avaliado quanto à presença de anticorpos de vetor antirretroviral. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos,

macacos com otimização do protocolo.

[001099] O controle negativo é soro de camundongo que foi depletado de IgM e IgG, e o controle positivo é soro derivado de um camundongo que recebeu múltiplas injeções de vetor retroviral gerado a partir de uma célula de origem xenogênica.

[001100] A fim de avaliar a presença de anticorpos pré-existentes que se ligam ao vetor retroviral, o soro de camundongos sem vetor retroviral é primeiramente descomplementado por aquecimento a 56 °C durante 30 minutos e posteriormente diluído em 33% em PBS que contém FCS a 3% e NaN3 a 0,1%. Quantidades iguais de soros e suspensões de vetor retroviral ($1x10^2$ a $1x10^8$ vetores retrovirais por ml) são incubadas durante 30 minutos a 4 °C e lavadas com PBS através de uma almofada de soro de bezerro.

[001101] Os anticorpos xenorreativos IgM são marcados por incubação do vetor retroviral com anticorpos de cabra conjugados com PE específicos para a porção Fc de IgM de camundongo (BD Bioscience) a 4 °C durante 45 minutos. Notavelmente, anticorpos secundários IgG1 ou IgG2 anticamundongo também podem ser usados. Os vetores retrovirais de todos os grupos são lavados duas vezes com PBS que contém FCS a 2% e, em seguida, analisados em um sistema FACS (BD Biosciences). Os dados de fluorescência são coletados por meio de amplificação logarítmica e expressos como intensidade fluorescente média.

[001102] Em uma modalidade, o soro de controle negativo mostrará fluorescência desprezível comparável aos controles sem soro ou secundários isolados. Em uma modalidade, o controle positivo mostrará mais fluorescência que o controle negativo e mais que os controles sem soro ou secundários isolados. Em uma modalidade, nos casos em que ocorre imunogenicidade, o soro de camundongos sem vetor retroviral exibirá mais fluorescência que o controle negativo. Em uma mo-

dalidade, nos casos em que a imunogenicidade não ocorre, o soro de camundongos sem vetor retroviral apresentará fluorescência semelhante em comparação ao controle negativo.

EXEMPLO 14: MEDIÇÃO DE RESPOSTAS DE ANTICORPOS IGG E IGM APÓS MÚLTIPLAS ADMINISTRAÇÕES DE VETORES RETROVIRAIS

[001103] Esse Exemplo descreve a quantificação da resposta humoral de um vetor retroviral modificado após múltiplas administrações do vetor retroviral modificado. Em uma modalidade, um vetor retroviral modificado, por exemplo, modificado por um método descrito no presente documento, terá uma resposta humoral reduzida (por exemplo, reduzida em comparação com a administração de um vetor retroviral não modificado) após múltiplas (por exemplo, mais de uma, por exemplo, 2 ou mais) administrações do vetor retroviral modificado.

[001104] Em algumas modalidades, uma medida de imunogenicidade para um vetor retroviral são as respostas de anticorpos. Em uma modalidade, as injeções repetidas de um vetor retroviral podem levar ao desenvolvimento de anticorpos de vetor antirretroviral, por exemplo, anticorpos que reconhecem o vetor retroviral. Em uma modalidade, os anticorpos que reconhecem o vetor retroviral podem se ligar de modo que possam limitar a atividade ou longevidade do vetor retroviral.

[001105] Nesse Exemplo, as titulações de anticorpos do vetor antirretroviral são examinadas após uma ou mais administrações do vetor retroviral. O vetor retroviral é produzido por qualquer um dos Exemplos anteriores. Nesse exemplo, os retrovirais são criados a partir de: células que não são submetidas à transfecção com uma proteína imunomoduladora (NMCs), células que são submetidas à transfecção com cDNA HLA-G ou HLA-E (doravante NMC-HLA-G) e células submetidas à transfecção com um controle de vetor vazio (doravante vetor sem NMC, controle negativo). Em algumas modalidades, os vetores retrovi-

rais são derivados de células que expressam outras proteínas imunorregulatórias.

[001106] O soro é retirado de coortes diferentes: camundongos injetados sistemicamente e/ou localmente com 1, 2, 3, 5, 10 injeções de veículo (grupo naïve de vetor retroviral), vetor retroviral NMC, vetor retroviral NMC-HLA-G ou vetor retroviral de vetores sem NMC.

[001107] A fim de avaliar a presença e abundância de anticorpos de vetores antirretrovirais, o soro dos camundongos é primeiramente descomplementado por aquecimento a 56 °C durante 30 minutos e, subsequentemente, diluído em 33% em PBS com FCS a 3% e NaN3 a 0,1%. Quantidades iguais de soro e vetor retroviral ($1x10^2$ a $1x10^8$ de vetor retroviral por ml) são incubadas durante 30 minutos a 4 °C e lavadas com PBS através de uma almofada de soro de bezerro.

[001108] Os anticorpos de IgM reativos ao vetor retroviral são marcados por incubação do vetor retroviral com anticorpos de cabra conjugados com PE específicos para a porção Fc de IgM de camundongo (BD Bioscience) a 4 °C durante 45 minutos. Notavelmente, anticorpos secundários IgG1 ou IgG2 anticamundongo também podem ser usados. Os vetores retrovirais de todos os grupos são lavados duas vezes com PBS que contém FCS a 2% e, em seguida, analisados em um sistema FACS (BD Biosciences). Os dados de fluorescência são coletados por meio de amplificação logarítmica e expressos como intensidade fluorescente média.

[001109] Em uma modalidade, os vetores retrovirais NMC-HLA-G terão titulações de anticorpos de IgM antivirais (ou IgG1/2) diminuídos (conforme medido pela intensidade de fluorescência em FACS) após as injeções, em comparação aos vetores retrovirais NMC ou vetores retrovirais vazios de NMC.

EXEMPLO 15: MEDIÇÃO DE RESPOSTAS DE ANTICORPOS DE TITULAÇÃO DE IGG E IGM PARA CÉLULAS RECEPTORAS DE

VETOR RETROVIRAL

[001110] Esse Exemplo descreve a quantificação de titulações de anticorpos contra células receptoras (células que se fundiram com vetores retrovirais) usando citometria de fluxo. Em algumas modalidades, uma medida da imunogenicidade das células receptoras é a resposta do anticorpo. Os anticorpos que reconhecem as células receptoras podem se ligar de modo que possam limitar a atividade celular ou longevidade. Em uma modalidade, as células receptoras não serão alvejadas por uma resposta de anticorpo, ou uma resposta de anticorpo estará abaixo de um nível de referência.

[001111] Nesse Exemplo, as titulações de anticorpos anticélulas receptoras em um indivíduo (por exemplo, humano, rato ou macaco) são testados. Além disso, o protocolo pode ser adaptado a qualquer tipo de célula para o qual existam marcadores de superfície adequados. Nesse exemplo, a célula receptora de destino é uma célula CD3+.

[001112] Os camundongos são tratados com vetores retrovirais produzidos por qualquer um dos métodos descritos no presente pedido ou com PBS (controle negativo) diariamente durante 5 dias. 28 dias após o tratamento final, o sangue periférico é coletado de camundongos que receberam vetores retrovirais e de camundongos que receberam tratamento com PBS. O sangue é coletado em 1 ml de PBS que contém EDTA a 5 μM e misturado imediatamente para evitar a coagulação. Os tubos são mantidos em gelo, e os glóbulos vermelhos são removidos com uma solução tamponada de cloreto de amônio (ACK). As células são marcadas com um anticorpo CD3-FITC de murganho (Catálogo Thermo Fisher n°: 11-0032-82), a 4 °C durante 30 minutos no escuro, após ser bloqueado com albumina sérica bovina durante 10 minutos. Após serem lavados duas vezes com PBS, as células são analisadas em um LSR II (BD Biosciences, San José, CA) com excitação a laser 488 nm e emissão coletada a 530+/-30nm executando o software

FACSDiva™ (BD Biosciences, San José, CA). As células CD3+ são classificadas.

[001113] As células CD3+ classificadas são, então, marcadas com anticorpos de IgM por incubação da mistura de reação com anticorpos de cabra conjugados com PE específicos para a porção Fc de IgM de camundongo (BD Bioscience) a 4 °C durante 45 minutos. Notavelmente, anticorpos secundários IgG1 ou IgG2 anticamundongo também podem ser usados. As células de todos os grupos são lavadas duas vezes com PBS que contém FCS a 2% e, em seguida, analisados em um sistema FACS (BD Biosciences). Os dados de fluorescência são coletados por meio de amplificação logarítmica e expressos como intensidade fluorescente média. A intensidade média de fluorescência é calculada para as células CD3 classificadas de camundongos tratados com vetores retrovirais e os camundongos tratados com PBS.

[001114] Uma baixa intensidade média de fluorescência é indicativa de uma resposta humoral baixa contra as células receptoras. Espera-se que camundongos tratados com PBS tenham baixa intensidade média de fluorescência. Em uma modalidade, a intensidade média de fluorescência será semelhante para células receptoras de camundongos tratados com vetores retrovirais e camundongos tratados com PBS.

EXEMPLO 16: MEDIÇÃO DA RESPOSTA FAGOCÍTICA ÀS CÉLULAS RECEPTORAS DO VETOR RETROVIRAL

[001115] Esse Exemplo descreve a quantificação da resposta de macrófagos contra células receptoras com um ensaio de fagocitose.

[001116] Em algumas modalidades, uma medida da imunogenicidade das células receptoras é a resposta do macrófago. Os macrófagos se envolvem na fagocitose, engolfando células e permitindo o sequestro e destruição de invasores estranhos, tais como bactérias ou células mortas. Em algumas modalidades, a fagocitose de células receptoras

por macrófagos reduziria sua atividade.

[001117] Em uma modalidade, as células receptoras não são alvejadas por macrófagos. Nesse Exemplo, a resposta do macrófago contra células receptoras em um indivíduo é testada. Além disso, o protocolo pode ser adaptado a qualquer tipo de célula para o qual existam marcadores de superfície adequados. Nesse exemplo, a célula receptora de destino é uma célula CD3+.

[001118] Os camundongos são tratados com vetores retrovirais produzidos por qualquer um dos métodos descritos no presente pedido ou com PBS (controle negativo) diariamente durante 5 dias. 28 dias após o tratamento final, o sangue periférico é coletado de camundongos que receberam vetores retrovirais e de camundongos que receberam tratamento com PBS. O sangue é coletado em 1 ml de PBS que contém EDTA a 5 μM e misturado imediatamente para evitar a coagulação. Os tubos são mantidos em gelo, e os glóbulos vermelhos são removidos com uma solução tamponada de cloreto de amônio (ACK).

[001119] As células são marcadas com um anticorpo CD3-FITC de murganho (Catálogo Thermo Fisher n°: 11-0032-82), a 4 °C durante 30 minutos no escuro, após ser bloqueado com albumina sérica bovina durante 10 minutos. Após serem lavados duas vezes com PBS, as células são analisadas em um LSR II (BD Biosciences, San José, CA) com excitação a laser 488 nm e emissão coletada a 530+/-30nm executando o software FACSDiva™ (BD Biosciences, San José, CA). Em seguida, as células CD3+ são classificadas.

[001120] Um ensaio de fagocitose é executado para avaliar a depuração imune mediada por macrófagos de acordo com o protocolo a seguir. Os macrófagos são semeados imediatamente após a colheita em pratos confocais com fundo de vidro. Os macrófagos são incubados em DMEM+10%FBS+1%P/S durante 1 horas para fixação. Uma quantidade apropriada de células CD3+ classificadas e marcadas com

FITC derivadas de camundongos que receberam vetores retrovirais e PBS são adicionados aos macrófagos, conforme indicado no protocolo, e é incubada durante 2 horas, por exemplo, conforme descrito no encarte de informações do produto Kit de Ensaio de Fagocitose Vybrant™ (Molecular Probes, revisado em 18 de março de 2001, encontrado em tools.thermofisher.com/content/sfs/manuals/mp06694.pdf).

[001121] Após 2 horas, a prato é cuidadosamente lavada, e a fluorescência intracelular é examinada. Para identificar macrófagos, as células são incubadas primeiramente com o anticorpo bloqueador do receptor Fc (número de catálogo da eBioscence Nº 14-0161-86, clone 93) durante 15 minutos em gelo para bloquear a ligação de mAbs identificados a receptores Fc, que são expressos abundantemente em macrófagos. Seguindo essa etapa, os anticorpos anti-F4/80-PE (número de catálogo de ThermoFisher 12-4801-82, clone BM8) e anti-CD11b-PerCP-Cy5.5 (número de catálogo BD Biosciences 550993, clone M1/70) conjugados são adicionados para marcar antígenos de superfície de macrófagos. As células são incubadas durante 30 minutos no escuro a 4 °C seguido por centrifugação e lavagem em PBS. As células são então ressuspensas em PBS. Em seguida, a citometria de fluxo de amostras é realizada, e os macrófagos são identificados por meio de sinal de fluorescência positiva para F4/80-PE e CD11b-PerCP-Cy5.5 usando excitação a laser 533 nm e 647 nm, respectivamente. Após o bloqueio para macrófagos, a fluorescência intracelular emitida por células receptoras engolfadas é avaliada por excitação a laser 488 nm. O número de macrófagos fagocitóticos positivos é quantificado usando software de imagem. Os dados são expressos como o índice fagocítico = (número total de células engolfadas/número total de macrófagos contados) × (número de macrófagos contendo células engolfadas/número total de macrófagos contados) × 100.

[001122] Um baixo índice fagocítico é indicativo de baixa fagocitose e alvejamento por macrófagos. Os camundongos tratados com PBS devem ter um índice fagocítico baixo. Em uma modalidade, o índice fagocítico será semelhante para células receptoras derivadas de camundongos tratados com vetores retrovirais e camundongos tratados com PBS.

EXEMPLO 17: MEDIÇÃO DA RESPOSTA DE PBMC A CÉLULAS RECEPTORAS DE VETOR RETROVIRAL

[001123] Esse Exemplo descreve a quantificação de uma resposta de PBMC contra células receptoras com um ensaio de lise celular.

[001124] Em algumas modalidades, uma medida da imunogenicidade das células receptoras é a resposta de PBMC. Em uma modalidade, a lise celular mediada por citotoxicidade de células receptoras por PBMCs é uma medida de imunogenicidade, pois a lise reduzirá, por exemplo, inibir ou interromper a atividade de um vetor retroviral.

[001125] Em uma modalidade, as células receptoras não elicitam uma resposta de PBMC. Neste Exemplo, a resposta de PBMC contra células receptoras em um indivíduo é testada.

[001126] Além disso, o protocolo pode ser adaptado a qualquer tipo de célula para o qual existam marcadores de superfície adequados. Nesse exemplo, a célula receptora de destino é uma célula CD3+.

[001127] Os camundongos são tratados com vetores retrovirais produzidos por qualquer um dos métodos descritos no presente pedido ou com PBS (controle negativo) diariamente durante 5 dias. 28 dias após o tratamento final, o sangue periférico é coletado de camundongos que receberam vetor retroviral e de camundongos que receberam tratamento com PBS. O sangue é coletado em 1 ml de PBS que contém EDTA a 5 μM e misturado imediatamente para evitar a coagulação. Os tubos são mantidos em gelo, e os glóbulos vermelhos são removidos com uma solução tamponada de cloreto de amônio (ACK). As células

são marcadas com um anticorpo CD3:APC-Cy7 de murganho (n° de Catálogo da Biolegend: 100330) ou um anticorpo de controle de isótipo APC-Cy7 (IC: APC-Cy7) (n° de Catálogo da Biolegend: 400230) a 4 °C durante 30 minutos no escuro, após ter Fc bloqueado (n° de Catálogo da Biolegend: 101319) em tampão de marcação de células (n° de Catálogo da Biolegend: 420201) por 10 minutos. Após serem lavadas duas vezes com PBS, as células são analisadas em um FACS Aria (BD Biosciences, San Jose, CA.) com excitação a laser de 640 nm e emissão coletada a 780 -/+ 60 nm executando o software FACSDiva™ (BD Biosciences, San Jose, CA) para definir portas negativas usando as células marcadas com anticorpo APC-Cy7 de controle de isótipo e, em seguida, as células positivas APC-Cy7 são classificadas e coletadas. Em seguida, as células CD3+ classificadas são marcadas com marcação de membrana de plasma CellMask™ Green (CMG, Número de Catálogo da ThermoFisher: C37608) ou DMSO como o controle negativo.

[001128] 7 dias antes do isolamento de células CD3+ dos camundongos tratados com vetor retroviral ou PBS, PBMCs são isoladas de camundongos tratados com vetor retroviral ou PBS de acordo com os métodos em Crop et al. Cell transplantation (20): 1.547 a 1.559; 2011 e simulado na presença de proteína de camundongo recombinante IL-2 (R&D Número de Catálogo da Systems: 402-ML-020) e microesferas CD3/CD28 (Número de Catálogo da ThermoFisher: 11456D) em uma placa de 96 poços de fundo redondo por 7 dias a 37 °C. No dia 7, as PBMCs estimuladas são coincubadas com células de controle CD3+/CMG+ ou CD3+/DMSO para 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 24, 48 horas em uma razão de plaqueamento de PBMC: CD3+/CMG+ ou PBMC: células de controle CD3+/DMSO variando de 1.000:1 a 1:1 e 1:1,25 a 1:1.000. Como controle negativo, um conjunto de poços receberia apenas células de controle CD3+/CMG+ e CD3+/DMSO, sem

PBMCs. Após a incubação, as placas são centrifugadas e processadas de modo que sejam marcadas com o anticorpo CD3:APC-Cy7 de murganho ou um anticorpo IC:APC-Cy7 conforme acima. Depois de serem lavadas duas vezes com PBS, as células são ressuspensas em PBS e analisadas em um FACS Aria (APC-Cy7: excitação/emissão de laser a 640 nm coletada a 780 -/+ 60 nm e excitação/emissão de laser CMG a 561 nm coletada a 585 -/+ 16 nm) executando o software FACSDiva™ (BD Biosciences, San Jose, CA). Os dados do evento FSC/SSC seriam então usados inicialmente para definir a porta para eventos rotulados como “células”. Essa porta de “células” seria então usada para exibir eventos para definir a tensão PMT para as amostras de análise de laser de 640 nm e 561 nm marcadas com IC: APC-Cy7/DMSO apenas. Essa amostra também seria usada para definir as portas para células negativas para APC-Cy7 e CMG. As células CD3+/CMG+ que não receberam nenhuma PBMC seriam então usadas para definir as portas positivas para células CD3+ e CMG+.

[001129] Os dados são analisados observando a porcentagem de células CD3+/CMG+ na população de células totais. Ao comparar os grupos de tratamento, uma porcentagem relativamente mais baixa de células CD3+/CMG+ em qualquer razão de ensaio de PBMC:células CD3+/CMG+ é indicativa de lise celular receptora. Em uma modalidade, a porcentagem de CD3+/CMG+ será semelhante para células receptoras derivadas de camundongos tratados com vetor retroviral e camundongos tratados com PBS.

EXEMPLO 18: MEDIÇÃO DA RESPOSTA DAS CÉLULAS NK ÀS CÉLULAS RECEPTORAS DO VETOR RETROVIRAL

[001130] Este Exemplo descreve a quantificação de uma resposta de células exterminadoras naturais contra células receptoras com um ensaio de lise celular.

[001131] Em algumas modalidades, uma medida da imunogenicida-

de das células receptoras é a resposta da célula exterminadora natural. Em uma modalidade, a lise celular mediada por citotoxicidade de células receptoras por células exterminadoras naturais é uma medida de imunogenicidade, uma vez que a lise irá reduzir, por exemplo, inibir ou parar, a atividade de um vetor retroviral.

[001132] Em uma modalidade, as células receptoras não elicitam uma resposta de células exterminadoras naturais. Nesse Exemplo, a resposta exterminadora natural contra células receptoras em um indivíduo é testada. Além disso, o protocolo pode ser adaptado a qualquer tipo de célula para o qual existam marcadores de superfície adequados. Nesse exemplo, a célula receptora de destino é uma célula CD3+.

[001133] Os camundongos são tratados com vetor retroviral, uma amostra de sangue é coletada, e as células CD3+ são classificadas conforme descrito acima no Exemplo 17. As células NK são isoladas, cultivadas com as células CD3+ e analisadas por FACS de acordo com o protocolo descrito acima no Exemplo 17, com exceção de que as células NK são usadas no lugar das células PBMC usadas no Exemplo 17.

[001134] Os dados são analisados observando a porcentagem de células CD3+/CMG+ na população de células totais. Ao comparar os grupos de tratamento, uma porcentagem relativamente menor de células CD3+/CMG+ em qualquer razão de ensaio de células NK:células CD3+/CMG+ é indicativa de lise celular receptora. Em uma modalidade, a porcentagem de CD3+/CMG+ será semelhante para células receptoras derivadas de camundongos tratados com vetor retroviral e camundongos tratados com PBS.

EXEMPLO 19: MEDIÇÃO DA RESPOSTA DAS CÉLULAS T CD8 ÀS CÉLULAS RECEPTORAS DO VETOR RETROVIRAL

[001135] Este Exemplo descreve a quantificação de uma resposta de células T CD8+ contra células receptoras (células que se fundiram

com vetores retrovirais) com um ensaio de lise celular.

[001136] Em algumas modalidades, uma medida da imunogenicidade das células receptoras é a resposta das células T CD8+. Em uma modalidade, a lise celular mediada por citotoxicidade de células receptoras por células T CD8+ é uma medida de imunogenicidade, uma vez que a lise irá reduzir, por exemplo, inibir ou parar, a atividade de um vetor retroviral.

[001137] Em uma modalidade, as células receptoras não elicitam uma resposta de células T CD8+. Nesse Exemplo, a resposta de células T CD8+ contra células receptoras em um indivíduo é testada. Além disso, o protocolo pode ser adaptado a qualquer tipo de célula para o qual existam marcadores de superfície adequados. Nesse exemplo, a célula receptora de destino é uma célula CD3+.

[001138] Os camundongos são tratados com vetor retroviral, uma amostra de sangue é coletada, e as células CD3+ são classificadas conforme descrito acima no Exemplo 17. As células T CD8+ são isoladas, cultivadas com as células CD3+ e analisadas por FACS de acordo com o protocolo descrito acima no Exemplo 17, com exceção de que as células T CD8+ são usadas no lugar das células PBMC usadas no Exemplo 17.

[001139] Os dados são analisados observando a porcentagem de células CD3+/CMG+ na população de células totais. Ao comparar os grupos de tratamento, uma porcentagem relativamente menor de células CD3+/CMG+ em qualquer razão de ensaio de células CD8+:células CD3+/CMG+ é indicativa de lise celular receptora. Em uma modalidade, a porcentagem de CD3+/CMG+ será semelhante para células receptoras derivadas de camundongos tratados com vetores retrovirais e camundongos tratados com PBS.

EXEMPLO 20: MEDIÇÃO DA ATIVIDADE ESPECÍFICA DE PROMOTOR DO FÍGADO

[001140] Esse exemplo descreve a medição da atividade de um promotor específico do fígado (um TCSRE positivo) em hepatócitos em comparação a células não alvo. Nesse exemplo, as células não alvo são células CD11c+.

[001141] As células CD11c+ são coletadas de camundongos conforme descrito em Li et al., *Journal of Immunology* 2008, 2.483 a 2.493. Os hepatócitos são derivados de camundongos, conforme descrito em Li et al., *Methods in Molecular Biology* 2010, 185 a 196.

[001142] Os dois tipos de células são cultivados separadamente e tratados com um vetor retroviral produzido, conforme descrito no presente documento. O vetor retroviral é pseudotipado com um VSV-G e codifica o repórter de proteína fluorescente tdtomato sob o controle de um promotor específico do fígado, por exemplo, um promotor específico do fígado da Tabela 3.

[001143] Dois dias após a transdução, a expressão do gene nas células é medida por meio de citometria de fluxo, e o número médio de cópias do vetor nas células é medido com PCR quantitativa. A expressão média do gene tdtomato por célula na população de células é normalizada para o número de cópias do vetor da população.

[001144] Em algumas modalidades, a população de hepatócitos terá uma razão entre a expressão de tdtomato e o número de cópias do vetor maior que a população de células CD11c+. Isso demonstrará que o promotor específico do fígado é mais ativo nas células do fígado.

EXEMPLO 21: MEDIÇÃO DA MUDANÇA NA EXPRESSÃO DO microRNA RESTRITIVO

[001145] Esse exemplo descreve a medição da atividade de um microRNA restritivo de células hematopoiéticas (um NTCSRE) em hepatócitos em comparação a células hematopoiéticas. Nesse exemplo, as células hematopoiéticas são células CD11c+.

[001146] As células CD11c+ são coletadas de camundongos con-

forme descrito em Li et al., *Journal of Immunology* 2008, 2.483 a 2.493. Os hepatócitos são derivados de camundongos, conforme descrito em Li et al., *Methods in Molecular Biology* 2010, 185 a 196.

[001147] Os dois tipos de células são cultivados separadamente e tratados com um vetor retroviral produzido, conforme descrito no presente documento. O vetor retroviral é pseudotipado com um VSV-G e codifica para o repórter de proteína tdtomato fluorescente sob o controle de um promotor ubiquamente ativo e um microRNA restritivo a células hematopoiéticas, por exemplo, um microRNA da Tabela 4.

[001148] Dois dias após a transdução, a expressão do gene nas células é medida por meio de citometria de fluxo, e o número médio de cópias do vetor nas células é medido com PCR quantitativa. A expressão média do gene tdtomato por célula na população de células é normalizada para o número de cópias do vetor da população.

[001149] Em algumas modalidades, a população de hepatócitos terá uma razão entre a expressão de tdtomato e o número de cópias do vetor maior que a população de células CD11c+. Isso demonstrará que o microRNA restritivo de células hematopoiéticas diminui a expressão em células CD11c+.

EXEMPLO 22: TRATAMENTO COM tdTOMATO COM VSV-G, PROMOTOR ESPECÍFICO DE HEPATÓCITOS E microRNA RESTRITIVO HEMATOPOIÉTICO

[001150] Esse Exemplo descreve a quantificação da expressão de um agente exógeno em células-alvo receptoras e células receptoras não alvo por expressão de agente exógeno em células únicas.

[001151] Em uma modalidade, os camundongos tratados terão expressão de agente exógeno em células não alvo semelhante àquelas de camundongos não tratados. Em uma modalidade, camundongos tratados terão uma porcentagem semelhante de células não alvo que expressam o agente exógeno em comparação àquelas de camundon-

gos não tratados. Em uma modalidade, camundongos tratados terão expressão de agente de proteína exógena em células-alvo maior que aquela de camundongos não tratados. Em uma modalidade, os camundongos tratados terão uma porcentagem de células-alvo que expressam o agente de proteína exógena maior que aquela de camundongos não tratados.

[001152] Nesse exemplo, a célula não alvo é uma célula CD11c+. Nesse exemplo, a célula-alvo é um hepatócito. No entanto, esse protocolo pode ser adaptado a qualquer tipo de célula para o qual existem marcadores de superfície adequados e que podem ser isolados do indivíduo. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo. Nesse exemplo, o agente exógeno é uma proteína fluorescente, e a expressão é medida por meio de citometria de fluxo. Em outras modalidades, a expressão de um agente de proteína exógena pode ser medida com imunomarcação para a proteína. Em outras modalidades, a expressão do agente de proteína exógena pode ser medida por meio de microscopia ou Western blot.

[001153] Os camundongos são tratados com vetor retroviral pseudotipado com VSV-G e que portam um agente de proteína fluorescente tdtomato sob o controle de um promotor hepatócio-específico (um TCSRE positivo) e sequência de microRNA restritivo de célula hematopeiética (um NTCSRE). O vetor retroviral é produzido por meio de qualquer um dos métodos descritos no presente documento. Os camundongos de controle negativo são tratados com PBS.

[001154] 28 dias após o tratamento, os camundongos são sacrificados e as células CD11c+ são coletadas de camundongos conforme descrito em Li et al., *Journal of Immunology* 2008, 2.483 a 2.493. Os hepatócitos são derivados de camundongos, conforme descrito em Li et al., *Methods in Molecular Biology* 2010, 185 a 196. A expressão de

Tdtomato nas células isoladas é testada por meio de citometria de fluxo em um FACS Aria (BD Biosciences, San José, CA.) executando o software FACSDiva™ (BD Biosciences, San José, CA). Uma porta negativa para expressão de tdtomato é definida com hepatócitos e células CD11c+ isoladas de camundongos tratados com veículo e com uma excitação a laser de 552 nm e uma emissão coletadas a 585 -/+ 42 nm.

[001155] A porcentagem de células CD11c+ que são tdtomato-positivas é medida. Em algumas modalidades, a porcentagem de células CD11c+ que são tdtomato-positivas será semelhante em células de camundongos tratados e não tratados. O nível mediano de fluorescência de tdtomato é medido em células CD11c+. Em algumas modalidades, o nível de fluorescência mediano de tdtomato em células CD11c+ será semelhante em células de camundongos tratados e não tratados. A porcentagem de hepatócitos que são tdtomato positive é medida. Em algumas modalidades, a porcentagem de hepatócitos que são tdtomato-positivos será maior em células em camundongos tratados do que em camundongos não tratados. O nível de fluorescência mediano de tdtomato é medido em hepatócitos. Em algumas modalidades, o nível de fluorescência mediano tdtomato em hepatócitos será maior em células de camundongos tratados do que camundongos não tratados. Em algumas modalidades, a porcentagem de hepatócitos tdtomato-positivos será maior que a porcentagem de CD11c+ que são tdtomato-positivos em células de camundongos tratados. Em algumas modalidades, o nível de fluorescência tdtomato mediano de hepatócitos será maior que o de CD11c+ em células de camundongos tratados.

EXEMPLO 23: TRATAMENTO PARA DEFICIÊNCIA DE ORNITINA TRANSCARBAMILASE COM PSEUDOTIPO DE VSV-G IN VITRO

[001156] Esse exemplo descreve a entrega de um transgene terapêutico às células *in vitro*. Nesse exemplo, o transgene terapêutico é

ornitina transcarbamilase (*otc*).

[001157] Os hepatócitos são derivados de camundongos spf^{ash}. Os hepatócitos são isolados conforme descrito em Li et al., *Methods in Molecular Biology* 2010, 185 a 196. Os hepatócitos são transduzidos com um vetor retroviral produzido conforme descrito no presente documento ou com PBS. O vetor retroviral é pseudotipado com VSV-G e carrega o gene *otc* sob o controle de um promotor específico de hepatócitos (um TCSRE positivo) e uma sequência de microRNA restritiva de células hematopoiéticas (um TCSRE negativo).

[001158] Após tempo suficiente para a expressão de OTC, as células são preparadas para a imagem. As células são fixadas, permeabilizadas, bloqueadas e imunomarcadas com um anticorpo anti-OTC (por exemplo, número de catálogo abcam ab91418). Após a imunomarcação, as células são contramarcadas com um anticorpo secundário conjugado a Alexa Fluor 488 (por exemplo, número de catálogo abcam ab150077). As células são fotografadas em um microscópio confocal Zeiss LSM 710 com um objetivo de imersão em óleo 63x, ao mesmo tempo que são mantidas a 37 °C e 5% de CO2. Alexa Fluor é submetido à excitação com laser de 488 nm teve a emissão capturada a 510±15 nm. A intensidade média da farinha Alexa por célula é calculada para determinar o nível de expressão de OTC por célula. Para cada grupo, pelo menos 30 a 40 células são fotografadas e analisadas.

[001159] Em algumas modalidades, o nível de expressão de OTC por célula será maior em hepatócitos tratados com o vetor retroviral que codifica o gene OTC do que em hepatócitos tratados com PBS.

EXEMPLO 24: TRATAMENTO PARA DEFICIÊNCIA DE ORNITINA TRANSCARBAMILASE COM RETROVÍRUS PSEUDOTIPADO VSV-G IN VIVO

[001160] Esse exemplo descreve a entrega de um transgene terapêutico às células *in vivo*. Nesse exemplo, o transgene terapêutico é

ornitina transcarbamilase (*otc*).

[001161] Os camundongos spf^{ash} são tratados com vetor retroviral pseudotipado com VSV-G e portando o agente do gene da *otc* sob o controle de um promotor específico do hepatócito (um TCSRE positivo) e uma sequência de microRNA restritiva de células hematopoiéticas (um NTCSRE). O vetor retroviral é produzido por meio de qualquer um dos métodos descritos no presente pedido. Os camundongos de controle negativo são tratados com PBS.

[001162] 28 dias após o tratamento, os hepatócitos são obtidos de camundongos tratados com retrovírus ou PBS e marcados para expressão de OTC, conforme descrito nos exemplos anteriores. Em algumas modalidades, o nível de expressão de OTC por célula será maior em hepatócitos derivados de camundongos tratados com o vetor retroviral que codifica o gene OTC do que em hepatócitos tratados com PBS.

[001163] Em um grupo separado de camundongos, 28 dias após o tratamento com retrovírus ou PBS, os camundongos são alimentados com uma dieta rica em proteínas por 1 semana. Os níveis de amônia no sangue e o ácido orótico urinário são medidos. Em algumas modalidades, o nível de amônia no sangue e/ou ácido orótico urinário será menor em camundongos tratados com retrovírus do que em camundongos tratados com PBS. Os camundongos são mantidos em dieta rica em proteínas por mais 28 dias. Em algumas modalidades, no grupo tratado com vetor retroviral sobreviverão mais camundongos durante todo o período de dieta rica em proteínas de 28 dias do que camundongos no grupo tratado com PBS. Em algumas modalidades, os camundongos do grupo tratado com vetor retroviral terão um tempo de sobrevivência significativamente maior que os camundongos do grupo tratado com PBS.

EXEMPLO 25: FALTA DE ATIVIDADE TRANSCRICIONAL EM FU-

SOSSOMAS

[001164] Esse Exemplo quantifica a atividade transcricional em fusossomas em comparação às células parentais, por exemplo, células de origem, usadas para geração de fusossomas. Em uma modalidade, a atividade transcricional será baixa ou estará ausente em fusossomas em comparação às células parentais, por exemplo, células de origem.

[001165] Os fusossomas são um chassi para a entrega do agente terapêutico. Agentes terapêuticos, tais como miRNA, mRNAs, proteínas e/ou organelas que podem ser entregues às células ou ambientes de tecidos locais com alta eficiência podem ser usados para modular as vias que não são normalmente ativas ou ativas em níveis patológicos baixos ou altos no tecido receptor. Em uma modalidade, a observação de que os fusossomas não têm capacidade para transcrição, ou de que os fusossomas têm atividade transcricional menor que a célula parental dos mesmos, demonstrará que a remoção do material nuclear ocorreu de maneira suficiente.

[001166] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. Em seguida, uma quantidade suficiente de fusossomas e de células parentais usados para gerar os fusossomas são colocados em uma placa com múltiplos poços de fixação baixa de 6 poços em DMEM que contém 20% de soro fetal bovino, 1x Penicilina/Estreptomicina e o alquino-nucleosídeo EU que pode receber tag fluorescente por 1 hora a 37 °C e 5% de CO2. Para controles negativos, uma quantidade suficiente de fusossomas e células parentais também é colocada em placa com múltiplos poços em DMEM que contém 20% de soro fetal bovino, 1x penicilina/estreptomicina, porém sem alquino-nucleosídeo EU.

[001167] Após a incubação de 1 hora, as amostras são processadas seguindo as instruções do fabricante para um kit de imagem (ThermoFisher Scientific). As amostras de células e fusossoma, incluindo os

controles negativos, são lavadas três vezes com tampão PBS 1x e ressuspensas em tampão 1xPBS e analisadas por citometria de fluxo (Becton Dickinson, San José, CA, EUA) com o uso de um laser de argônio 488 nm para excitação e o 530 +/- Emissão de 30 nm. O software BD FACSDiva foi usado para aquisição e análise. Os canais de dispersão de luz são definidos em ganhos lineares e os canais de fluorescência são definidos em escala logarítmica, com um mínimo de 10.000 células analisadas em cada condição.

[001168] Em uma modalidade, a atividade transcricional medida pela emissão de 530 +/- 30 nm nos controles negativos será nula devido à omissão do alquino-nucleosídeo EU. Em algumas modalidades, os fusossomas terão menos de aproximadamente 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1% ou menos atividade transcricional do que as células parentais.

[001169] Consultar, também, Proc Natl Acad Sci U S A, 2008, 14 de outubro 14;105(41):15.779 a 15.784. doi: 10.1073/pnas.0808480105. Publicação eletrônica de 7 de outubro de 2008.

EXEMPLO 26: FALTA DE REPLICAÇÃO DE DNA OU ATIVIDADE DE REPLICAÇÃO

[001170] Esse exemplo quantifica a replicação de DNA em fusossomas. Em uma modalidade, os fusossomas replicarão o DNA a uma taxa baixa em comparação às células.

[001171] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. A atividade de replicação do DNA da célula parental e do fusossoma é avaliada pela incorporação de um nucleotídeo que pode receber tag fluorescente (ThermoFisher Scientific n° C10632). Fusossomas e um número equivalente de células são incubados com EdU a uma concentração final de 10 μM por 2 horas, após a preparação de uma solução-mãe de EdU com dimetilsulfóxido. As amostras são então fixadas por 15 minutos usando PFA

3,7%, lavadas com tampão 1xPBS, pH 7,4 e permeabilizadas por 15 min em solução de detergente a 0,5% em tampão 1xPBS, pH 7,4.

[001172] Após a permeabilização, fusossomas e células em suspensão em tampão PBS que contém 0,5% de detergente são lavados com tampão 1xPBS, pH 7,4 e incubados durante 30 minutos a 21 °C em coquetel de reação, tampão 1xPBS, CuSO4 (Componente F), azida-fluor 488, 1x aditivo de tampão de reação.

[001173] Um controle negativo para fusossoma e atividade de replicação de DNA celular é feito com amostras tratadas da mesma forma que acima, porém sem azida-fluor 488 no coquetel de reação 1x.

[001174] As amostras de células e fusossomas são então lavadas e ressuspensas em tampão 1xPBS e analisadas por citometria de fluxo. A citometria de fluxo é feita com um citômetro FACS (Becton Dickinson, San José, CA, EUA) com excitação a laser de argônio 488 nm, e um espectro de emissão 530 +/- 30 nm é coletado. O software de análise FACS é usado para aquisição e análise. Os canais de dispersão de luz são definidos em ganhos lineares e os canais de fluorescência são definidos em escala logarítmica, com um mínimo de 10.000 células analisadas em cada condição. A atividade de replicação do DNA relativa é calculada com base na intensidade média de azida-fluor 488 em cada amostra. Todos os eventos são capturados nos canais de dispersão direta e lateral (alternativamente, um portão pode ser aplicado para selecionar apenas a população fusossoma). O valor de intensidade de fluorescência normalizado para os fusossomas é determinado subtraindo do valor de intensidade de fluorescência mediana do fusossoma o valor de intensidade de fluorescência mediana da respectiva amostra de controle negativo. Em seguida, a atividade de replicação de DNA relativa normalizada para as amostras de fusossomas é normalizada para as respectivas amostras de células nucleadas, a fim de gerar medições quantitativas para a atividade de replica-

ção de DNA.

[001175] Em uma modalidade, os fusossomas têm menos atividade de replicação de DNA do que as células parentais.

[001176] Consultar, também, Salic, 2.415 a 2.420, doi: 10.1073/pnas.0712168105.

EXEMPLO 27: QUANTIFICAÇÃO DE FUSOGÊNIOS

[001177] Esse exemplo descreve a quantificação do número absoluto de fusogênios por fusossoma.

[001178] Uma composição de fusossoma é produzida por qualquer um dos métodos descritos nos Exemplos anteriores, exceto o fusossoma é projetado conforme descrito em um Exemplo anterior para expressar um fusogênio (VSV-G) com tag de com GFP. Além disso, um fusossoma de controle negativo é projetado sem que fusogênio (VSV-G) ou GFP estejam presentes.

[001179] Os fusossomas com o fusogênio com tag de GFP e o controle negativo (ou controles negativos) são, em seguida, testados quanto ao número absoluto de fusogênios conforme a seguir. GFP recombinante adquirido comercialmente é diluído em série para gerar uma curva de calibração de concentração de proteína. A fluorescência GFP da curva de calibração e uma amostra de fusossomas de quantidade conhecida é então medida em um fluorímetro com o uso de um cubo de luz GFP (filtro de excitação 469/35 e um filtro de emissão 525/39) para calcular a concentração molar média de moléculas de GFP em a preparação fusossoma. Em seguida, a concentração molar é convertida no número de moléculas de GFP e dividida pelo número de fusossomas por amostra para atingir um número médio de moléculas de fusogênio marcadas com GFP por fusossoma e, assim, fornece uma estimativa relativa do número de fusogênios por fusossoma.

[001180] Em uma modalidade, a fluorescência de GFP será maior nos fusossomas com tag GFP em comparação aos controles negati-

vos, em que nenhum fusogênio ou GFP está presente. Em uma modalidade, GFP a fluorescência é relativa ao número de moléculas de fusogênio presentes.

[001181] Alternativamente, fusossomas individuais são isolados com o uso de um sistema de preparação de célula única (Fluidigm) de acordo com as instruções do fabricante, e a qRT-PCR é realizada com o uso de um conjunto de sondas disponíveis comercialmente (Taqman) e a mistura principal projetada para quantificar níveis de cDNA de fusogênio ou de GFP com base no valor C_t. Um padrão de RNA da mesma sequência que o fragmento clonado do gene de fusogênio ou do gene de GFP é gerado por síntese (Amsbio) e, em seguida, adicionado à reação experimental de qRT-PCR do sistema de preparação de célula única em diluições em série para estabelecer uma curva padrão de C_t vs. a concentração de RNA de fusogênio ou de GFP.

[001182] O valor C_t dos fusossomas é comparado à curva-padrão para determinar a quantidade de RNA de fusogênio ou de GFP por fusossoma.

[001183] Em uma modalidade, o RNA de fusogênio e de GFP será maior nos fusossomas submetidos à engenharia para expressar os fusogênios em comparação aos controles negativos, em que nenhum fusogênio ou GFP está presente.

[001184] Os fusogênios podem ser quantificados adicionalmente na bicamada lipídica analisando-se a estrutura da bicamada lipídica, conforme descrito anteriormente, e quantificando-se os fusogênios na bicamada lipídica por LC-MS, conforme descrito em outros exemplos no presente documento.

EXEMPLO 28: MEDIÇÃO DO TAMANHO MÉDIO DOS FUSOSSOMAS

[001185] Esse exemplo descreve a medição do tamanho médio dos

fusossomas.

[001186] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. Os fusossomas medidos para determinar o tamanho médio com o uso dos sistemas disponíveis comercialmente (iZON Science). O sistema é usado com software de acordo com as instruções do fabricante e um nanoporo projetado para analisar partículas na faixa de tamanho de 40 nm a 10 μm. Fusossomas e células parentais são suspensos novamente em solução salina tamponada com fosfato (PBS) para uma faixa de concentração final de 0,01 a 0,1 μg de proteína/ml. Outras configurações do instrumento são ajustadas, conforme indicado na seguinte tabela:

TABELA 6: PARÂMETROS E DEFINIÇÕES DA MEDIÇÃO DE FUSOSSOMA

Parâmetro de Medição	Definição
Pressão	6
Tipo de nanoporo	NP300
Amostra de calibração	CPC400_6P
Análise do padrão-ouro	não
Assistente de captura	nenhum

[001187] Todos os fusossomas são analisados dentro de 2 horas de isolamento. Em uma modalidade, os fusossomas terão um tamanho dentro de aproximadamente 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que as células parentais.

EXEMPLO 29: MEDIÇÃO DA DISTRIBUIÇÃO DE TAMANHO MÉDIO DE FUSOSSOMAS

[001188] Esse exemplo descreve a medição da distribuição de tamanho de fusossomas.

[001189] Fusossomas são gerados por qualquer um dos métodos

descritos nos Exemplos anteriores e são testados para determinar o tamanho médio das partículas com o uso de um sistema disponível comercialmente, tal conforme descrito em um Exemplo anterior. Em uma modalidade, os limites de tamanho para 10%, 50% e 90% dos fusossomas centrados em torno da mediana são comparados às células parentais para avaliar a distribuição de tamanho dos fusossomas.

[001190] Em uma modalidade, os fusossomas terão menos que aproximadamente 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% ou menos da variabilidade da célula parental em tamanho distribuição em 10%, 50% ou 90% da amostra.

EXEMPLO 30: VOLUME MÉDIO DE FUSOSSOMAS

[001191] Esse exemplo descreve a medição do volume médio dos fusossomas. Sem querer se limitar à teoria, a variação do tamanho (por exemplo, volume) dos fusossomas pode tornar os mesmos versáteis para carregamento de carga, projeto terapêutico ou aplicação distintos.

[001192] Os fusossomas são preparados conforme descrito nos exemplos anteriores. O controle positivo são células HEK293 ou esferas de poliestireno com um tamanho conhecido. O controle negativo são células HEK293 que são passadas por uma agulha de calibre 36 aproximadamente 50 vezes.

[001193] A análise com um microscópio eletrônico de transmissão, conforme descrito em um Exemplo anterior, é usada para determinar o tamanho dos fusossomas. O diâmetro do fusossoma é medido, e o volume é, então, calculado.

[001194] Em uma modalidade, os fusossomas terão um tamanho médio de aproximadamente 50 nm ou maior em diâmetro.

EXEMPLO 31: DENSIDADE MÉDIA DE FUSOSSOMAS

[001195] A densidade de fusossoma é medida por meio de um ensaio de centrifugação de gradiente de sacarose contínuo, conforme

descrito em Théry et al., Curr Protoc Cell Biol. abril de 2006; Capítulo 3: Unidade 3.22. Os fusossomas são obtidos conforme descrito nos exemplos anteriores.

[001196] Primeiramente, um gradiente de sacarose é preparado. Uma solução 2 M e com sacarose a 0,25 é gerada misturando-se 4 ml da solução-mãe de HEPES/sacarose e 1 ml de solução-mãe de HEPES ou 0,5 ml de solução-mãe de HEPES/sacarose e 4,5 ml de solução-mãe de HEPES, respectivamente. Essas duas frações são carregadas no gerador de gradiente com todas as venezianas fechadas, a solução de sacarose 2 M no compartimento proximal com uma barra de agitação magnética e a solução de sacarose 0,25 M no compartimento distal. O fabricante de gradiente é colocado em uma placa de agitação magnética, o obturador entre os compartimentos proximal e distal é aberto e a placa de agitação magnética é ligada. A solução-mãe HEPES é produzida de acordo com o seguinte: 2,4 g de ácido N-2-hidroxietilpiperazina-N'-2-etanossulfônico (HEPES; 20 mM final), 300 H2O, ajustar o pH para 7,4 com NaOH 10 N e finalmente ajustar o volume para 500 ml com H2O. A solução-mãe de HEPES/sacarose é produzida de acordo com o seguinte: 2,4 g de ácido hidroxietilpiperazina-N'-2-etanossulfônico (HEPES; 20 mM final), 428 g de sacarose livre de protease (ICN; 2,5 M final), 150 ml de H2O, ajustar o pH para 7,4 com NaOH 10 N e finalmente ajuste o volume para 500 ml com H2O.

[001197] Os fusossomas são suspensos novamente em 2 ml de solução-mãe HEPES/sacarose e são despejados no fundo de um tubo de centrífuga SW 41. O tubo externo é colocado no tubo SW 41, logo acima dos 2 ml dos fusossomas. A veneziana externa é aberta e um gradiente contínuo de sacarose de 2 M (parte inferior) a 0,25 M (parte superior) é lentamente derramado no topo dos fusossomas. O tubo SW 41 é abaixado à medida que o gradiente é vertido, de modo que o tubo esteja sempre ligeiramente acima do topo do líquido.

[001198] Todos os tubos com gradientes são equilibrados entre si ou com outros tubos com o mesmo peso de soluções de sacarose. Os gradientes são centrifugados de um dia para o outro (≥14 horas) a 210.000 xg, 4 °C, no rotor de balde oscilante SW 41 com o freio definido como baixo.

[001199] Com uma micropipeta, onze frações de 1 ml, de cima para baixo, são coletadas e colocadas em um tubo de 3 ml para o rotor TLA-100.3. As amostras são reservadas e, em poços separados de uma placa de 96 poços, 50 µl de cada fração são usados para medir o índice de refração. A placa é coberta com película adesiva para evitar a evaporação e armazenada durante, no máximo, 1 hora sob temperatura ambiente. Um refratômetro é usado para medir o índice de refração (logo, a concentração de sacarose e a densidade) de 10 a 20 µl de cada fração do material armazenado na placa de 96 poços.

[001200] Uma tabela para converter o índice de refração em g/ml está disponível no catálogo de ultracentrifugação, disponível para transferência por download no site da Beckman.

[001201] Cada fração é, então, preparada para análise do teor da proteína. Dois mililitros de HEPES 20 mM, pH 7,4, são adicionados a cada fração de gradiente de 1 ml e misturados por pipetagem para cima e para baixo duas a três vezes. Um lado de cada tubo é marcado com um marcador permanente, e os tubos são colocados com o lado marcado para cima em um rotor TLA-100.3.

[001202] Os tubos de 3 ml com frações diluídas são centrifugados durante 1 hora a 110.000 xg, 4 °C. O rotor TLA-100.3 comporta seis tubos, então duas centrifugações para cada gradiente são realizadas com os outros tubos mantidos a 4 °C até que possam ser centrifugados.

[001203] O sobrenadante é aspirado de cada um dos tubos de 3 ml, deixando uma gota no topo do pélete. O pélete provavelmente não é

visível, porém sua localização pode ser inferida a partir da marca no tubo. O sedimento invisível é suspenso novamente e transferido para tubos de microcentrífuga. Metade de cada fração ressuspensa é usada para análise de teor da proteína por ensaio de ácido bicinconínico, descrito em outro exemplo. Isso fornece uma distribuição entre as várias frações de gradiente da preparação fusossoma. Essa distribuição é usada para determinar a densidade média dos fusossomas. A segunda metade do volume é armazenada a -80 °C e usada para outros fins (por exemplo, análise funcional ou purificação adicional por imunoisolamento), uma vez que a análise de proteína revelou a distribuição de fusossoma entre as frações.

[001204] Em uma modalidade, com o uso desse ensaio, a densidade média dos fusossomas será 1,25 g/ml +/- 0,05 de desvio padrão. Em uma modalidade, a densidade média dos fusossomas estará na faixa de 1 a 1,1, 1,05 a 1,15, 1,1 a 1,2, 1,15 a 1,25, 1,2 a 1,3 ou 1,25 a 1,35. Em uma modalidade, a densidade média dos fusossomas será inferior a 1 ou superior a 1,35.

EXEMPLO 32: MEDIÇÃO DO CONTEÚDO DO ENVELOPE NUCLEAR

[001205] Esse exemplo descreve uma medição do teor do envelope nuclear em fusossomas enucleados. O envelope nuclear isola o DNA do citoplasma da célula.

[001206] Em uma modalidade, uma composição de fusossoma purificada compreende uma célula de mamífero, tal como HEK-293Ts (293 [HEK-293] (ATCC® CRL-1573™), que foi enucleada, conforme descrito no presente documento. Esse exemplo descreve a quantificação de diferentes proteínas de membrana nuclear como uma representação para medir a quantidade de membrana nuclear intacta que permanece após a geração de fusossoma.

[001207] Nesse exemplo, HEK-293Ts a $10x10^6$ e a quantidade equi-

valente de fusossomas preparados a partir de HEK-293Ts a $10x10^6$ são fixadas durante 15 minutos com o uso de 3,7% de PFA, lavadas com tampão 1xPBS, pH 7,4 e permeabilizadas simultaneamente e, em seguida, bloqueadas durante 15 minutos com o uso do tampão 1xPBS que contém 1% de albumina sérica bovina e 0,5% de Triton® X-100, pH 7,4. Após a permeabilização, os fusossomas e células são incubados durante 12 horas a 4 °C com diferentes anticorpos primários, por exemplo (o anticorpo anti-RanGAP1 [EPR3295] (Abcam - ab92360), o anticorpo anti-NUP98 [EPR6678] - marcador de poro nuclear (Abcam - ab124980), anticorpo de proteínas de complexo de poros antinucleares [Mab414] - (Abcam- ab24609), anticorpo anti-importina 7 (Abcam - ab213670), nas concentrações sugeridas pelo fabricante em tampão 1xPBS que contém 1% de albumina sérica bovina e 0,5% de Triton® X-100, pH 7,4. Os fusossomas e células são, então, lavados com tampão 1xPBS, pH 7,4, e incubados durante 2 horas a 21 °C com um anticorpo secundário fluorescente apropriado que detecta o anticorpo primário especificado anteriormente nas concentrações sugeridas pelo fabricante diluídas em tampão 1xPBS que contém 1% de albumina de soro bovino e 0,5% de detergente, pH 7,4. Os fusossomas e as células são, então, lavados com tampão 1xPBS, suspenso novamente em tampão 1xPBS a 300 μl, pH 7,4, que contém 1 μg/ml de Hoechst 33342, filtrado através de um tubo FACS de 20 μm e analisado por citometria de fluxo.

[001208] Os controles negativos são gerados usando o mesmo procedimento de marcação, porém sem adição de anticorpo primário. A citometria de fluxo é realizada em um citômetro FACS (Becton Dickinson, San Jose, CA, EUA) com excitação com laser de argônio 488 nm, e um espectro de emissão 530 +/- 30 nm é coletado. O software de aquisição FACS é usado para aquisição e análise. Os canais de dispersão de luz são definidos em ganhos lineares e os canais de fluo-

rescência são definidos em escala logarítmica, com um mínimo de 10.000 células analisadas em cada condição. O teor relativo da membrana nuclear intacta é calculado com base na intensidade média de fluorescência em cada amostra. Todos os eventos são capturados nos canais de dispersão direta e lateral.

[001209] O valor de intensidade de fluorescência normalizado para os fusossomas é determinado subtraindo do valor de intensidade de fluorescência mediana do fusossoma o valor de intensidade de fluorescência mediana da respectiva amostra de controle negativo. Em seguida, a fluorescência normalizada para as amostras de fusossomas é normalizada para as respectivas amostras de células nucleadas, a fim de gerar medições quantitativas do teor da membrana nuclear intacta.

[001210] Em uma modalidade, os fusossomas enucleados compreenderão menos de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% da intensidade de fluorescência ou conteúdo do envelope nuclear em comparação às células parentais nucleadas.

EXEMPLO 33: MEDIÇÃO DOS NÍVEIS DE CROMATINA

[001211] Esse exemplo descreve a medição da cromatina em fusossomas enucleados.

[001212] O DNA pode ser condensado em cromatina para permitir que seja encaixado dentro do núcleo. Em uma modalidade, uma composição de fusossoma purificada, conforme produzida por qualquer um dos métodos descritos no presente documento, compreenderá baixos níveis de cromatina.

[001213] Fusossomas enucleados preparados por qualquer um dos métodos descritos anteriormente e células de controle positivo (por exemplo, células parentais) são testados quanto ao teor de cromatina com o uso de um ELISA com anticorpos que são específicos para proteína histona H3 ou proteína histona H4. As histonas são o principal

componente de proteína da cromatina, em que H3 e H4 são as proteínas histonas predominantes.

[001214] As histonas são extraídas da preparação de fusossoma e da preparação de células com o uso de um kit comercial (por exemplo, Kit de Extração de Histona Abcam (ab113476)) ou outros métodos conhecidos na técnica. Essas alíquotas são armazenadas a -80 °C até o uso. Uma diluição em série do padrão é preparada diluindo-se a proteína histona purificada (H3 ou H4) de 1 a 50 ng/µl em uma solução do tampão de ensaio. O tampão de ensaio pode ser derivado de um kit fornecido por um fabricante (por exemplo, Kit de quantificação total Abcam Histone H4 (ab156909) ou Kit de quantificação total Abcam Histone H3 (ab115091)). O tampão de ensaio é adicionado a cada poço de uma placa de 48 ou 96 poços, que é revestida com um anticorpo anti-histona H3 ou anti-H4, e a amostra ou controle padrão é adicionado ao poço para trazer o volume total de cada poço a 50 µl. A placa é, então, coberta e incubada a 37 graus durante 90 a 120 minutos.

[001215] Após a incubação, qualquer histona ligada ao anticorpo anti-histona fixada à placa é preparada para detecção. O sobrenadante é aspirado e a placa é lavada com 150 µl de tampão de lavagem. Em seguida, o tampão de captura, que inclui um anticorpo de captura anti-histona H3 ou anti-H4, é adicionado à placa em um volume de 50 µl e a uma concentração de 1 µg/ml. A placa é, então, incubada à temperatura ambiente em um agitador orbital por 60 minutos.

[001216] Em seguida, a placa é aspirada e lavada 6 vezes com o uso do tampão de lavagem. A molécula repórter de sinal ativável pelo anticorpo de captura é, então, adicionada a cada poço. A placa é tapada e incubada à temperatura ambiente durante 30 minutos. A placa é, então, aspirada e lavada 4 vezes com o uso de um tampão de lavagem. A reação é interrompida pela adição de solução de terminação. A absorvância de cada poço na placa é lida a 450 nm e a concentração de

histonas em cada amostra é calculada de acordo com a curva padrão de absorvância a 450 nm vs. concentração de histona em amostras padrão.

[001217] Em uma modalidade, as amostras de fusossoma compreenderão menos de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% da concentração de histonas das células parentais nucleadas.

EXEMPLO 34: MEDIÇÃO DO CONTEÚDO DE miRNA EM FUSOSSOMAS

[001218] Esse exemplo descreve a quantificação de microRNAs (miRNAs) em fusossomas. Em uma modalidade, um fusossoma compreende miRNAs.

[001219] MiRNAs são elementos regulatórios que, entre outras atividades, controlam a taxa pela qual RNAs mensageiros (mRNAs) são traduzidos em proteínas. Em uma modalidade, fusossomas carregando miRNA podem ser usados para entregar o miRNA a locais alvo.

[001220] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. O RNA de fusossomas ou células parentais é preparado conforme descrito anteriormente. Pelo menos um gene de miRNA é selecionado a partir do Registro de mIRNA do Centro Sanger em www.sanger.ac.uk/Software/Rfam/mirna/ index.shtml. miRNA é preparado conforme descrito em Chen *et al*, *Nucleic Acids Research*, 33(20), 2005. Todos os ensaios de miRNA TaqMan estão disponíveis através da Thermo Fisher (A25576, Waltham, MA).

[001221] qPCR é realizado de acordo com as especificações do fabricante no miRNA cDNA, e os valores C_T são gerados e analisados com o uso de um sistema de PCR em tempo real, conforme descrito no presente documento.

[001222] Em uma modalidade, o teor de miRNA de fusossomas será

pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou superior que aquele das células parentais.

EXEMPLO 35: QUANTIFICAÇÃO DA EXPRESSÃO DE UM RNA ENDÓGENO OU RNA SINTÉTICO EM FUSOSSOMAS

[001223] Esse exemplo descreve a quantificação dos níveis de RNA endógeno com expressão alterada, ou um RNA sintético que é expresso em um fusossoma.

[001224] O fusossoma ou célula parental é projetado para alterar a expressão de um RNA endógeno ou sintético que medeia uma função celular para os fusossomas.

[001225] Os vetores de transposase (System Biosciences, Inc.) incluem o quadro aberto de leitura do gene de resistência à puromicina juntamente com um quadro aberto de leitura de um fragmento clonado de um agente de proteína. Os vetores são eletroporados em 293Ts com o uso de um eletroporador (Amaxa) e um kit de transfecção nuclear específico de linhagem celular 293T (Lonza).

[001226] Após a seleção com puromicina por 3 a 5 dias em DMEM que contém 20% de Soro Fetal Bovino e 1x Penicilina/Estreptomicina, os fusossomas são preparados a partir da linhagem celular de expressão estável por qualquer um dos métodos descritos nos Exemplos anteriores.

[001227] Os fusossomas individuais são isolados e o agente de proteína ou RNA por fusossoma é quantificado, conforme descrito em um Exemplo anterior.

[001228] Em uma modalidade, os fusossomas terão pelo menos 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3, 5,0 x 10^3, 10^4, 5,0 x 10^4, 10^5, 5,0 x 10^5, 10^6, 5,0 x 10^6 ou mais do RNA por fusossoma.

EXEMPLO 36: MEDIÇÃO DA COMPOSIÇÃO PROTEÔMICA EM FUSOSSOMAS

[001229] Esse exemplo descreve a quantificação da composição de

proteína de fusossomas. Em uma modalidade, a composição de proteína de fusossomas será semelhante às células das quais eles são derivados.

[001230] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. Os fusossomas são suspensos novamente em tampão de lise (ureia 7 M, tioureia 2 M, Chaps a 4% (p/v) em Tris 50 mM pH 8,0) e incubados durante 15 minutos à temperatura ambiente com vórtice ocasional. Em seguida, as misturas são lisadas por sonicação por 5 minutos em um banho de gelo e centrifugadas por 5 minutos a 13.000 RPM. O teor de proteína é determinado por um ensaio colorimétrico (Pierce), e a proteína de cada amostra é transferida para um novo tubo e o volume é equalizado com Tris 50 mM pH 8.

[001231] As proteínas são reduzidas durante 15 minutos a 65 Celsius com DTT 10 mM e alquiladas com iodoacetamida 15 mM durante 30 minutos à temperatura ambiente no escuro. As proteínas são precipitadas com adição gradual de 6 volumes de acetona fria (-20 Celsius) e incubadas de um dia para o outro a -80 Celsius. Péletes de proteína são lavadas 3 vezes com metanol frio (-20 Celsius). As proteínas são ressuspensas em Tris 50 mM pH 8,3.

[001232] Em seguida, a tripsina/lysC é adicionada às proteínas durante as primeiras 4 horas de digestão a 37 °C com agitação. As amostras são diluídas com Tris 50 mM pH 8, e o desoxicolato de sódio 0,1% é adicionado com mais tripsina/lysC para digestão durante a noite a 37 °C com agitação. A digestão é interrompida e o desoxicolato de sódio é removido pela adição de ácido fórmico a 2% v/v. As amostras são submetidas a vórtice e limpas por centrifugação durante 1 minuto a 13.000 RPM. Os peptídeos são purificados por extração em fase sólida de fase reversa (SPE) e secos. As amostras são reconstituídas em 20 µl de DMSO a 3%, ácido fórmico a 0,2% em água e analisadas por

LC-MS.

[001233] Para ter medições quantitativas, um padrão de proteína também é executado no instrumento. Peptídeos padrão (Pierce, equimolar, grau LC-MS, n° 88342) são diluídos a 4, 8, 20, 40 e 100 fmol/ul e são analisados por LC-MS/MS. A AUC (área sob a curva) média dos 5 melhores peptídeos por proteína (3 MS/transição de MS/peptídeo) é calculada para cada concentração para gerar uma curva padrão.

[001234] A aquisição é realizada com um espectrômetro de massa de alta resolução (ABSciex, Foster City, CA, EUA) equipado com uma interface de eletroaspersão com um capilar iD de 25 µm e acoplado a cromatografia líquida de micro-ultra-alta performance (µUHPLC) (Eksigent, Redwood City, CA, EUA). O software de análise é usado para controlar o instrumento e para processamento e aquisição de dados. A tensão da fonte é ajustada para 5,2 kV e mantida a 225 °C, o gás de cortina é ajustado para 0,19 MPa (27 psi), o gás um a 0,08 MPa (12 psi) e o gás dois a 0,07 MPa (10 psi). A aquisição é realizada no modo de Aquisição Dependente de Informações (IDA) para o banco de dados de proteínas e no modo de aquisição SWATH para as amostras. A separação é realizada em uma coluna de fase reversa de 0,3 µm i.d., partículas de 2,7 µm, 150 mm de comprimento (Advance Materials Technology, Wilmington, DE) que é mantida a 60 °C. As amostras são injetadas por transbordamento de alça em uma alça de 5 µl. Para o gradiente de LC de 120 minutos (amostras), a fase móvel inclui o seguinte: solvente A (ácido fórmico a 0,2% em v/v e DMSO a 3% em v/v em água) e solvente B (ácido fórmico a 0,2% em v/v e DMSO a 3% em EtOH) a uma taxa de fluxo de 3 µl/min.

[001235] Para a quantificação absoluta das proteínas, uma curva padrão (5 pontos, R2> 0,99) é gerada com o uso da soma da AUC dos 5 melhores peptídeos (3 íons MS/MS por peptídeo) por proteína. Para gerar um banco de dados para a análise das amostras, o algoritmo

DIAUmpire é executado em cada uma das 12 amostras e combinado com os arquivos MGF de saída em um banco de dados. Esse banco de dados é utilizado com software (ABSciex) para quantificar as proteínas em cada uma das amostras, com o uso de 5 transições/peptídeos e 5 peptídeos/proteínas no máximo. Um peptídeo é considerado adequadamente medido, caso a pontuação calculada seja superior a 1,5 ou tenha um FDR <1%. A soma da AUC de cada um dos peptídeos medidos adequadamente é mapeada na curva padrão e é relatada como fmol.

[001236] Os dados de quantificação de proteína resultantes são, então, analisados para determinar os níveis de proteína e as proporções de classes conhecidas de proteínas de acordo com o seguinte: as enzimas são identificadas como proteínas que são observadas com um número da Comissão de Enzimas (EC) number; as proteínas associadas a ER são identificadas como proteínas que tiveram uma classificação do compartimento celular de Ontologia Genética (GO; http://www.geneontology.org) de ER e não mitocôndrias; proteínas associadas a exossomas são identificadas como proteínas que têm uma classificação de compartimento celular de Ontologia Genética de exossomas e não mitocôndrias; e as proteínas mitocondriais são identificadas como proteínas que são identificadas como mitocondriais no banco de dados MitoCarta (Calvo et al., NAR 2015l doi:10.1093/nar/gkv1003). As razões molares de cada uma dessas categorias são determinadas como a soma das quantidades molares de todas as proteínas em cada classe dividida pela soma das quantidades molares de todas as proteínas identificadas em cada amostra.

[001237] A composição proteômica de fusossoma é comparada à composição proteômica da célula parental. Em uma modalidade, composições proteômicas semelhantes entre fusossomas e células parentais serão observadas quando > 50% das proteínas identificadas esti-

verem presentes no fusossoma e, dentre essas proteínas identificadas, o nível é > 25% do nível de proteína correspondente na célula parental.

EXEMPLO 37: MEDIÇÃO DE GAPDH EM FUSOSSOMAS

[001238] Esse ensaio descreve a quantificação do nível de gliceraldeído 3-fosfato desidrogenase (GAPDH) nos fusossomas e o nível relativo de GAPDH nos fusossomas em comparação às células parentais.

[001239] A GAPDH é medida nas células parentais e nos fusossomas com o uso de um ELISA padrão comercialmente disponível para GAPDH (ab176642, Abcam) de acordo com as instruções do fabricante.

[001240] Os níveis de proteína total são medidos de forma similar por meio do ensaio de ácido bicinconínico, conforme descrito anteriormente no mesmo volume de amostra usado para medir GAPDH. Nas modalidades, com o uso desse ensaio, o nível de GAPDH por proteína total nos fusossomas será <100 ng de GAPDH/µg de proteína total. Da mesma forma, nas modalidades, a diminuição nos níveis de GAPDH em relação à proteína total das células parentais para os fusossomas será maior que uma diminuição de 10%.

[001241] Em uma modalidade, o teor de GAPDH na preparação em ng GAPDH/µg de proteína total será inferior a 500, inferior a 250, inferior a 100, inferior a 50, inferior a 20, inferior a 10, inferior a 5 ou inferior a 1.

[001242] Em uma modalidade, a diminuição em GAPDH por proteína total em ng/µg da célula parental para a preparação será superior a 1%, superior a 2,5%, superior a 5%, superior a 10%, superior a 15%, mais superior a 20%, superior a 30%, superior a 40%, superior a 50%, superior a 60%, superior a 70%, superior a 80% ou superior a 90%.

EXEMPLO 38: MEDIÇÃO DE CALNEXINA EM FUSOSSOMAS

[001243] Esse ensaio descreve a quantificação do nível de calnexina

(CNX) nos fusossomas e do nível relativo de CNX nos fusossomas em comparação às células parentais.

[001244] A calnexina é medida nas células iniciais e na preparação com o uso de um ELISA padrão comercialmente disponível para calnexina (MBS721668, MyBioSource) de acordo com as instruções do fabricante.

[001245] Os níveis de proteína total são medidos de forma similar por meio do ensaio de ácido bicinconínico, conforme descrito anteriormente no mesmo volume de amostra usado para medir a calnexina. Nas modalidades, o uso desse ensaio, o nível de calnexina por proteína total nos fusossomas será <100 ng de calnexina/µg de proteína total. Da mesma forma, nas modalidades, o aumento nos níveis de calnexina em relação à proteína total da célula parental para os fusossomas será maior que um aumento de 10%.

[001246] Em uma modalidade, o teor de calnexina na preparação em ng calnexina/µg de proteína total será inferior a 500, 250, 100, 50, 20, 10, 5 ou 1.

[001247] Em uma modalidade, a diminuição da calnexina por proteína total em ng/µg da célula parental para a preparação será mais do que 1%, 2,5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%.

EXEMPLO 39: COMPARAÇÃO DA MASSA SOLÚVEL DA PROTEÍNA À MASSA INSOLÚVEL DA PROTEÍNA

[001248] Esse exemplo descreve a quantificação da razão solúvel:insolúvel da massa da proteína em fusossomas. Em uma modalidade, a razão solúvel:insolúvel de massa da proteína em fusossomas será semelhante às células nucleadas.

[001249] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. A preparação de fusossoma é testada para determinar a razão proteína solúvel:insolúvel com o uso

de um ensaio do ácido bicinconínico padrão (BCA) (por exemplo, com o uso do kit de ensaio de proteína Pierce™ BCA disponível comercialmente, produto Thermo Fischer nº 23.225). As amostras de proteínas solúveis são preparadas suspendendo-se os fusossomas preparados ou células parentais a uma concentração de 1x10^7 células ou fusossomas/ml em PBS e centrifugando-se a 1.600 g para peletizar os fusossomas ou células. O sobrenadante é coletado como a fração de proteína solúvel.

[001250] Os fusossomas ou células no pélete são lisados por pipetagem vigorosa e vórtice em PBS com 2% de Triton-X-100. A fração lisada representa a fração de proteína insolúvel.

[001251] Uma curva padrão é gerada com o uso do BSA fornecido, de 0 a 20 µg de BSA por poço (em triplicada). O fusossoma ou preparação celular é diluído de modo que a quantidade medida esteja dentro da faixa dos padrões. A preparação fusossoma é analisada em triplicada e o valor médio é usado. A concentração de proteína solúvel é dividida pela concentração de proteína insolúvel para produzir a razão proteína solúvel:insolúvel.

[001252] Em uma modalidade, a razão proteína solúvel:insolúvel de fusossoma estará dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais em comparação às células parentais.

EXEMPLO 40: MEDIÇÃO DE LPS EM FUSOSSOMAS

[001253] Esse exemplo descreve a quantificação dos níveis de lipopolissacarídeos (LPS) em fusossomas em comparação às células parentais. Em uma modalidade, os fusossomas terão níveis mais baixos de LPS em comparação às células parentais.

[001254] Os LPS são um componente das membranas bacterianas e um potente indutor de respostas imunes inatas.

[001255] As medições de LPS se baseiam em espectrometria de

massa, conforme descrito nos exemplos anteriores.

[001256] Em uma modalidade, menos de 5%, 1%, 0,5%, 0,01%, 0,005%, 0,0001%, 0,00001% ou menos do teor de lipídios dos fusossomas será LPS.

EXEMPLO 41: RAZÃO ENTRE LIPÍDIOS E PROTEÍNAS EM FUSOSSOMAS

[001257] Esse exemplo descreve a quantificação da razão entre massa de lipídios e massa de proteína em fusossomas. Em uma modalidade, os fusossomas terão uma razão entre massa de lipídios e massa de proteína que é semelhante às células nucleadas.

[001258] O teor total de lipídios é calculado como a soma do teor molar de todos os lipídios identificados nos dados lipidômicos definidos em um exemplo anterior. O teor total da proteína dos fusossomas é medido por meio de ensaio de ácido bicinconínico, conforme descrito no presente documento.

[001259] Alternativamente, a razão entre lipídeos e proteínas pode ser descrita como uma razão entre uma espécie de lipídio particular e uma proteína específica. A espécie lipídica particular é selecionada a partir dos dados lipidômicos produzidos em um exemplo anterior. A proteína específica é selecionada a partir dos dados proteômicos produzidos em um exemplo anterior. Diferentes combinações de espécies de lipídeos e proteínas selecionadas são usadas para definir razões específicas de lipídio:proteína.

EXEMPLO 42: RAZÃO ENTRE PROTEÍNAS E DNA EM FUSOSSOMAS

[001260] Esse exemplo descreve a quantificação da razão entre massa de proteína e massa de DNA em fusossomas. Em uma modalidade, os fusossomas terão uma razão entre massa de proteína e massa de DNA que é muito maior que as células.

[001261] O teor total da proteína dos fusossomas e células é medido

conforme descrito em um Exemplo anterior. A massa de DNA de fusossomas e células é medida conforme descrito em um Exemplo anterior. A razão entre proteínas e ácidos nucleicos totais é, então, determinada dividindo-se o teor total da proteína pelo teor total do DNA para produzir uma razão dentro de um determinado intervalo para uma preparação típica de fusossoma.

[001262] Alternativamente, a razão entre proteínas e ácidos nucleicos é determinada definindo-se os níveis de ácido nucleico como o nível de um gene de manutenção específico, como GAPDH, com o uso de PCR em tempo real semiquantitativa (RT-PCR).

[001263] Então, a razão entre proteínas e ácidos nucleicos de GAPDH é determinada dividindo-se o teor total da proteína pelo teor total do DNA de GAPDH para definir uma faixa específica de razão proteína:ácido nucleico para uma preparação de fusossoma típica.

EXEMPLO 43: MEDIÇÃO DE FUSÃO COM UMA CÉLULA-ALVO

[001264] Esse exemplo descreve a quantificação da fusão de fusossoma com uma célula-alvo em comparação a uma célula não alvo.

[001265] Em uma modalidade, a fusão de fusossoma com uma célula-alvo permite a entrega específica da célula de uma carga, transportada dentro do lúmen do fusossoma, para o citosol da célula receptora. Os fusossomas produzidos pelos métodos descritos no presente documento são testados quanto à taxa de fusão com uma célula-alvo conforme a seguir.

[001266] Nesse exemplo, o fusossoma compreende uma célula HEK293T que expressa Myomaker em sua membrana plasmática. Além disso, o fusossoma expressa a proteína fluorescente mTagBFP2 e a Cre recombinase. A célula-alvo é uma célula mioblástica que expressa Myomaker e Myomixer, e a célula não alvo é uma célula fibroblástica, que não expressa, não expressa Myomaker tampouco Myomixer. Prevê-se que um fusossoma que expressa o Myomaker es-

tá se funde com a célula-alvo que expressa Myomaker e Myomixer, porém não com a célula não alvo (Quinn et al., 2017, Nature Communications, 8, 15.665. doi.org/10.1038/ncomms15665) (Millay et al., 2013, Nature, 499(7.458), 301 a 305. doi.org/10.1038/nature12343). Ambos os tipos de células-alvo e não alvo são isolados de camundongos e expressam de forma estável o cassete “LoxP-stop-Loxp-tdTomato” sob um promotor CMV que, após a recombinação por Cre ativa a expressão de tdTomato, indicando fusão.

[001267] As células-alvo ou não alvo receptoras são colocadas em uma placa de 96 poços de fundo transparente preto. Ambas as células-alvo e não alvo são colocadas em placa para os diferentes grupos de fusão. Em seguida, 24 horas após as células receptoras serem colocadas em placas, os fusossomas que expressam a proteína Cre recombinase e Myomaker são aplicados às células-alvo ou não alvo receptoras em meio DMEM. A dose de fusossomas é correlacionada ao número de células receptoras colocadas em placa no poço. Após a aplicação dos fusossomas, a placa de células é centrifugada a 400 g durante 5 minutos para ajudar a iniciar o contato entre os fusossomas e as células receptoras.

[001268] Começando quatro horas após a aplicação de fusossoma, os poços de células são imageados para identificar positivamente as células positivas para RFP versus células positivas para GFP no campo ou poço.

[001269] Nesse exemplo, as placas de células são imageadas com o uso de um microscópio automático (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). A população total de células em um determinado poço é determinada marcando-se primeiramente as células com Hoechst 33342 em meio DMEM durante 10 minutos. O Hoechst 33342 marca os núcleos de células intercalando-se em DNA e, portanto, é usado para

identificar células individuais. Após a marcação, o meio Hoechst é substituído por meio DMEM regular.

[001270] A imagem da Hoechst é obtida com o uso de LED de 405 nm e cubo de filtro DAPI. A imagem da GFP é obtida com o uso do LED de 465 nm e cubo de filtro GFP, ao passo que a imagem de RFP é obtida com o uso do LED de 523 nm e do cubo de filtro de RFP. Imagens de poços de células-alvo e não alvo são obtidas estabelecendo-se primeiramente a intensidade do LED e os tempos de integração em um poço de controle positivo; isto é, células receptoras tratadas com adenovírus que codificam para Cre recombinase em vez de fusossomas.

[001271] As configurações de aquisição são definidas de forma que as intensidades de RFP e GFP estejam nos valores máximos de intensidade de pixel, porém não saturadas. Os poços de interesse são, então, fotografados com o uso das configurações estabelecidas. Os poços são fotografados a cada 4 horas para adquirir dados de curso de tempo para taxas de atividade de fusão.

[001272] A análise de poços positivos para GFP e RFP é realizada com software fornecido com o microscópio fluorescente ou outro software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, EUA, rsb.info.nih.gov/ij/, 1997 a 2007).

[001273] As imagens são pré-processadas com o uso de um algoritmo de subtração de fundo de bola rolante com uma largura de 60 μm. A máscara celular total é definida nas células Hoechst-positivas. As células com intensidade Hoechst significativamente acima das intensidades de fundo são submetidas a um limiar, e áreas muito pequenas ou grandes para serem células Hoechst-positivas são excluídas.

[001274] Dentro da máscara de célula total, as células GFP e RFP-positivas são identificadas novamente pelo limiar para células significativamente acima do fundo e estendendo-se as máscaras Hoechst (nú-

cleos) para toda a área celular para incluir toda a fluorescência celular de GFP e RFP. O número de células positivas para RFP identificadas em poços de controle que contêm células-alvo receptoras ou não alvo é usado para subtrair do número de células positivas para RFP nos poços que contêm fusossoma (para subtrair para recombinação Loxp não específica). O número de células positivas para RFP (células receptoras fundidas) é, então, dividido pela soma das células positivas para GFP (células receptoras que não se fundiram) e células positivas para RFP em cada ponto de tempo para quantificar a taxa de fusão de fusossoma dentro da população de células receptoras. A taxa é normalizada para a dose determinada de fusossoma aplicada às células receptoras. Para taxas de fusão alvejada (fusão de fusossoma às células alvejadas), a taxa de fusão à célula não alvo é subtraída da taxa de fusão à célula-alvo a fim de quantificar as taxas de fusão alvejada.

[001275] Em uma modalidade, a taxa média de fusão para os fusossomas com as células-alvo estará na faixa de 0,01 a 4,0 células RFP/GFP por hora para fusão de células-alvo ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que células receptoras não alvo com fusossomas. Em uma modalidade, os grupos sem fusossoma aplicado mostrarão uma taxa de fundo de <0,01 células RFP/GFP por hora.

EXEMPLO 44: FUSÃO IN VITRO PARA ENTREGAR UMA PROTEÍNA DE MEMBRANA

[001276] Esse exemplo descreve a fusão de fusossoma com uma célula *in vitro*. Em uma modalidade, a fusão de fusossoma com uma célula *in vitro* resulta na entrega de uma proteína de membrana ativa para a célula receptora.

[001277] Nesse exemplo, os fusossomas são gerados a partir de uma célula HEK293T que expressa a proteína HVJ-E do vírus Sendai (Tanaka et al., 2015, Gene Therapy, 22(de outubro de 2014), 1 a 8.

doi.org/10.1038/gt.2014.12). Em uma modalidade, os fusossomas são gerados para expressar a proteína de membrana, GLUT4, que é encontrada principalmente nos tecidos musculares e adiposos e é responsável pelo transporte de glicose regulado por insulina para as células. Os fusossomas com e sem GLUT4 são preparados a partir de células HEK293T, conforme descrito por qualquer um dos métodos descritos em um Exemplo anterior.

[001278] As células musculares, tais como células C2C12, são, então, tratadas com fusossomas que expressam GLUT4, fusossomas que não expressam GLUT4, PBS (controle negativo) ou insulina (controle positivo). A atividade de GLUT4 em células C2C12 é medida pela absorção do análogo 2-desoxiglicose fluorescente, 2-[N-(7-nitrobenz-2-oxa-1,3-diaxol-4-il) amino]-2-desoxiglicose (2-NBDG). A fluorescência de células C2C12 é avaliada por meio de microscopia com o uso dos métodos descritos nos exemplos anteriores.

[001279] Em uma modalidade, as células C2C12 que são tratadas com fusossomas que expressam GLUT4 e insulina devem demonstrar fluorescência aumentada em comparação a células C2C12 tratadas com PBS ou fusossomas que não expressam GLUT4.

[001280] Consultar, também, Yang et al., *Advanced Materials* 29, 1605604, 2017.

EXEMPLO 45: MEDIÇÃO DO EXTRAVASAMENTO DOS VASOS SANGUÍNEOS

[001281] Esse exemplo descreve a quantificação de extravasamento de fusossoma através de uma monocamada endotelial, conforme testado com um sistema microfluídico *in vitro* (JS Joen et al. 2013, journals.plos.org/plosone/article?id=10.1371/journal.pone.0056910).

[001282] As células extravasam da vasculatura para o tecido circundante. Sem se ater a qualquer teoria, o extravasamento é uma forma de os fusossomas atingirem os tecidos extravasculares.

[001283] O sistema inclui três canais de mídia endereçáveis independentemente, separados por câmaras nas quais um gel que imita ECM pode ser injetado. Em suma, o sistema microfluídico moldou PDMS (polidimetilsiloxano; Silgard 184; Dow Chemical, MI) através do qual as portas de acesso são perfuradas e ligadas a uma tampa de vidro para formar canais microfluídicos. As dimensões da seção transversal do canal são 1 mm (largura) por 120 µm (altura). Para aumentar a adesão à matriz, os canais de PDMS são revestidos com uma solução de PDL (bromidrato de poli-D-lisina; 1 mg/ml; Sigma-Aldrich, St. Louis, MO).

[001284] Em seguida, a solução de colágeno tipo I (BD Biosciences, San Jose, CA, EUA) (2,0 mg/ml) com solução salina tamponada com fosfato (PBS; Gibco) e NaOH é injetada nas regiões de gel do dispositivo através de quatro portas de enchimento separadas e incubado durante 30 minutos para formar um hidrogel. Quando o gel é polimerizado, o meio de células endoteliais (adquirido de fornecedores como Lonza ou Sigma) é imediatamente pipetado para os canais para evitar a desidratação do gel. Após a aspiração do meio, solução diluída de hidrogel (BD science) (3,0 mg/ml) é introduzida no canal da célula, e o excesso de solução de hidrogel é removido por lavagem com o uso de meio frio.

[001285] As células endoteliais são introduzidas no canal do meio e podem se estabelecer para formar um endotélio. Dois dias após a semeadura das células endoteliais, fusossomas ou células de macrófagos (controle positivo) são introduzidos no mesmo canal no qual as células endoteliais formaram uma monocamada completa. Os fusossomas são introduzidos de forma que aderem e transmigram através da monocamada para a região de gel. As culturas são mantidas em uma incubadora umidificada a 37 °C e 5% CO_2. Uma versão que expressa GFP do fusossoma é usada para permitir a geração de ima-

gens de células vivas por meio de microscopia fluorescente. No dia seguinte, as células são fixadas e marcadas para núcleos com o uso da marcação DAPI na câmara, e várias regiões de interesse são fotografadas com o uso do microscópio confocal para determinar quantos fusossomas passaram através da monocamada endotelial.

[001286] Em uma modalidade, a marcação com DAPI indicará que fusossomas e células de controle positivo têm capacidade para passar através da barreira endotelial após a semeadura.

EXEMPLO 46: MEDIÇÃO DA MOBILIDADE CELULAR QUIMIOTÁTICA

[001287] Esse Exemplo descreve a quantificação da quimiotaxia de fusossoma. As células podem se mover perto ou para longe de um gradiente químico por meio de quimiotaxia. Em uma modalidade, a quimiotaxia permitirá que os fusossomas se alojem em um sítio de lesão ou rastreiem um patógeno. Uma composição de fusossoma purificada produzida por qualquer um dos métodos descritos nos Exemplos anteriores é testada quanto às suas capacidades quimiotáticas conforme a seguir.

[001288] Uma quantidade suficiente de fusossomas ou células de macrófagos (controle positivo) são carregados em um poço de microlâmina de acordo com o protocolo fornecido pelo fabricante em mídia DMEM (ibidi.com/img/cms/products/labware/channel_slides/S_8032X_Chemotaxis/ IN_8032X_Chemotaxis.pdf). Os fusossomas são deixados a 37 °C e 5% de CO2 durante 1 hora para fixação. Após a fixação da célula, DMEM (controle negativo) ou DMEM que contém quimioatraente MCP1 é carregado em reservatórios adjacentes do canal central, e os fusossomas são fotografados continuamente durante 2 horas com o uso de um microscópio de campo amplo invertido Zeiss. As imagens são analisadas com o uso do software ImageJ (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland,

EUA, http://rsb.info.nih.gov/ij/, 1997 a 2007). Os dados de coordenação da migração para cada fusossoma ou célula observada são adquiridos com o plugin de rastreamento manual (Fabrice Cordelières, Institut Curie, Orsay, França). Os gráficos de quimiotaxia e as velocidades de migração são determinados com a ferramenta de quimiotaxia e migração (ibidi).

[001289] Em uma modalidade, a distância média acumulada e a velocidade de migração de fusossomas estarão dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais do que a resposta das células de controle positivo à quimiocina. A resposta de células a uma quimiocina é descrita, por exemplo, em in Howard E. Gendelman et al., *Journal of Neuroimmune Pharmacology*, 4(1): 47 a 59, 2009.

EXEMPLO 47: MEDIÇÃO DO POTENCIAL DE HOMING

[001290] Esse Exemplo descreve o homing de fusossomas a um sítio de lesão. As células podem migrar de um sítio distal e/ou acumular em um sítio específico, por exemplo, a origem de um local. Normalmente, o sítio é um sítio de lesão. Em uma modalidade, os fusossomas se alojarão, por exemplo, migrar para ou se acumular em um sítio de lesão.

[001291] Camundongos C57BL/6J de oito semanas de idade (Jackson Laboratories) são dosados com notexina (NTX) (Accurate Chemical & Scientific Corp), uma miotoxina, em solução salina estéril por injeção intramuscular (IM) com o uso de uma agulha 30G no músculo tibial anterior direito (TA) na concentração de 2 μg/ml. A pele sobre o músculo tibial anterior (TA) é preparada depilando-se a área com um removedor químico de pelos por 45 segundos, seguido de 3 enxágues com água. Essa concentração é escolhida para garantir degeneração máxima das miofibras, assim como danos mínimos às células satélite, aos axônios motores e aos vasos sanguíneos.

[001292] No dia 1 após a injeção de NTX, os camundongos recebem uma injeção IV de fusossomas ou células que expressam a luciferase do vagalume. Os fusossomas são produzidos a partir de células que expressam de forma estável a luciferase do pirilampo por qualquer um dos métodos descritos nos Exemplos anteriores. Um sistema de imageamento por bioluminescência (Perkin Elmer) é usado para obter imagens de bioluminescência de animais inteiros em 0, 1, 3, 7, 21 e 28 após a injeção.

[001293] Cinco minutos antes da imagem, os camundongos recebem uma injeção intraperitoneal de substrato bioluminescente (Perkin Elmer) a uma dose de 150 mg/kg, a fim de visualizar a luciferase. O sistema de imagem é calibrado para compensar todas as configurações do dispositivo. O sinal bioluminescente é medido usando Radiance Photons, com Total Flux usado como um valor medido. A região de interesse (ROI) é gerada cercando-se o sinal da ROI para fornecer um valor em fótons/segundo. Uma ROI é avaliada tanto no músculo de TA tratado com NTX quanto no músculo TA contralateral, e a razão de fótons/segundo entre músculos TA tratados com NTX e não tratados com NTX é calculada como uma medida do homing para o músculo tratado com NTX.

[001294] Em uma modalidade, a razão de fótons/segundo entre os músculos TA tratados com NTX e não tratados com NTX em fusossomas e células será maior que 1 indicando acúmulo específico de sítio de fusossomas que expressam luciferase na lesão.

[001295] Consultar, por exemplo, Plant et al., *Muscle Nerve* 34(5)L 577 a 585, 2006.

EXEMPLO 48: MEDIÇÃO DA ATIVIDADE FAGOCÍTICA

[001296] Esse exemplo demonstra a atividade fagocítica de fusossomas. Em uma modalidade, os fusossomas têm atividade fagocítica, por exemplo, são com capacidade para fagocitose. As células iniciam

a fagocitose, engolfando partículas, o que permite o sequestro e destruição de invasores estranhos, tais como bactérias ou células mortas.

[001297] Uma composição de fusossoma purificada, conforme produzida por qualquer um dos métodos descritos nos Exemplos anteriores, que compreende um fusossoma de um macrófago de mamífero com inativação nuclear parcial ou completa teve capacidade para fagocitose testada por meio de biopartículas de patógenos. Essa estimativa foi feita com o uso de um ensaio de fagocitose fluorescente de acordo com o seguinte protocolo.

[001298] Macrófagos (controle positivo) e fusossomas foram colocados em prato imediatamente após a colheita em pratos confocais de fundo de vidro separados. Os macrófagos e fusossomas foram incubados em DMEM+10%de FBS+1% de P/S durante 1 hora para fixação. A *E. coli* K12 identificada com fluoresceína e a *Escherichia coli* K-12 não identificada com fluoresceína (controle negativo) foram adicionadas aos macrófagos/fusossomas, conforme indicado no protocolo do fabricante, e foram incubadas durante 2 horas, tools.thermofisher.com/content/ sfs/manuais/mp06694.pdf. Após 2 horas, as partículas fluorescentes livres foram extintas pela adição de azul de Tripano. A fluorescência intracelular emitida por partículas engolfadas foi imageada por microscopia confocal a uma excitação 488. O número de fusossomas fagocitóticos positivos foram quantificados com o uso do software de imagem J.

[001299] O número médio de fusossomas fagocitóticos foi de pelo menos 30% 2 horas após a introdução da biopartícula, e foi maior que 30% nos macrófagos de controle positivo.

EXEMPLO 49: MEDIÇÃO DO POTENCIAL DE SECREÇÃO DE PROTEÍNA

[001300] Esse exemplo descreve a quantificação da secreção por fusossomas. Em uma modalidade, os fusossomas terão capacidade

para secreção, por exemplo, secreção de proteína. As células podem descartar ou descarregar material por meio de secreção. Em uma modalidade, os fusossomas interagirão quimicamente e se comunicarão em seu ambiente por meio de secreção.

[001301] A capacidade que os fusossomas têm para secretar uma proteína a uma determinada taxa é determinada com o uso do ensaio de flash de Gaussia luciferase da ThermoFisher Scientific (n° de catálogo 16158). Células de fibroblastos embrionários de camundongo (controle positivo) ou fusossomas produzidos por qualquer um dos métodos descritos nos Exemplos anteriores são incubados em meio de crescimento, e as amostras dos meios são coletadas a cada 15 minutos peletizando-se primeiramente os fusossomas a 1.600 g durante 5 minutos e, em seguida, coletando-se o sobrenadante. As amostras coletadas são pipetadas para uma placa de 96 poços de fundo transparente. Uma solução de trabalho de tampão de ensaio é, então, preparada de acordo com as instruções do fabricante.

[001302] Em suma, a colenterazina, uma luciferina ou molécula emissora de luz, é misturada com o tampão de ensaio instantâneo, e a mistura é pipetada para cada poço da placa de 96 poços contendo as amostras. Os poços de controle negativo que não têm células ou fusossomas incluem meio de crescimento ou tampão de ensaio para determinar o sinal de fundo da Gaussia luciferase. Além disso, uma curva padrão de Gaussia luciferase purificada (Athena Enzyme Systems, n° de catálogo 0308) é preparada a fim de converter o sinal de luminescência em moléculas de secreção de Gaussia luciferase por hora.

[001303] A placa é testada para luminescência, com o uso da integração de 500 ms. O sinal da Gaussia luciferase de fundo é subtraído de todas as amostras e, em seguida, uma curva linear de melhor ajuste é calculada para a curva padrão da Gaussia luciferase. Caso as leituras da amostra não se encaixarem na curva padrão, elas serão diluí-

das de forma adequada e testadas novamente. Com o uso desse ensaio, é determinada a capacidade que os fusossomas têm para secretar a Gaussia luciferase a uma taxa (moléculas/hora) dentro de um determinado intervalo.

[001304] Em uma modalidade, os fusossomas terão capacidade para secretar proteínas a uma taxa de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais que as células de controle positivo.

EXEMPLO 50: MEDIÇÃO DO POTENCIAL DE TRANSDUÇÃO DE SINAL

[001305] Esse exemplo descreve a quantificação da transdução de sinal em fusossomas. Em uma modalidade, os fusossomas são com capacidade para transdução de sinal. As células podem enviar e receber sinais moleculares do ambiente extracelular por meio de cascatas de sinalização, tais como a fosforilação, em um processo conhecido como transdução de sinal. Uma composição de fusossoma purificada como produzida por qualquer um dos métodos descritos nos Exemplos anteriores compreendendo um fusossoma de uma célula de mamífero com inativação nuclear parcial ou completa tem capacidade para transdução de sinal induzida pela insulina. A transdução de sinal induzida pela insulina é avaliada medindo-se os níveis de fosforilação de AKT, uma via principal na cascata de sinalização do receptor de insulina e absorção de glicose em resposta à insulina.

[001306] Para medir a fosforilação de AKT, células, por exemplo, Fibroblastos Embrionários de Camundongo (MEFs) (controle positivo) e fusossomas são colocados em placas de 48 poços e deixados durante 2 horas em um incubador umidificado a 37 °C e 5% de CO_2. Após a adesão das células, a insulina (por exemplo, a 10 nM), ou uma solução de controle negativo sem insulina, é adicionada ao poço contendo células ou fusossomas durante 30 minutos. Após 30 minutos, o lisado de

proteína é produzido a partir dos fusossomas ou células, e os níveis de fosfo-AKT são medidos por western blotting em amostras estimuladas por insulina e controle não estimuladas.

[001307] A absorção de glicose em resposta à insulina ou solução de controle negativo é medida, conforme explicado na seção de absorção de glicose com o uso de glicose identificada (2-NBDG). (S. Galic et al., *Molecular Cell Biology* 25(2): 819 a 829, 2005).

[001308] Em uma modalidade, os fusossomas aumentarão a fosforilação de AKT e a absorção de glicose em resposta à insulina em relação aos controles negativos em pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais.

EXEMPLO 51: MEDIÇÃO DA CAPACIDADE DE TRANSPORTAR GLICOSE ATRAVÉS DA MEMBRANA CELULAR

[001309] Esse Exemplo descreve a quantificação dos níveis de um 2-NBDG (2-(*N*-(7-Nitrobenz-2-oxa-1,3-diazol-4-il)Amino)-2-Desoxiglicose), um análogo de glicose fluorescente que pode ser usado para monitorar a absorção de glicose em células vivas e, assim, medir o transporte ativo através da bicamada lipídica. Em uma modalidade, esse ensaio pode ser usado para medir o nível de absorção de glicose e transporte ativo através da bicamada lipídica do fusossoma.

[001310] Uma composição de fusossoma é produzida por qualquer um dos métodos descritos nos Exemplos anteriores. Um número suficiente de fusossomas são, então, incubados em DMEM sem glicose, 20% de Soro Fetal Bovino e 1x Penicilina/Estreptomicina por 2 horas a 37 °C e 5% de CO_2. Após um período de privação de glicose de 2 horas, o meio é alterado de modo a incluir DMEM sem glicose, 20% de soro fetal bovino, 1x penicilina/estreptomicina e 2-NBDG 20 uM (ThermoFisher) e incubado por mais 2 horas a 37 °C e 5% de CO_2.

[001311] Fusossomas de controle negativo são tratados da mesma maneira, com exceção de que uma quantidade igual de DMSO é adi-

cionada no lugar de 2-NBDG.

[001312] Os fusossomas são, então, lavados três vezes com 1xPBS e suspensos novamente em um tampão apropriado e transferidos para uma placa de imagem de 96 poços. A fluorescência 2-NBDG é, então, medida em um fluorímetro com o uso de um cubo de luz GFP (filtro de excitação 469/35 e um filtro de emissão 525/39) para quantificar a quantidade de 2-NBDG que foi transportado através da membrana de fusossoma e acumulado no fusossoma no período de carregamento de 1 hora.

[001313] Em uma modalidade, fluorescência 2-NBDG será mais alta no fusossoma com tratamento com 2-NBDG em comparação ao controle negativo (DMSO). A medida de fluorescência com um filtro de emissão 525/39 se correlacionará com o número de moléculas 2-NBDG presentes.

EXEMPLO 52: MEDIÇÃO DA ATIVIDADE ESTERASE NO CITOSOL

[001314] Esse exemplo descreve a quantificação da atividade esterase, como um substituto para a atividade metabólica, em fusossomas. A atividade da esterase citosólica em fusossomas é determinada por avaliação quantitativa da marcação com calceína-AM (Bratosin et al., Cytometry 66 (1): 78 a 84, 2005).

[001315] O corante permeável à membrana, calceína-AM (Molecular Probes, Eugene OR EUA), é preparado como uma solução-mãe de 10 mM em dimetilsulfóxido e como uma solução de trabalho de 100 mM em tampão PBS, pH 7,4. Os fusossomas produzidos por qualquer um dos métodos descritos nos Exemplos anteriores ou células de fibroblastos embrionários de camundongos parentais de controle positivo são suspensos em tampão PBS e incubados durante 30 minutos com solução de trabalho de calceína-AM (concentração final em calceína-AM: 5 mM) a 37 °C no escuro e, em seguida, diluída em tampão PBS para análise citométrica de fluxo imediata de retenção de fluores-

cência de calceína.

[001316] Fusossomas e células de fibroblastos embrionários de camundongo parentais de controle são permeabilizadas experimentalmente como um controle negativo para atividade esterase zero com saponina, conforme descrito em (Jacob et al., *Cytometry* 12(6): 550 a 558, 1991). Os fusossomas e células são incubados durante 15 minutos em solução de saponina a 1% em tampão PBS, pH 7,4, que contém azida de sódio a 0,05%. Devido à natureza reversível da permeabilização da membrana plasmática, a saponina é incluída em todos os tampões usados para as etapas de marcação e lavagem subsequentes. Após a permeabilização da saponina, fusossomas e células são suspensos em tampão PBS que contém 0,1% de saponina e 0,05% de azida de sódio e incubados (37 °C no escuro durante 45 minutos) com calceína-AM a uma concentração final de 5 mM, lavados três vezes com o mesmo tampão PBS que contém 0,1% de saponina e 0,05% de azida de sódio, e analisado por citometria de fluxo. As análises de citometria de fluxo são realizadas em um citômetro FACS (Becton Dickinson, San José, CA, EUA) com excitação com laser de argônio 488 nm, e a emissão é coletada a 530 +/- 30 nm. O software FACS é usado para aquisição e análise. Os canais de dispersão de luz são definidos em ganhos lineares e os canais de fluorescência são definidos em uma escala logarítmica, com um mínimo de 10.000 células analisadas em cada condição. As atividades de esterase relativa são calculadas com base na intensidade da calceína-AM em cada amostra. Todos os eventos são capturados nos canais de dispersão direta e lateral (alternativamente, um portão pode ser aplicado para selecionar apenas a população fusossoma). O valor de intensidade de fluorescência (FI) para os fusossomas é determinado subtraindo-se o valor FI da respectiva amostra tratada com saponina de controle negativo. A atividade de esterase normalizada para as amostras de fusossomas é

normalizada para as respectivas amostras de células de controle positivo, a fim de gerar medições quantitativas para atividades de esterase citosólica.

[001317] Em uma modalidade, uma preparação de fusossoma terá a atividade de esterase dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais em comparação à célula de controle positivo.

[001318] Consultar, também, Bratosin D, Mitrofan L, Palii C, Estaquier J, Montreuil J. Novel fluorescence assay using calcein-AM for the determination of human erythrocyte viability and aging. Cytometry A. Julho de 2005;66(1):78 a 84; e Jacob BC, Favre M, Bensa JC. Permeabilização de células membranares com saponina e análise multiparamétrica por citometria de fluxo. Cytometry 1991;12:550 a 558.

EXEMPLO 53: MEDIÇÃO DA ATIVIDADE DA ACETILCOLINESTERASE EM FUSOSSOMAS

[001319] A atividade da acetilcolinesterase é medida com o uso de um kit (MAK119, SIGMA) que segue um procedimento descrito anteriormente (Ellman, et al., Biochem. Pharmacol. 7, 88, 1961) e seguindo as recomendações do fabricante.

[001320] Em suma, os fusossomas são suspensos em acetiltiocolina 1,25 mM em PBS, pH 8, misturado com ácido 5,5-ditio-bis (ácido 2-nitrobenzoico) 0,1 mM em PBS, pH 7. A incubação é realizada em temperatura ambiente, porém os fusossomas e a solução de substrato são pré-aquecidos a 37 °C durante 10 minutos antes de iniciar as leituras de densidade óptica.

[001321] As alterações na absorção são monitoradas a 450 nm durante 10 minutos com um espectrofotômetro de leitor de placas (ELX808, BIO-TEK Instruments, Winooski, VT, EUA). Separadamente, uma amostra é usada para determinar o teor de proteína dos fusossomas via ensaio de ácido bicinconínico para normalização. Com uso

desse ensaio, os fusossomas são determinados como tendo <100 unidades de atividade AChE/µg de proteína.

[001322] Em uma modalidade, as unidades de atividade de AChE/µg de valores de proteína serão inferiores a 0,001, 0,01, 0,1, 1, 10, 100 ou 1.000.

EXEMPLO 54: MEDIÇÃO DO NÍVEL DE ATIVIDADE METABÓLICA

[001323] Esse exemplo descreve a quantificação da medição da atividade da citrato sintase em fusossomas.

[001324] A citrato sintase é uma enzima do ciclo do ácido tricarboxílico (TCA) que catalisa a reação entre o oxaloacetato (OAA) e o acetil-CoA para gerar citrato. Após a hidrólise da acetil-CoA, ocorre a liberação de CoA com um grupo tiol (CoA-SH). O grupo tiol reage com um reagente químico, 5,5-Ditiobis (ácido 2-nitrobenzoico) (DTNB), para formar o ácido 5-tio-2-nitrobenzoico (TNB), que é um produto amarelo que pode ser medido por espectrofotometria a 412 nm (Green 2008). Kits comercialmente disponíveis, tais como o kit de ensaio de atividade de citrato sintase humano Abcam (produto n° ab119692) fornecem todos os reagentes necessários para realizar essa medição.

[001325] O ensaio é realizado de acordo com as recomendações do fabricante. Os lisados de amostra de fusossoma são preparados coletando-se os fusossomas, conforme produzidos por qualquer um dos métodos descritos nos exemplos anteriores e solubilizando-os em tampão de extração (Abcam) por 20 minutos em gelo. Os sobrenadantes são coletados após a centrifugação, e o teor de proteína é avaliado por ensaio de ácido bicinconínico (BCA, ThermoFisher Scientific) e a preparação permanece em gelo até o seguinte protocolo de quantificação ser iniciado.

[001326] Em suma, as amostras de lisado de fusossoma são diluídas em tampão de incubação 1X (Abcam) nos poços da microplaca fornecidos, em que um conjunto de poços recebe apenas tampão de incu-

bação 1X. A placa é vedada e incubada durante 4 horas à temperatura ambiente com agitação a 300 rpm. O tampão é, então, aspirado dos poços e é adicionado tampão de lavagem 1X. Essa etapa de lavagem é repetida mais uma vez. Em seguida, a solução 1X Activity é adicionada a cada poço e a placa é analisada em um leitor de microplaca medindo-se a absorbância a 412 nm a cada 20 segundos durante 30 minutos, com agitação entre as leituras.

[001327] Os valores de fundo (poços com apenas 1X tampão de incubação) são subtraídos de todos os poços, e a atividade da citrato sintase é expressa como a mudança na absorbância por minuto por µg de amostra de lisado de fusossoma carregada (ΔmOD @ 412 nm/min/ug de proteína). Apenas a parte linear de 100 a 400 segundos da medição cinética é usada para calcular a atividade.

[001328] Em uma modalidade, uma preparação de fusossoma terá dentro de 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais atividade da sintase em comparação à célula de controle.

[001329] Consultar, por exemplo, Green HJ et al. Metabolic, enzymatic, and transporter response in human muscle during three consecutive days of exercise and recovery. Am J Physiol Regul Integr Comp Physiol 295: R1238 a R1250, 2008.

EXEMPLO 55: MEDIÇÃO DOS NÍVEIS DE RESPIRAÇÃO

[001330] Esse exemplo descreve a quantificação da medição do nível de respiração em fusossomas. O nível de respiração nas células pode ser uma medida do consumo de oxigênio, que estimula o metabolismo. A respiração fusossoma é medida para as taxas de consumo de oxigênio por um analisador de fluxo extracelular Seahorse (Agilent) (Zhang 2012).

[001331] Os fusossomas produzidos por qualquer um dos métodos descritos nos Exemplos anteriores ou células são semeados em uma

microplaca Seahorse de 96 poços (Agilent). A microplaca é centrifugada brevemente para peletizar os fusossomas e as células no fundo dos poços. Os ensaios de consumo de oxigênio são iniciados removendo-se o meio de crescimento, substituindo por um meio mínimo DMEM com baixo tamponamento que contém glicose 25 mM e glutamina 2 mM (Agilent) e incubando-se a microplaca a 37 °C por 60 minutos para permitir o equilíbrio de temperatura e pH.

[001332] A microplaca é, então, testada em um analisador de fluxo extracelular (Agilent) que mede as mudanças no oxigênio extracelular e no pH da mídia imediatamente ao redor dos fusossomas e células aderentes. Após obter o consumo de oxigênio em estado estacionário (taxa de respiração basal) e as taxas de acidificação extracelular, oligomicina (5 µM), que inibe a ATP sintase, e ionóforo de prótons FCCP (cianeto de carbonila 4-(trifluorometoxi)fenil-hidrazona; 2 µM), que desacopla mitocôndrias, são adicionados a cada poço da microplaca para obter os valores das taxas máximas de consumo de oxigênio.

[001333] Finalmente, antimicina A a 5 µM (inibidor do complexo III da mitocôndria) é adicionada para confirmar que as alterações na respiração se devem principalmente à respiração mitocondrial. A taxa mínima de consumo de oxigênio após a adição de antimicina A é subtraída de todas as medições de consumo de oxigênio para remover o componente de respiração não mitocondrial. As amostras de células que não respondem adequadamente à oligomicina (pelo menos uma redução de 25% na taxa de consumo de oxigênio do valor basal) ou FCCP (pelo menos um aumento de 50% na taxa de consumo de oxigênio após a oligomicina) são excluídas da análise. O nível de respiração dos fusossomas é, então, medido como pmol O2/min/1e4 fusossomas.

[001334] Esse nível de respiração é, então, normalizado para o respectivo nível de respiração celular. Em uma modalidade, os fusossomas terão pelo menos um nível de respiração de 1%, 2%, 3%, 4%,

5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou superior em comparação às respectivas amostras de células.

[001335] Consultar, por exemplo, Zhang J, Nuebel E, Wisidagama DRR, et al. Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. Nature protocols. 2012;7(6):10.1038/nprot.2012.048. doi:10.1038/nprot.2012.048.

EXEMPLO 56: MEDIÇÃO DOS NÍVEIS DE FOSFATIDILSERINA DE FUSOSSOMAS

[001336] Esse exemplo descreve a quantificação do nível de ligação da anexina-V à superfície dos fusossomas.

[001337] As células que estão morrendo podem exibir fosfatidilserina na superfície celular, que é um marcador de apoptose na via de morte celular programada. A anexina-V se liga à fosfatidilserina e, portanto, a ligação da anexina-V é uma representação para a viabilidade nas células.

[001338] Os fusossomas foram produzidos, conforme descrito no presente documento. Para a detecção de sinais de apoptose, os fusossomas ou células de controle positivo foram marcados com 5% de anexina V fluor 594 (A13203, Thermo Fisher, Waltham, MA). Cada grupo (detalhado na tabela abaixo) incluiu um braço experimental que foi tratado com um indutor de apoptose, menadiona. Menadiona foi adicionada em 100 μM DE menadiona durante 4 horas. Todas as amostras foram executadas em citômetro de fluxo (Thermo Fisher, Waltham, MA) e a intensidade de fluorescência foi medida com o laser YL1 em um comprimento de onda de 561 nm e um filtro de emissão de 585/16 nm. A presença de fosfatidilserina extracelular foi quantificada pela comparação da intensidade de fluorescência da anexina V em todos os grupos.

[001339] Os fusossomas não marcados de controle negativo não foram positivos para a marcação com anexina V.

[001340] Em uma modalidade, os fusossomas tiveram capacidade para regular positivamente a exibição de fosfatidilserina na superfície celular em resposta à menadiona, indicando que os fusossomas não estimulados por menadiona não estão passando por apoptose. Em uma modalidade, as células de controle positivo que foram estim μladas com menadiona demonstraram níveis mais elevados de marcação com anexina V do que fusossomas não estim μlados com menadiona.

TABELA 7: PARÂMETRO DE MARCAÇÃO DE ANEXINA V

Braço Experimental	Intensidade média de fluorescência do sinal de anexina V (e desvio padrão)
Fusossomas sem marcação (controle negativo)	941 (937)
Fusossomas marcados	11.257 (15.826)
Fusossomas marcados + menadiona	18.733 (17.146)
Macrófagos marcados + menadiona (controle positivo)	14.301 (18.142)

EXEMPLO 57: MEDIÇÃO DOS NÍVEIS DE SINALIZAÇÃO DE JUSTACRINA

[001341] Esse exemplo descreve a quantificação da sinalização de justacrina em fusossomas.

[001342] As células podem formar sinalização dependente do contato celular por meio da sinalização de justacrina. Em uma modalidade, a presença de sinalização de justacrina em fusossomas demonstrará que os fusossomas podem estimular, reprimir e geralmente se comunicar com as células imediatamente mais próximas.

[001343] Os fusossomas produzidos por qualquer um dos métodos descritos nos Exemplos anteriores a partir de células estromais da medula óssea de mamíferos (BMSCs) com inativação nuclear parcial ou completa desencadeiam a secreção de IL-6 via sinalização de justacrina em macrófagos. Macrófagos primários e BMSCs são cocultivados. Os macrófagos derivados da medula óssea são semeados primei-

ramente em placas de 6 poços e incubados durante 24 horas, em seguida, os fusossomas primários de camundongo BMSC ou células BMSC (células parentais de controle positivo) são colocados nos macrófagos em um meio DMEM com 10% de FBS. O sobrenadante é coletado em diferentes instantes do tempo (2, 4, 6, 24 horas) e analisado quanto à secreção de IL-6 por ensaio ELISA. (Chang J. et al., 2015).

[001344] Em uma modalidade, o nível de sinalização de justacrina induzida por fusossomas BMSC é medido por um aumento nos níveis de IL-6 secretada por macrófagos no meio. Em uma modalidade, o nível de sinalização de justacrina será de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais do que os níveis induzidos pelas células estromais da medula óssea de controle positivo (BMSCs).

EXEMPLO 58: MEDIÇÃO DOS NÍVEIS DE SINALIZAÇÃO DE PARÁCRINA

[001345] Esse exemplo descreve a quantificação da sinalização de parácrina em fusossomas.

[001346] As células podem se comunicar com outras células no microambiente sítio por meio de sinalização de parácrina. Em uma modalidade, os fusossomas terão capacidade para sinalização de parácrina para, por exemplo, se comunicar com as células em seu ambiente local. Em uma modalidade, a capacidade que os fusossomas têm para desencadear a sinalização de Ca^{2+} em células endoteliais por meio de secreção derivada de parácrina com o seguinte protocolo medirá a sinalização de Ca^{2+} por meio do indicador de cálcio, fluo-4 AM.

[001347] Para preparar a prato experimental, as células endoteliais microvasculares pulmonares de murganho (MPMVECs) são semeadas em um prato confocal com fundo de vidro de 25 mm revestida com gelatina a 0,2% (80% de confluência). As MPMVECs são incubadas em temperatura ambiente durante 30 minutos em ECM que contém 2% de

BSA e 0,003% de ácido plurônico com uma concentração final de fluo-4 AM a 5 μM (Invitrogen) para permitir o carregamento de fluo-4 AM. Após o carregamento, as MPMVECs são lavadas com solução de imagem experimental (ECM que contém 0,25% de BSA) que contém sulfinpirazona para minimizar a perda de corante. Após o carregamento de fluo-4, 500 μl da solução de imagem experimental pré-aquecida são adicionados à placa, e a placa é fotografada por um sistema de imagem confocal Zeiss.

[001348] Em um tubo separado, os macrófagos de murganho isolados recentemente são tratados com 1 μg/ml de LPS em meio de cultura (DMEM + FBS a 10%) ou não tratados com LPS (controle negativo). Após a estimulação, os fusossomas são gerados a partir de macrófagos por qualquer um dentre os métodos descritos nos Exemplos anteriores.

[001349] Os fusossomas ou macrófagos parentais (controle positivo) são, então, identificados com vermelho rastreador de células, CMTPX (Invitrogen), em ECM que contém BSA a 2% e ácido plurônico 0,003%. Os fusossomas e macrófagos são, então, lavados e suspensos novamente em solução de imagem experimental. Os fusossomas e macrófagos identificados são adicionados às MPMVECs carregadas com fluo-4 AM na placa confocal.

[001350] O sinal de fluorescência verde e vermelho é registrado a cada 3 s durante 10 a 20 minutos com o uso do sistema de imagem confocal Zeiss com fonte de laser de íon argônio com excitação a 488 e 561 nm para fluo-4 AM e fluorescência vermelha do rastreador de células, respectivamente. As mudanças de intensidade de fluorescência Fluo-4 são analisadas com o uso do software de imageamento (Mallilankaraman, K. et al., *J Vis Exp.* (58): 3.511, 2.011). O nível de intensidade de Fluo-4 medido em fusossoma de controle negativo e grupos de células é subtraído de fusossoma e grupos de células esti-

mulados por LPS.

[001351] Em uma modalidade, fusossomas, por exemplo, fusossomas ativados, induzirão um aumento na intensidade de fluorescência Fluo-4 que é de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou superior aos grupos de células de controle positivo.

EXEMPLO 59: MEDIÇÃO DA CAPACIDADE PARA POLIMERIZAR ACTINA PARA MOBILIDADE

[001352] Esse exemplo descreve a quantificação de componentes do citoesqueleto, tal como actina, em fusossomas. Em uma modalidade, os fusossomas compreendem componentes do citoesqueleto, tais como actina, e têm capacidade para polimerização de actina.

[001353] As células usam actina, que é um componente do citoesqueleto, para a motilidade e outros processos citoplasmáticos. O citoesqueleto é essencial para criar forças conduzidas pela motilidade e coordenar o processo de movimento.

[001354] As células C2C12 foram enucleadas conforme descrito no presente documento. Os fusossomas obtidos das camadas de Ficoll a 12,5% e 15% foram agrupados e identificados como “Leves”, ao passo que fusossomas das camadas de 16 a 17% foram agrupadas e identificadas como “Médias”. Fusossomas ou células (células C2C12 parentais, controle positivo) foram suspensos novamente em DMEM + Glutamax + 10% de Soro Fetal Bovino (FBS), plaqueados em placas de fixação ultrabaixa de 24 poços (n° 3473, Corning Inc, Corning, NY) e incubados em 37 °C + 5% de CO_2. As amostras foram retiradas periodicamente (5,25 horas, 8,75 horas, 26,5 horas) e marcadas com Rodamina faloidina a 165 μM (controle negativo não foi marcado) e medida em um citômetro de fluxo (n° A24858, Thermo Fisher, Waltham, MA) com um laser FC YL1 (561 nm com filtro 585/16) para medir o conteúdo do citoesqueleto de F-actina. A intensidade de fluorescência

da rodamina faloidina em fusossomas foi medida junto com fusossomas não marcados e células C2C12 parentais marcadas.

[001355] A intensidade da fluorescência dos fusossomas foi maior (A Figura 1) que o controle negativo em todos os instantes do tempo e os fusossomas tiveram capacidade para polimerizar actina a uma taxa semelhante às células C2C12 parentais.

[001356] Componentes adicionais do citoesqueleto, tais como os listados na tabela abaixo, são medidos por meio de sistemas ELISA disponíveis no mercado (Cell Signaling Technology e MyBioSource), de acordo com as instruções do fabricante.

TABELA 8: COMPONENTES DO CITOESQUELETO

Proteína citoesquelética medida	Tipo de Kit Comercial	ID do kit
Actina	Kit ELISA Sanduíche de B-Actina Total de Varredura de Trajetória	Sinalização Celular, 7880
Arp2/3	Kit ELISA de Subunidade de Complexo Proteína 2/3 Relacionada à Actina (APRC2) Humana	MyBioSource, MBS7224740
Formina	**Proteína de Formação de Formina 1 (FNBP1), Kit ELISA**	MyBioSource, MBS9308864
Coronina	Kit ELISA de Coronina Humana 1	MyBioSource, MBS073640
Distrofina	Kit ELISA de Distrofina Humana	MyBioSource MBS722223
Queratina	Kit ELISA de Queratina Humana 5	MyBioSource, MBS081200
Miosina	Kit ELISA de Miosina Humana IG (MYO1G)	MyBioSource, MBS9312965
Tubulina	Kit ELISA de Beta Tubulina Humana 3	MyBioSource, MBS097321

[001357] Em seguida, 100 ul de lisado diluído apropriadamente são adicionados ao poço apropriado das tiras de micropoços. Os micropoços são vedados com fita e incubados durante 2 horas a 37 °C. Após a

incubação, a fita de vedação é removida, e o teor é descartado. Cada micropoço é lavado quatro vezes com 200 ul de Tampão de Lavagem 1X. Após cada lavagem individual, as placas são colocadas em um pano absorvente para que a solução de lavagem residual seja removida de cada poço. No entanto, os poços não estão completamente secos em nenhum momento durante o experimento.

[001358] Em seguida, 100 ul do Anticorpo de Detecção reconstituído (verde) são adicionados a cada poço individual, exceto para os poços de controle negativo. Em seguida, os poços são vedados e incubados durante 1 hora a 37 °C. O procedimento de lavagem é repetido após a incubação ser concluída. 100 µl de anticorpo secundário ligado a HRP reconstituído (vermelho) são adicionados a cada um dos poços. Os poços são vedados com fita e incubados durante 30 minutos a 37 °C. Em seguida, a fita de vedação é removida, e o procedimento de lavagem é repetido. 100 ul de substrato TMB são, então, adicionados a cada poço. Os poços são vedados com fita adesiva e, em seguida, incubados durante 10 minutos a 37 °C. Após essa incubação final ter sido concluída, 100 µl de solução de terminação são adicionados a cada um dos poços, e a placa é agitada suavemente por vários segundos.

[001359] A análise espectrofotométrica do ensaio é realizada 30 minutos após a adição da solução de terminação. A parte inferior dos poços é limpa com um tecido sem fiapos e, em seguida, a absorbância é lida a 450 nm. Em uma modalidade, as amostras de fusossoma que foram marcadas com o anticorpo de detecção absorverão a 450 nm mais luz que as amostras de fusossoma de controle negativo e absorverão menos luz que as amostras de células que foram marcadas com o anticorpo de detecção.

EXEMPLO 60: MEDIÇÃO DO POTENCIAL MÉDIO DE MEMBRANA

[001360] Esse exemplo descreve a quantificação do potencial de

membrana mitocondrial de fusossomas. Em uma modalidade, os fusossomas que compreendem uma membrana mitocondrial manterão o potencial de membrana mitocondrial.

[001361] A atividade metabólica mitocondrial pode ser medida pelo potencial de membrana mitocondrial. O potencial de membrana da preparação de fusossoma é quantificado com o uso de um corante disponível comercialmente, TMRE, para avaliar o potencial de membrana mitocondrial (TMRE: tetrametil rodamina, éster etílico, perclorato, Abcam, n° de Catálogo T669).

[001362] Os fusossomas são gerados por qualquer um dos métodos descritos nos Exemplos anteriores. Fusossomas ou células parentais são diluídos em meio de crescimento (DMEM livre de vermelho de fenol com 10% de soro fetal bovino) em 6 alíquotas (triplicatas não tratadas e tratadas com FCCP). Uma alíquota das amostras é incubada com FCCP, um desacoplador que elimina o potencial da membrana mitocondrial e evita a marcação de TMRE. Para amostras tratadas com FCCP, 2 µM de FCCP é adicionado às amostras e incubado por 5 minutos antes da análise. Os fusossomas e células parentais são, então, marcados com TMRE a 30 nM. Para cada amostra, uma amostra não marcada (sem TMRE) também é preparada paralelamente. As amostras são incubadas a 37 °C durante 30 minutos. As amostras são, então, analisadas em um citômetro de fluxo com laser de argônio a 488 nm, e a excitação e a emissão são coletadas a 530 +/- 30nm.

[001363] Os valores do potencial da membrana (em milivolts, mV) são calculados com base na intensidade de TMRE. Todos os eventos são capturados nos canais de dispersão frontal e lateral (alternativamente, uma porta pode ser aplicada para excluir pequenos detritos). O valor de intensidade de fluorescência (FI) para as amostras não tratadas e tratadas com FCCP são normalizados subtraindo-se a média geométrica da intensidade de fluorescência da amostra não marcada

da média geométrica da amostra não tratada e tratada com FCCP. O estado do potencial de membrana para cada preparação é calculado usando os valores de intensidade fluorescente normalizados com uma equação de Nernst modificada (veja abaixo) que pode ser usada para determinar o potencial de membrana mitocondrial dos fusossomas ou células com base na fluorescência TMRE (como TMRE se acumula nas mitocôndrias de maneira nernstiana).

[001364] O fusossoma ou potencial de membrana celular é calculado com a seguinte fórmula: (mV) = -61,5 * log(normalizado-não tratado por FI/normalizado-não tratado por FIFCCP). Em uma modalidade, com o uso desse ensaio em preparações de fusossoma de células de mioblasto de camundongo C2C12, o estado potencial de membrana da preparação de fusossoma estará dentro de aproximadamente 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% ou mais do que as células parentais. Em uma modalidade, a faixa de potencial de membrana é de aproximadamente -20 a -150 mV.

EXEMPLO 61: MEDIÇÃO DA MEIA-VIDA DE PERSISTÊNCIA EM UM INDIVÍDUO

[001365] Esse exemplo descreve a medição da meia-vida de fusossoma.

[001366] Os fusossomas são derivados de células que expressam a luciferase de Gaussia produzida por qualquer um dos métodos descritos nos Exemplos anteriores, e são feitas diluições puras de 1:2, 1:5 e 1:10 em solução tamponada. Uma solução tamponada sem fusossomas é usada como um controle negativo.

[001367] Cada dose é administrada a três camundongos C57BL/6J machos com oito semanas de idade (Jackson Laboratories) por via intravenosa. O sangue é coletado da veia retro-orbital em 1, 2, 3, 4, 5, 6, 12, 24, 48 e 72 horas após a administração intravenosa dos fusossomas. Os animais são sacrificados no final do experimento com ina-

lação de CO_2.

[001368] O sangue é centrifugado durante 20 minutos em temperatura ambiente. As amostras de soro são imediatamente congeladas a -80 °C até a bioanálise. Em seguida, cada amostra de sangue é usada para realizar um ensaio de atividade de Gaussia luciferase após misturar as amostras com substrato de Gaussia luciferase (Nanolight, Pinetop, AZ). Resumidamente, a colenterazina, uma molécula de emissão de luciferina ou de luz, é misturada com tampão de ensaio instantâneo, e a mistura é pipetada em poços que contêm amostras de sangue em uma placa de 96 poços. Os poços de controle negativo sem sangue contêm tampão de ensaio para determinar o sinal de fundo da Gaussia luciferase.

[001369] Além disso, uma curva padrão de Gaussia luciferase purificada por controle positivo (Athena Enzyme Systems, n° de catálogo 0308) é preparada a fim de converter o sinal de luminescência em moléculas de secreção de Gaussia luciferase por hora. A placa é testada para luminescência, com o uso da integração de 500 ms. O sinal da Gaussia luciferase de fundo é subtraído de todas as amostras e, em seguida, uma curva linear de melhor ajuste é calculada para a curva padrão da Gaussia luciferase. Caso as leituras da amostra não se encaixarem na curva padrão, elas serão diluídas de forma adequada e testadas novamente. O sinal de luciferase de amostras tomadas em 1, 2, 3, 4, 5, 6, 12, 24, 48 e 72 horas é interpolado para a curva padrão. A constante de taxa de eliminação k_e (h^{-1}) é calculada com o uso da equação a seguir de um modelo de um compartimento: $C(t) = C_0 \times e^{-kext}$, no qual $C(t)$ (ng/ml) é a concentração de fusossomas no tempo t (h) e C_0 é a concentração de fusossomas no tempo = 0 (ng/ml). A meia-vida de eliminação $t_{1/2,e}$ (h) é calculada como $\ln(2)/k_e$.

[001370] Em uma modalidade, os fusossomas terão uma meia-vida de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%,

60%, 70%, 80%, 90%, 100% ou mais do que as células de controle negativo.

EXEMPLO 62: ENTREGA DE FUSOSSOMAS POR VIA NÃO ENDOCÍTICA

[001371] Esse exemplo descreve a quantificação da entrega de fusossoma da Cre a uma célula receptora por meio de uma via não endocítica.

[001372] Em uma modalidade, os fusossomas entregarão agentes por meio de uma via não endocítica mediada por fusossomas. Sem se ater a nenhuma teoria, a entrega de um agente, por exemplo, Cre, que é transportado dentro do lúmen dos fusossomas, diretamente para o citosol das células receptoras, sem qualquer necessidade de absorção mediada por endocitose dos fusossomas, ocorrerá através uma distribuição de via não endocítica mediada por fusossoma.

[001373] Nesse exemplo, o fusossoma compreende uma célula HEK293T que expressa a proteína H e F do vírus Sendai em sua membrana plasmática (Tanaka et al., 2015, Gene Therapy, 22(de outubro de 2014), 1 a 8. https://doi.org/10.1038/gt.2014.123). Além disso, o fusossoma expressa a proteína fluorescente mTagBFP2 e a Cre recombinase. A célula-alvo é uma célula RPMI8226 que expressa de forma estável o cassete “LoxP-GFP-stop-LoxP-RFP” sob um promotor CMV, que após recombinação por Cre muda de GFP para expressão de RFP, indicando fusão e Cre, como um marcador, entrega.

[001374] Os fusossomas produzidos pelos métodos descritos no presente documento são testados para entrega de Cre por meio de uma via não endocítica conforme a seguir. As células receptoras são colocadas em uma placa de 96 poços de fundo transparente preto. Em seguida, 24 horas após as células receptoras terem sido colocadas em placas, os fusossomas que expressam a proteína Cre recombinase e que possuem a proteína fusogênica particular são aplicados às células

receptoras em meio DMEM. Para determinar o nível de entrega de Cre por meio de uma via não endocítica, um grupo paralelo de células receptoras que recebe fusossomas é tratado com um inibidor de acidificação endossômica, cloroquina (30 μg/ml). A dose de fusossomas é correlacionada ao número de células receptoras colocadas em placa no poço. Após a aplicação dos fusossomas, a placa de células é centrifugada a 400 g durante 5 minutos para ajudar a iniciar o contato entre os fusossomas e as células receptoras. As células são, então, incubadas durante 16 horas, e a entrega do agente, Cre, é avaliada por meio de imagens.

[001375] As células são fotografadas para identificar positivamente células positivas para RFP versus células positivas para GFP no campo ou poço. Nesse exemplo, placas de células são imageadas com o uso de um microscópio de fluorescência automatizado. A população total de células em um determinado poço é determinada marcando-se primeiramente as células com Hoechst 33342 em meio DMEM durante 10 minutos. O Hoechst 33342 marca os núcleos de células intercalando-se em DNA e, portanto, é usado para identificar células individuais. Após a marcação, o meio Hoechst é substituído por meio DMEM regular.

[001376] A imagem da Hoechst é obtida com o uso de LED de 405 nm e cubo de filtro DAPI. A imagem da GFP é obtida com o uso do LED de 465 nm e cubo de filtro GFP, ao passo que a imagem de RFP é obtida com o uso do LED de 523 nm e do cubo de filtro de RFP. As imagens dos diferentes grupos de células são adquiridas estabelecendo-se primeiramente a intensidade do LED e os tempos de integração em um poço de controle positivo; isto é, células receptoras tratadas com adenovírus que codificam para Cre recombinase em vez de fusossomas.

[001377] As configurações de aquisição são definidas de forma que

as intensidades de RFP e GFP estejam nos valores máximos de intensidade de pixel, porém não saturadas. Os poços de interesse são, então, fotografados com o uso das configurações estabelecidas.

[001378] A análise de poços positivos para GFP e RFP é realizada com software dotado do microscópio de fluorescência ou outro software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, EUA, 1997 a 2007). As imagens são pré-processadas com o uso de um algoritmo de subtração de fundo de bola rolante com uma largura de 60 µm. A máscara celular total é definida nas células Hoechst-positivas. As células com intensidade Hoechst significativamente acima das intensidades de fundo são usadas para definir um limite e áreas muito pequenas ou grandes para serem células Hoechst-positivas são excluídas.

[001379] Dentro da máscara de célula total, células GFP e RFP-positivas são identificadas ajustando-se novamente um limite para células significativamente acima do fundo e estendendo-se as máscaras Hoechst (núcleos) para toda a área celular para incluir toda a fluorescência celular de GFP e RFP.

[001380] O número de células positivas para RFP identificadas em poços de controle que contêm células receptoras é usado para subtrair do número de células positivas para RFP nos poços que contêm fusossomas (a fim de subtrair para recombinação Loxp não específica). O número de células positivas para RFP (células receptoras que receberam Cre) é, então, dividido pela soma de células positivas para GFP (células receptoras que não receberam Cre) e células positivas para RFP para quantificar a fração do fusossoma Cre entregue à população de células receptoras. O nível é normalizado para a dose determinada de fusossomas aplicada às células receptoras. A fim de calcular o valor do fusossoma Cre entregue por uma via não endocítica, o nível de entrega de fusossoma Cre na presença de cloroquina (FusL+CQ) é

determinado, assim como o nível de entrega de fusossoma Cre na ausência de cloroquina (FusL-CQ). Para determinar o valor normalizado de fusossoma Cre entregue por uma via não endocítica, a seguinte equação é usada: [(FusL-CQ)-(FusL+CQ)]/(FusL-CQ).

[001381] Em uma modalidade, o nível médio de fusossoma Cre entregue por meio de uma via não endocítica para um determinado fusossoma estará na faixa de 0,1 a 0,95, ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais que as células receptoras tratadas com cloroquina.

EXEMPLO 63: ENTREGA DE FUSOSSOMAS POR VIA ENDOCÍTICA

[001382] Esse exemplo descreve a entrega de fusossoma de Cre a uma célula receptora por meio de uma via endocítica.

[001383] Em uma modalidade, os fusossomas entregarão agentes por meio de uma via endocítica mediada por fusossomas. Sem se ater a nenhuma teoria, a entrega de um agente, por exemplo, uma carga, transportada no lúmen dos fusossomas, para as células receptoras com a rota de absorção sendo dependente de endocitose ocorrerá através de uma entrega de via endocítica mediada por fusossoma.

[001384] Nesse exemplo, o fusossoma compreende microvesículas que foram produzidas pela extrusão de uma célula HEK293T que expressa uma proteína fusogênica em sua membrana plasmática através de um filtro de 2 μm (Lin et al., 2016, Biomedical Microdevices, 18(3). doi.org/10.1007/s10544-016-0066-y)(Riedel, Kondor-Koch, & Garoff, 1984, The EMBO Journal, 3(7), 1.477 a 1483. Obtido em www.ncbi. nlm. nih. gov/ pubmed/6086326). Além disso, o fusossoma expressa a proteína fluorescente mTagBFP2 e a Cre recombinase. A célula-alvo é uma célula PC3 que expressa de forma estável o cassete “LoxP-GFP-stop-LoxP-RFP” sob um promotor CMV, que após recombinação por Cre muda de GFP para expressão de RFP, indicando fusão e Cre,

como um marcador, entrega.

[001385] Os fusossomas produzidos pelos métodos descritos no presente documento são testados para entrega de Cre por meio de uma via endocítica conforme a seguir. As células receptoras são colocadas em uma placa de cultura de células de múltiplos poços compatível com o sistema de imageamento a ser usado (nesse exemplo, as células são colocadas em uma placa de 96 poços de fundo transparente preto). Em seguida, 24 horas após as células receptoras terem sido colocadas em placas, os fusossomas que expressam a proteína Cre recombinase e que possuem a proteína fusogênica particular são aplicados às células receptoras em meio DMEM. Para determinar o nível de entrega de Cre por meio de uma via endocítica, um grupo paralelo de células receptoras que recebem fusossomas é tratado com um inibidor de acidificação endossômica, cloroquina (30 μg/ml). A dose de fusossomas é correlacionada ao número de células receptoras colocadas em placa no poço. Após a aplicação dos fusossomas, a placa de células é centrifugada a 400 g durante 5 minutos para ajudar a iniciar o contato entre os fusossomas e as células receptoras. As células são, então, incubadas durante 16 horas, e a entrega do agente, Cre, é avaliada por meio de imagens.

[001386] As células são fotografadas para identificar positivamente células positivas para RFP versus células positivas para GFP no campo ou poço. Nesse exemplo, as placas de células são fotografadas com o uso de um microscópio fluorescente automatizado. A população total de células em um determinado poço é determinada marcando-se primeiramente as células com Hoechst 33342 em meio DMEM durante 10 minutos. O Hoechst 33342 marca os núcleos de células intercalando-se em DNA e, portanto, é usado para identificar células individuais. Após a marcação, a mídia Hoechst é substituída por mídia DMEM regular.

[001387] A imagem da Hoechst é obtida com o uso de LED de 405 nm e cubo de filtro DAPI. A imagem da GFP é obtida com o uso do LED de 465 nm e cubo de filtro GFP, ao passo que a imagem de RFP é obtida com o uso do LED de 523 nm e do cubo de filtro de RFP. As imagens dos diferentes grupos de células são adquiridas estabelecendo-se primeiramente a intensidade do LED e os tempos de integração em um poço de controle positivo; isto é, células receptoras tratadas com adenovírus que codificam para Cre recombinase em vez de fusossomas.

[001388] As configurações de aquisição são definidas de forma que as intensidades de RFP e GFP estejam nos valores máximos de intensidade de pixel, porém não saturadas. Os poços de interesse são, então, fotografados com o uso das configurações estabelecidas.

[001389] A análise de poços positivos para GFP e RFP é realizada com software dotado do microscópio fluorescente ou outro software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, EUA, 1997 a 2007). As imagens são pré-processadas com o uso de um algoritmo de subtração de fundo de bola rolante com uma largura de 60 μm. A máscara celular total é definida nas células Hoechst-positivas. As células com intensidade Hoechst significativamente acima das intensidades de fundo são submetidas a um limiar, e áreas muito pequenas ou grandes para serem células Hoechst-positivas são excluídas.

[001390] Dentro da máscara de célula total, as células GFP e RFP-positivas são identificadas novamente pelo limiar para células significativamente acima do fundo e estendendo-se as máscaras Hoechst (núcleos) para toda a área celular para incluir toda a fluorescência celular de GFP e RFP.

[001391] O número de células positivas para RFP identificadas em poços de controle que contêm células receptoras é usado para subtrair

do número de células positivas para RFP nos poços que contêm fusossomas (a fim de subtrair para recombinação Loxp não específica). O número de células positivas para RFP (células receptoras que receberam Cre) é, então, dividido pela soma de células positivas para GFP (células receptoras que não receberam Cre) e células positivas para RFP para quantificar a fração do fusossoma Cre entregue à população de células receptoras. O nível é normalizado para a dose determinada de fusossomas aplicada às células receptoras. A fim de calcular o valor do fusossoma Cre entregue por uma via endocítica, o nível de entrega de fusossoma Cre na presença de cloroquina (FusL+CQ) é determinado, assim como o nível de entrega de fusossoma Cre na ausência de cloroquina (FusL-CQ). Para determinar o valor normalizado de fusossoma Cre entregue através de uma via endocítica, a seguinte equação é usada: (FusL+CQ)/(FusL-CQ).

[001392] Em uma modalidade, o nível médio de fusossoma Cre entregue por meio de uma via endocítica para um determinado fusossoma estará na faixa de 0,01 a 0,6 ou pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do que as células receptoras tratadas com cloroquina.

EXEMPLO 64: ENTREGA DE FUSOSSOMAS POR MEIO DE UMA VIA MEDIADA POR DINAMINA, UMA VIA DE MACROPINOCITOSE OU UMA VIA MEDIADA POR ACTINA

[001393] Esse exemplo descreve a entrega de fusossoma de Cre a uma célula receptora por meio de uma via mediada por dinamina. Um fusossoma que compreende uma microvesícula pode ser produzido conforme descrito no exemplo anterior. Fusossomas são testados para a entrega de Cre por meio de uma mediada por dinamina de acordo com o exemplo anterior, com exceção de que um grupo de células receptoras que recebem fusossomas é tratado com um inibidor de dinamina, Dynasore (120 μM). Para calcular o valor de fusossoma Cre en-

tregue por meio de uma via mediada por dinamina, o nível de entrega de fusossoma Cre na presença de Dynasore (FusL+DS) é determinado, assim como o nível de entrega de fusossoma Cre na ausência de Dynasore (FusL-DS). O valor normalizado de fusossoma Cre entregue pode ser calculado conforme descrito no exemplo anterior.

[001394] Esse exemplo também descreve a entrega de Cre a uma célula receptora por meio de macropinocitose. Um fusossoma que compreende uma microvesíula pode ser produzido conforme descrito no exemplo anterior. Os fusossomas são testados quanto à entrega de Cre via macropinocitose de acordo com o exemplo anterior, com exceção de que um grupo de células receptoras que recebem fusossomas é tratado com um inibidor de macropinocitose, 5-(*N*-etil-*N*-isopropil) amilorida (EIPA) (25 μM). A fim de calcular o valor de fusossoma Cre entregue por macropinocitose, o nível de entrega de fusossoma Cre na presença de EIPA (FusL+EPIA) é determinado, assim como o nível de entrega de fusossoma Cre na ausência de EPIA (FusL-EIPA). O valor normalizado de fusossoma Cre entregue pode ser calculado conforme descrito no exemplo anterior.

[001395] Esse exemplo também descreve a entrega de fusossoma de Cre a uma célula receptora por meio de uma via mediada por actina. Um fusossoma que compreende uma microvesíula pode ser produzido conforme descrito no exemplo anterior. Fusossomas são testados para entrega de Cre por meio de macropinocitose de acordo com o exemplo anterior, com exceção de que um grupo de células receptoras que recebem fusossomas é tratado com um inibidor de polimerização de actina, Latrunculina B (6 μM). Para calcular o valor de fusossoma Cre entregue por uma via mediada por actina, o nível de entrega de fusossoma Cre na presença de Latrunculina B (FusL+LatB) é determinado, assim como o nível de entrega de fusossoma Cre na ausência de Latrunculina B (FusL-LatB). O valor normalizado de fusos-

soma Cre entregue pode ser calculado conforme descrito no exemplo anterior.

EXEMPLO 65: ENTREGA IN VIVO DE PROTEÍNA

[001396] Esse exemplo descreve a entrega de agentes terapêuticos ao olho por fusossomas.

[001397] Os fusossomas são derivados de células-tronco hematopoiéticas e progenitoras que usam qualquer um dos métodos descritos nos exemplos anteriores e são carregados com uma proteína que é deficiente em um knockout de camundongo.

[001398] Os fusossomas são injetados no olho direito por via subretiniana de um camundongo que é deficiente da proteína e o controle do veículo é injetado no olho esquerdo dos camundongos. Um subconjunto dos camundongos é submetido à eutanasia quando chegam aos 2 meses de idade.

[001399] A histologia e a marcação H&E do tecido retiniano colhido são conduzidas para contar o número de células resgatadas em cada retina dos camundongos (descrito em Sanges et al., The Journal of Clinical Investigation, 126(8): 3.104 a 3.116, 2016).

[001400] O nível da proteína injetada é medido em retinas colhidas de camundongos sacrificados aos 2 meses de idade por meio de um western blot com um anticorpo específico para a proteína PDE6B.

[001401] Em uma modalidade, os olhos esquerdos de camundongos, aos quais são administrados fusossomas, terão um número aumentado de núcleos presentes no nível nuclear externo da retina em comparação com os olhos direitos de camundongos, que são tratados com veículo. O aumento da proteína é sugestivo de complementação da proteína PBE6B mutante.

EXEMPLO 66: AVALIAÇÃO DA FORMAÇÃO DE TERATOMA APÓS ADMINISTRAÇÃO DE FUSOSSOMA

[001402] Esse exemplo descreve a ausência de formação de terato-

ma com um fusossoma. Em uma modalidade, um fusossoma não resultará na formação de teratoma quando administrado a um indivíduo.

[001403] Os fusossomas são produzidos por qualquer um dos métodos descritos em um exemplo anterior. Fusossomas, células tumorais (controle positivo) ou veículo (controle negativo) são injetados por via subcutânea em PBS no flanco esquerdo de camundongos (12 a 20 semanas de idade). O teratoma, por exemplo, tumor, o crescimento é analisado 2 a 3 vezes por semana determinando-se o volume do tumor por medições do calibrador por oito semanas após fusossoma, célula tumoral ou injeção de veículo.

[001404] Em uma modalidade, os camundongos administrados com fusossomas ou veículo não terão uma formação tumoral mensurável, por exemplo, teratoma, por meio de medições de calibrador. Em uma modalidade, os animais de controle positivo tratados com células tumorais demonstrarão um tumor observável, por exemplo, teratoma, tamanho medido por compassos ao longo das oito semanas de observação.

EXEMPLO 67: MEDIÇÃO DO RNA TOTAL EM UM FUSOSSOMA E CÉLULA DE ORIGEM

[001405] Esse exemplo descreve um método para quantificar a quantidade de RNA em um fusossoma em relação a uma célula de origem. Em uma modalidade, um fusossoma terá níveis de RNA semelhantes à célula de origem. Nesse ensaio, os níveis de RNA são determinados medindo-se o RNA total.

[001406] Os fusossomas são preparados por qualquer um dos métodos descritos nos Exemplos anteriores. As preparações da mesma massa medida pela proteína de fusossomas e células de origem são usadas para isolar o RNA total (por exemplo, com o uso de um kit como o Qiagen RNeasy n° de catálogo 74104), seguido pela determinação da concentração de RNA com o uso de métodos espectroscópicos

padrão para avaliar a absorbância de luz por RNA (por exemplo, com Thermo Scientific NanoDrop).

[001407] Em uma modalidade, a concentração de RNA em fusossomas será 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% daquela das células de origem por massa de proteína.

EXEMPLO 68: ISOLAMENTO DE MICROVESÍCULAS FUSOGÊNICAS LIBERADAS LIVREMENTE DAS CÉLULAS

[001408] Esse exemplo descreve o isolamento de microvesículas fusogênicas liberadas livremente das células. Microvesículas fusogênicas foram isoladas conforme a seguir. 9,2 x 10^6 HEK-293T (ATCC, n° de catálogo CRL-3216) foram submetidas à transfecção reversa com o uso do agente de transfecção Xfect (Takara, n° de catálogo 631317) com 10 μg do plasmídeo de expressão pcDNA3.1 que contém quadro aberto de leitura para VSVg e 15 μg do plasmídeo de expressão pcDNA3.1 que contém o quadro aberto de leitura para a Cre Recombinase do bacteriógrafo P1 com uma sequência de localização nuclear SV40 em 7,5 ml do meio completo (Meio Eagle Modificado de Dulbecco (DMEM) suplementado com GlutaMAX (ThermoFisher), 10% de soro de bezerro bovino (ThermoFisher) e antibióticos de penicilina/estreptomoicina (ThermoFisher)) em uma prato revestida com colágeno a 100 mm (Corning). Doze horas após a semeadura, mais 7,5 ml de meio completo foram adicionados cuidadosamente. As células foram separadas do meio de cultura por centrifugação a 200 *x g* durante 10 minutos. Os sobrenadantes foram coletados e centrifugados sequencialmente duas vezes a 500 xg durante 10 minutos, uma vez a 2.000 xg por 15 minutos, uma vez a 10.000 xg durante 30 minutos e uma vez a 70.000 xg por 60 minutos. Os fusossomas liberados livremente foram peletizados durante a etapa final de centrifugação, suspensos novamente em PBS e peletizados novamente a 70.000 x g. O pélete final foi suspenso novamente em PBS.

[001409] Consultar, também, Wubbolts R *et al*. Proteomic and Biochemical Analyses of Human B Cell-derived Exosomes: Potential Implications for their Function and Multivesicular Body Formation. *J. Biol. Chem*. 278:10.963 a 10.972 2003.

EXEMPLO 69: MEDIÇÃO DA DISTRIBUIÇÃO DE TAMANHO MÉDIO DE FUSOSSOMAS

[001410] Esse exemplo descreve a medição da distribuição de tamanho de fusossomas.

[001411] Os fusossomas foram preparados, conforme descrito no presente documento, por transfecção transitória de HEK293T com VSV-G, enucleação e fracionamento subsequente com Ficoll. Os fusossomas foram medidos para determinar a distribuição de tamanho com o uso do método do Exemplo 28, conforme mostrado na Figura 3. É contemplado que os fusossomas podem ter menos que aproximadamente 50%, 40%, 30%, 20%, 10%, 5% ou menos da variabilidade da célula parental na distribuição de tamanho dentro de 90% da amostra. É contemplado que os fusossomas podem ter 58% a menos da variabilidade da célula parental na distribuição de tamanho em 90% da amostra.

EXEMPLO 70: VOLUME MÉDIO DE FUSOSSOMAS

[001412] Esse exemplo descreve a medição do volume médio dos fusossomas. A variação do tamanho (por exemplo, volume) dos fusossomas pode torná-los versáteis para carregamento de carga, projeto terapêutico ou aplicação distinta.

[001413] Os fusossomas foram preparados, conforme descrito no presente documento, por transfecção transitória de HEK293T com VSV-G, enucleação e fracionamento subsequente com Ficoll. O controle positivo foram células HEK293T.

[001414] A análise com uma combinação de NTA e microscopia confocal conforme descrito no Exemplo 28 foi usada para determinar o

tamanho dos fusossomas. O diâmetro dos fusossomas foi medido e o volume calculado, conforme mostrado na Figura 4. Contempla-se que os fusossomas podem ter um tamanho médio superior a 50 nm de diâmetro. Contempla-se que os fusossomas podem ter um tamanho médio a 129 nm de diâmetro.

EXEMPLO 71: COMPARAÇÃO DA MASSA DE PROTEÍNA SOLÚVEL A INSOLÚVEL

[001415] Esse exemplo descreve a quantificação da razão solúvel:insolúvel da massa da proteína em fusossomas. A razão solúvel:insolúvel de massa de proteína em fusossomas pode, em alguns casos, ser semelhante à de células nucleadas.

[001416] Os fusossomas foram preparados conforme descrito no presente documento por transfecção transitória de HEK293T com VSV-G, enucleação e fracionamento subsequente com Ficoll. A preparação de fusossoma foi testada para determinar a razão de proteína solúvel:insolúvel com o uso de um ensaio do ácido bicinconínico padrão (BCA) (Kit de Ensaio de Proteína BCA Pierce™, produto Thermo Fischer n° 23225). As amostras de proteínas solúveis foram preparadas suspendendo-se os fusossomas ou células parentais preparados a uma concentração de $1x10^7$ células ou ~ 1 mg/ml de fusossomas totais em PBS e centrifugando a 1.500 xg para sedimentar as células ou 16.000 xg para peletizar os fusossomas. O sobrenadante foi coletado como a fração de proteína solúvel.

[001417] Os fusossomas ou células foram então suspensos novamente em PBS. Essa suspensão representa a fração de proteína insolúvel.

[001418] Uma curva padrão foi gerada com o uso do BSA fornecido, de 0 a 15 µg de BSA por poço (em duplicata). O fusossoma ou preparação celular foi diluído de modo que a quantidade medida esteja dentro da faixa dos padrões. A preparação de fusossoma foi analisada em

duplicata, e o valor médio foi usado. A concentração de proteína solúvel foi dividida pela concentração de proteína insolúvel para produzir a razão proteína solúvel:insolúvel (Figura 5).

EXEMPLO 72: MEDIÇÃO DA FUSÃO COM UMA CÉLULA-ALVO

[001419] Os fusossomas derivados de células HEK-293T que expressam a glicoproteína hemaglutinina modificada do vírus do sarampo (MvH) e a proteína de fusão (F) na superfície celular e que contém a proteína Cre recombinase foram gerados, conforme descrito no presente documento. O MvH foi projetado de modo que sua ligação natural ao receptor seja eliminada e a especificidade da célula-alvo seja fornecida através de um anticorpo de cadeia única (scFv) que reconhece o antígeno de superfície celular, nesse caso, o scFv é projetado para atingir CD8, um correceptor para o receptor de células T. Foi usado um fusossoma de controle derivado de células HEK-293T que expressa o fusogênio VSV-G em sua superfície e que contém a proteína Cre recombinase. A célula-alvo era uma célula HEK-293T projetada para expressar um cassete “Loxp-GFP-stop-Loxp-RFP” sob o promotor CMV, assim como projetada para expressar em excesso os correceptores CD8a e CD8b. A célula não alvo era a mesma célula HEK-293T que expressa o cassete “Loxp-GFP-stop-Loxp-RFP”, porém sem superexpressão de CD8a/b. As células destinatárias alvo ou não alvo foram semeadas com 30.000 células/poço em uma placa de 96 poços de fundo transparente preto e cultivadas em meio DMEM com 10% de soro fetal bovino a 37 °C e 5% de CO_2. Quatro a seis horas após as células receptoras serem colocadas em placas, os fusossomas que expressam a proteína Cre recombinase e MvH+F foram aplicadas às células-alvo ou não alvo receptoras em meio DMEM. As células receptoras foram tratadas com 10 µg de fusossomas e incubadas durante 24 horas a 37 °C e 5% de CO_2.

[001420] As placas celulares foram fotografadas com o uso de um

microscópio automatizado (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). A população total de células em um determinado poço foi determinada marcando-se primeiramente as células com Hoechst 33342 em meio DMEM durante 10 minutos. O Hoechst 33342 marca os núcleos de células intercalando-se em DNA e, portanto, é usado para identificar células individuais. A imagem da Hoechst foi obtida com o uso de LED de 405 nm e cubo de filtro DAPI. A imagem da GFP foi obtida com o uso do LED de 465 nm e cubo de filtro GFP, ao passo que a imagem de RFP foi obtida com o uso do LED de 523 nm e do cubo de filtro de RFP. Imagens de poços de células-alvo e não alvo foram obtidas estabelecendo-se primeiramente a intensidade do LED e os tempos de integração em um poço de controle positivo; isto é, células receptoras tratadas com adenovírus que codificam para Cre recombinase em vez de fusossomas.

[001421] As configurações de aquisição foram definidas de modo que as intensidades Hoescht, RFP e GFP estivessem nos valores máximos de intensidade de pixel, porém não saturados. Em seguida, os poços de interesse foram fotografados com o uso das configurações estabelecidas. O foco foi definido em cada poço por autofoco no canal Hoescht e, em seguida, com o uso do plano focal estabelecido para os canais GFP e RFP. A análise de células positivas para GFP e RFP foi realizada com software Gen5 fornecido com microscópio fluorescente automatizado (https://www.biotek.com/products/software-robotics-software/ gen5-microplate-reader-and-imager-software/).

[001422] As imagens foram pré-processadas com o uso de um algoritmo de subtração em segundo plano com esfera rolante com uma largura de 60 µm. As células com intensidade de GFP significativamente acima das intensidades de fundo foram submetidas a um limiar, e áreas muito pequenas ou grandes para serem células GFP-positivas

foram excluídas. As mesmas etapas de análise foram aplicadas ao canal RFP. O número de células positivas para RFP (células receptoras que recebem Cre) foi, então, dividido pela soma das células positivas para GFP (células receptoras que não mostraram entrega) e células positivas para RFP para quantificar a conversão percentual de RFP, que descreve a quantidade de fusão de fusossoma dentro da pop µlação de células-alvo receptoras e não alvo. Para quantidades de fusão alvejada (fusão de fusossoma com células receptoras alvejadas), o valor de conversão de RFP percentual é normalizado para a porcentagem de células receptoras que são células receptoras-alvo (isto é, expressando CD8), que foi avaliada por marcação com anticorpo anti-CD8 conjugado à ficoeritrina (PE) e analisada por citometria de fluxo. Finalmente, a quantidade absoluta de fusão alvejada foi determinada subtraindo-se a quantidade de fusão de células não alvo da quantidade de fusão de células alvo (qualquer valor <0 foi considerado como 0).

[001423] Com esse ensaio, o fusossoma derivado de uma célula HEK-293T que expressa o MvH(CD8)+F modificado em sua superfície e que contém a proteína Cre recombinase mostrou uma conversão de RFP percentual de 25,2 +/- 6,4% quando a célula receptora era a célula-alvo HEK-293T que expressa o cassete “Loxp-GFP-stop-Loxp-RFP”, e 51,1% dessas células receptoras foram observadas como positivas para CD8. A partir desses resultados, a conversão de RFP de porcentagem normalizada ou quantidade de fusão direcionada foi determinada como sendo 49,3 +/- 12,7% para a fusão direcionada. O mesmo fusossoma mostrou uma conversão de RFP percentual de 0,5 +/- 0,1% quando o receptor era a célula não alvo HEK-293T que expressa “Loxp-GFP-stop-Loxp-RFP”, porém sem expressão de CD8. Com base no supracitado, a quantidade absoluta de fusão direcionada para o fusossoma MvH(CD8)+F determinada como sendo 48,8% e a quantida-

de absoluta de fusão direcionada para o fusossoma VSV-G de controle foi determinada como 0% (Figura 6).

EXEMPLO 73: MEDIÇÃO DA CAPACIDADE DE TRANSPORTAR GLICOSE ATRAVÉS DA MEMBRANA CELULAR

[001424] Os fusossomas de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular e que expressa a proteína Cre recombinase foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter fusossomas de partículas pequenas, conforme descrito no presente documento. A fim de medir a capacidade que os fusossomas têm para transportar glicose através da membrana celular, os níveis de um análogo de glicose fluorescente e um 2-NBDG (2-(*N*-(7-Nitrobenz-2-oxa-1,3-diazol-4-il)Amino)-2-desoxiglicose), que pode ser usado para monitorar a absorção de glicose em células vivas, foram quantificados para avaliar o transporte ativo através da bicamada lipídica. Um kit comercialmente disponível da Biovision Inc. (n° de catálogo K682) foi usado para o ensaio de acordo com as instruções do fabricante.

[001425] Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Em seguida, 40 ug de total de proteína de fusossoma foram peletizados por centrifugação a 3.000 g durante 5 minutos em uma centrífuga de mesa, seguido de ressuspensão em 400 µl de DMEM suplementado com 0,5% de soro fetal bovino. Isso foi feito em duplicata para cada amostra, e uma das duplicatas foi tratada com 4 ul de floretina (fornecida com o kit), um fenol natural que inibe a absorção de glicose, como controle da inibição da absorção de glicose. Em seguida, as amostras foram incubadas durante 1 hora à temperatura ambiente. Após a incubação, a amostra de fusossoma foi peletizada e sus-

pensa novamente em 400 ul de mistura de absorção de glicose preparada anteriormente (consultar a Tabela 12 abaixo para formulação). As amostras pré-tratadas com floretina foram suspensas novamente em mistura de absorção de glicose com floretina; as amostras não pré-tratadas foram suspensas novamente na mistura de absorção de glicose com 20 ul de PBS em vez de floretina. Além disso, um conjunto paralelo de amostras de fusossoma foi suspenso novamente em meio DMEM com 0,5% de FBS apenas como um controle negativo para análise de citometria de fluxo.

TABELA 12: FORMULAÇÃO DA MISTURA DE ABSORÇÃO DE GLICOSE

Reagente	**Volume (ul)**
Meios DMEM com 0,5% de FBS	1.880
Reagente 2-NBDG	20
Intensificador da absorção de glicose	100
Opcional: Floretina	20

[001426] As amostras foram, então, incubadas a 37 °C com 5% de CO_2 durante 30 minutos. Após a incubação, as células foram peletizadas, lavadas uma vez com 1 ml de tampão de análise 1X (dotado do kit), peletizadas novamente e suspensas novamente em 400 ul de tampão de análise 1X.

[001427] As amostras foram, então, medidas para a absorção de 2-NBDG por análise de citometria de fluxo com o uso de um citômetro de focalização acústica Invitrogen Attune NxT. 2-NBDG foi submetido à excitação com um laser de 488 nm e teve a emissão capturada a 513 ± 26 nm. As portas de dispersão frontal e lateral foram inicialmente usadas para capturar eventos do tamanho de fusossoma e descartar pequenos detritos. Eventos positivos para 2-NBDG foram determinados por meio de gating no nível mínimo para os quais a amostra de

controle negativa de 2-NBDG mostrou <0,5% de eventos positivos para sinalização de 2-NBDG. As células submetidas a gating positivas para fluorescência de 2-NBDG foram, então, avaliadas quanto à intensidade de fluorescência média (F.I.) de 2-NBDG a fim de calcular um valor para absorção de glicose para os fusossomas com e sem tratamento com floretina.

[001428] Com esse ensaio, o fusossoma derivado de uma célula HEK-293T que expressa a VSV-G e Cre mostrou uma F.I. média de 2-NBDG de 631,0 +/- 1,4 sem tratamento com floretina e uma F.I. média de 565,5 +/- 4,9 com tratamento com floretina (Figura 7).

EXEMPLO 74: MEDIÇÃO DA ATIVIDADE DE ESTERASE NO CITOSOL

[001429] Os fusossomas de células C2C12 foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter fusossomas de partículas pequenas, conforme descrito no presente documento. A fim de medir a atividade da esterase no citosol dos fusossomas, as amostras foram marcadas com Calceína AM (BD Pharmigen, n° de catálogo 564061), um derivado de fluoresceína e corante vital não fluorescente que atravessa passivamente a membrana celular de células viáveis e é convertido por esterases citosólicas em calceína fluorescente verde, que é retida pelas células com membranas intactas e proteína de resistência a múltiplos fármacos.

[001430] Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Em seguida, 20 µg de toda a proteína do fusossoma foram peletizados por centrifugação a 3.000 g durante 5 minutos em uma centrífuga de mesa, seguido de suspensão nova em 400 µl de DMEM suplementado com 0,5% de soro fetal bovino. O corante permeável à

membrana, calceína-AM, foi preparado como uma solução-mãe de 10 mM em dimetilsulfóxido e como uma solução de trabalho de 1 mM em tampão PBS, pH 7,4. Os fusossomas VSV-G foram marcados com solução 1 µM de calceína-AM diluída em meio DMEM. As amostras foram incubadas a 37 °C no escuro durante 30 minutos e, em seguida, peletizadas por centrifugação. Após lavagem duas vezes com tampão PBS, os fusossomas foram suspensos novamente em PBS e analisados por citometria de fluxo.

[001431] As amostras foram medidas para retenção de fluorescência de calceína com o uso de um citômetro de focalização acústica Invitrogen Attune NxT. A calceína AM foi submetida à excitação com um laser de 488 nm e teve a emissão capturada a 513 ± 26 nm. As portas de dispersão frontal e lateral foram inicialmente usadas para capturar eventos do tamanho de fusossoma e descartar pequenos detritos. Eventos positivos para calceína foram determinados por meio de gating no nível mínimo para os quais a amostra de controle negativa de calceína mostrou <0,5% de eventos positivos para sinalização de calceína. As células submetidas a gating positivas para fluorescência de calceína foram, então, avaliadas quanto à intensidade de fluorescência (F.I.) de calceína a fim de calcular um valor para atividade de esterase no citosol de fusossomas.

[001432] Com esse ensaio, o fusossoma derivado de uma célula C2C12 mostrou uma atividade de esterase (média de calceína F.I.) de 631,0 +/- 1,4 (Figura 8).

EXEMPLO 75: MEDIÇÃO DA ATIVIDADE DA ACETILCOLINESTERASE EM FUSOSSOMAS

[001433] Fusossomas de células HEK-293T que expressam a proteína de fusão célula-célula placentária sincitina-1 (Syn1) na superfície celular e que expressam a proteína Cre recombinase foram gerados conforme descrito no presente documento. A atividade da acetilcoli-

nesterase foi medida com o uso do Kit de Quantificação FluoroCet (System Biosciences, n° de catálogo FCET96A-1) seguindo as recomendações do fabricante.

[001434] Resumidamente, os fusossomas foram peletizados por ultracentrifugação a 120.000 g por 90 minutos e suspensos novamente cuidadosamente em solução salina tamponada com fosfato (PBS). Em seguida, os fusossomas foram quantificados quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Após a quantificação de BCA da concentração de proteína, 1.000 ng de proteína fusossoma total foi diluído com PBS para um volume de 60 ul, seguido pela adição de 60 ul de tampão de lise para lisar as partículas. Após uma incubação de 30 minutos em gelo, as amostras estavam prontas para serem processadas no ensaio FluoroCet.

[001435] Em poços duplicados de uma placa de 96 poços, 50 ul de amostra de fusossoma lisada foram misturados com 50 ul de estoque de Trabalho de Tampão A e 50 ul de estoque de trabalho de Tampão B. Em paralelo, uma curva padrão foi preparada pipetando-se 2 ul do padrão fornecido em 126 ul de tampão de reação 1X. Essa solução padrão foi, então, diluída em série 5X para fazer uma curva padrão de seis pontos que consiste em 2,0E+08, 1,0E+08, 5,0E+07, 2,5E+07, 1,25E+07, e 6,25E+06 equivalentes de exossoma da atividade da acetilcolinesterase. 50 ul de cada padrão foram, então, misturados com 50 ul de estoque de trabalho de Tampão A e 50 ul de estoque de trabalho de Tampão B em poços duplicados da placa de 96 poços. 50 ul de tampão de reação 1X foi usado como um branco. A placa foi misturada batendo-se nas laterais seguido de incubação no escuro por 20 minutos em temperatura ambiente. A placa foi, então, medida imediatamente com o uso de um leitor de placa de fluorescência ajustado para Excitação: 530 a 570 nm e Emissão: 590 a 600 nm. A placa foi agitada

durante 30 segundos antes da leitura.

[001436] As unidades de fluorescência relativa (RFU) foram, então, representadas graficamente contra os equivalentes de exossoma conhecidos da atividade da acetilcolinesterase após subtrair os valores de RFU dos poços em branco. Em seguida, uma linha de regressão linear foi calculada e a equação usada para determinar a atividade da acetilcolinesterase (em equivalentes de exossoma) para as amostras de fusossoma a partir dos valores de RFU medidos. A atividade de acetilcolinesterase medida para fusossomas Syn1 é mostrada na Tabela 13:

TABELA 13: ATIVIDADE DA ACETILCOLINESTERASE EM FUSOSSOMAS E PARTÍCULAS DE CONTROLE

Amostra	**Atividade de acetilcolinesterase (equivalentes de exossoma)**
Fusossomas de Syn1	6,83E+05 +/- 2,21E+05

EXEMPLO 76: MEDIÇÃO DO NÍVEL DE ATIVIDADE METABÓLICA

[001437] Os fusossomas de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular e que expressa a proteína Cre recombinase foram gerados, conforme descrito no presente documento. A fim de determinar o nível de atividade metabólica da preparação de fusossoma, a atividade da citrato sintase foi avaliada com o uso de um kit comercialmente disponível da Sigma (n° de catálogo CS0720) que fornece todos os reagentes necessários. A citrato sintase é uma enzima do ciclo do ácido tricarboxílico (TCA) que catalisa a reação entre o oxaloacetato (OAA) e o acetil-CoA para gerar citrato. Após a hidrólise da acetil-CoA, ocorre a liberação de CoA com um grupo tiol (CoA-SH). O grupo tiol reage com um reagente químico, 5,5-Ditiobis (ácido 2-nitrobenzoico) (DTNB), para formar o ácido 5-tio-2-nitrobenzoico (TNB), que tem um produto amarelo que pode ser medido por espectrofotometria a 412

nm.

[001438] O ensaio foi realizado de acordo com as recomendações do fabricante. Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Em seguida, 400 µg de total de proteína de fusossoma foram peletizados por centrifugação a 3.000 g durante 5 minutos em uma centrífuga de mesa. Os fusossomas foram lavados uma vez por peletização novamente e suspensos novamente em PBS gelado. Os fusossomas foram peletizados novamente e o sobrenadante foi removido. O sedimento foi lisado em 100 µl de tampão CelILytic M com inibidores de protease 1X. Após mistura por pipetagem, a amostra lisada foi incubada durante 15 minutos à temperatura ambiente para completar a lise. A amostra foi, então, centrifugada a 12.000 g durante 10 minutos, e o sobrenadante foi transferido para um novo tubo de microcentrífuga e armazenado a -80 °C até o ensaio subsequente ser realizado.

[001439] A fim de iniciar o ensaio de atividade da citrato sintase, todas as soluções de ensaio foram aquecidas à temperatura ambiente antes de usar. A amostra de fusossoma lisada foi misturada com soluções de ensaio de acordo com a Tabela 14 abaixo:

TABELA 14: ESQUEMA DE REAÇÃO PARA MEDIÇÃO DA ATIVIDADE DA CITRATO SINTASE EM PLACA DE 96 POÇOS

Amostra	**Tampão de ensaio**	**Solução de Acetil CoA 30 mM**	**Solução de DTNB 10 mM**	**Solução de OAA 10 mM (adicionada por último)**
4 ul	182 ul	2 ul	2 ul	10 ul

[001440] Os volumes na Tabela 14 representam os volumes para um único poço de uma placa de 96 poços. As amostras foram medidas em duplicatas. Todos os componentes da reação foram misturados e pipetados em um único poço de uma placa de 96 poços. A absorbância a

412 nm foi, então, analisada em um leitor de microplacas por 1,5 minuto para medir a reação de linha de base. Em seguida, 10 ul da solução de OAA 10 mM foram adicionados a cada poço para iniciar a reação. A placa foi agitada por 10 segundos no leitor de microplacas antes da leitura da absorbância a 412 nm por 1,5 minutos com medição a cada 10 segundos.

[001441] A fim de calcular a atividade da citrato sintase, a absorbância a 412 nm foi plotada em função do tempo para cada reação. A mudança na absorbância por minuto foi calculada para a faixa linear do gráfico para antes (atividade endógena) e depois (atividade total) da adição de OAA. A atividade da citrato sintase líquida foi, então, calculada subtraindo-se a atividade endógena da atividade total para a amostra. Este valor foi, então, usado para calcular a atividade da citrato sintase com base na equação e nos valores constantes fornecidos pelo fabricante. A atividade de citrato sintase medida para os fusossomas de VSV-G foi 1,57E-02 +/- 1,86E-03 µmol/ug de fusossoma/minuto.

EXEMPLO 77: MEDIÇÃO DOS NÍVEIS DE RESPIRAÇÃO

[001442] Os fusossomas de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter fusossomas de partículas pequenas, conforme descrito no presente documento. O nível de respiração na preparação de fusossoma foi determinado medindo-se as taxas de consumo de oxigênio mitocondrial por um analisador de fluxo extracelular Seahorse (Agilent).

[001443] Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Os próximos 20 µg de total de proteína de fusossoma foram

peletizados por centrifugação a 3.000 g durante 5 minutos em uma centrífuga de mesa, seguido por ressuspensão (em quadruplicatas) em 150 µl de meio de ensaio XF (Agilent n° de catálogo 103575-100) suplementado com glicose 25 mM e glutamina 2 mM (pH 7,4). As amostras suspensas novamente foram, então, adicionadas a um poço de uma placa Seahorse de 96 poços (Agilent).

[001444] Os ensaios de consumo de oxigênio foram iniciados incubando-se a placa Seahorse de 96 poços com amostras a 37 °C durante 60 minutos para permitir que a temperatura e o pH atinjam o equilíbrio. A microplaca foi, então, testada no Analisador de Fluxo Extracelular XF96 (Agilent) para medir as mudanças no fluxo extracelular de oxigênio e pH na mídia imediatamente ao redor dos fusossomas. Após obter o consumo de oxigênio em estado estacionário e as taxas de acidificação extracelular, a oligomicina (5 µM), que inibe a ATP sintase, e o ionóforo de prótons FCCP (cianeto de carbonila 4- (trifluorometoxi)fenil-hidrazona; 2 µM), que desacopla mitocôndrias, foram injetados sequencialmente por meio de câmaras de distribuição de reagente para cada poço na microplaca para obter valores para as taxas máximas de consumo de oxigênio. Por fim, 5 µM de antimicina A (inibidor do complexo mitocondrial III) foram injetados para confirmar que as alterações respiratórias eram devidas principalmente à respiração mitocondrial. As taxas de respiração da antimicina A foram subtraídas das outras três taxas de respiração para determinar as taxas de respiração mitocondrial basal, desacoplada (resistente à oligomicina) e máxima (induzida por FCCP).

[001445] Com o uso desse ensaio, foi determinado que os fusossomas do doador VSV-G apresentaram taxas de consumo de oxigênio (respiração) basal, desacoplada e máxima de acordo com a Tabela 15 abaixo.

TABELA 15: TAXAS DE RESPIRAÇÃO DOS FUSOSSOMAS VSV-G

Estado de respiração	**Taxa de consumo de oxigênio mitocondrial (respiração) (pmol/min/20 µg de fusossoma) AVG ± SEM**
Basal	11,3 ± 3,0
Desacoplado	10,1 ± 2,3
Maximal	20,0 ± 1,9

EXEMPLO 78: MEDIÇÃO DOS NÍVEIS DE FOSFATIDILSERINA DE FUSOSSOMAS

[001446] Os fusossomas de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular e que expressa a proteína Cre recombinase foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter fusossomas de partículas pequenas, conforme descrito no presente documento. A fim de medir os níveis de fosfatidilserina dos fusossomas, a marcação com anexina V foi realizada com o uso da anexina V disponível comercialmente conjugada com corante Alexa Fluor 647 (n° de catálogo A23204) de acordo com as instruções do fabricante. A anexina V é uma proteína celular que pode se ligar à fosfatidilserina quando exposta no folheto externo da membrana plasmática; assim, a leitura da ligação da anexina V a uma amostra pode fornecer uma avaliação dos níveis de fosfatidilserina na amostra.

[001447] Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Em seguida, 40 µg de total de proteína de fusossoma foram peletizados por centrifugação (em triplicatas de amostra) a 3.000 g durante 5 minutos em uma centrífuga de mesa, seguido de ressuspensão em 400 ul de DMEM suplementado com 2% de soro fetal bovino. Uma amostra foi tratada com antimicina A a 40 µM. As amostras foram, en-

tão, incubadas durante 1 hora a 37 °C. Após incubação as amostras foram, então, peletizadas por centrifugação novamente e suspensas novamente em 100 de tampão de ligação a anexina (ABB; HEPES a 10 mM, NaCl a 140 mM, $CaCl_2$ a 2,5 mM, pH 7,4). Em seguida, 5 µl de anexina V conjugada com Alexa Fluor 647 foram adicionados a cada amostra (com exceção do controle negativo sem marcação com anexina V). As amostras foram incubadas durante 15 minutos à temperatura ambiente, seguido da adição de 400 µl de ABB.

[001448] As amostras foram, então, medidas para a marcação com anexina V por análise de citometria de fluxo com o uso de um citômetro de focalização acústica Invitrogen Attune NxT. Anexina V conjugada com Alexa Fluor 647 foi submetida à excitação com um laser de 638 nm e teve a emissão capturada a 670 ± 14 nm. As portas de dispersão frontal e lateral foram inicialmente usadas para capturar eventos do tamanho de fusossoma e descartar pequenos detritos. Os eventos positivos para a marcação de Alexa Fluor 647 (anexina V) foram determinados por passagem no nível mínimo para o qual a amostra de controle negativa com anexina V não marcada apresentou <0,5% de eventos positivos para a marcação de Alexa Fluor 647. Os eventos fechados positivos para a marcação Alexa Fluor 647 foram, então, avaliados quanto à porcentagem de eventos positivos para anexina V da população parental total (eventos de tamanho de fusossoma no portão de dispersão frontal/lateral) e esse valor foi usado como a quantificação dos níveis de fosfatidilserina na amostra de fusossoma.

[001449] Com esse ensaio, o fusossoma derivado de uma célula HEK-293T que expressa a VSV-G e Cre mostrou uma % de fusossomas positivos para anexina V de 63,3 ± 2,3% sem tratamento com antimicina A e porcentagem de fusossomas positivos para anexina V de 67,6 ± 5,7% com tratamento com antimicina A.

EXEMPLO 79: MEDIÇÃO DO POTENCIAL MÉDIO DE MEMBRANA MITOCONDRIAL

[001450] Os fusossomas de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular e que expressa a proteína Cre recombinase foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter fusossomas de partículas pequenas, conforme descrito no presente documento. A fim de medir os níveis médios de potencial de membrana mitocondrial dos fusossomas, um corante disponível comercialmente que é sensível ao potencial de membrana mitocondrial, tetrametil rodamina, éster etílico, perclorato (TMRE; Abcam, n° de catálogo T669) foi usado para avaliar o potencial de membrana mitocondrial. Para normalizar a intensidade de fluorescência de TMRE (FI) para a quantidade de mitocôndrias na amostra, o corante MitoTracker Green FM (MTG; ThermoFisher, n° de catálogo M7514) foi usado para comarcar amostras a fim de normalizar F.I. de TMRE para F.I. de MTG e, assim, para a quantidade de mitocôndrias na amostra. Além disso, o cianeto de carbonil-p-trifluorometoxifenil-hidrazona (FCCP; n° de catálogo Sigma C2920) foi usado para tratar um conjunto paralelo de amostras a fim de despolarizar totalmente o potencial de membrana mitocondrial e, assim, permitir a quantificação do potencial de membrana mitocondrial em milivolts com base na diminuição de F.I. de TMRE.

[001451] Resumidamente, a amostra de fusossoma foi medida quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, ThermoFisher, n° de catálogo 23225) de acordo com as instruções do fabricante. Em seguida, 40 µg de total de proteína de fusossoma foram peletizados por centrifugação (em quadruplicatos de amostra para duplicatas não tratadas e tratadas com FCCP) a 3.000 g durante 5 minutos em uma centrífuga de mesa, seguido por ressuspensão em 100 µl

de DMEM suplementado com 2% de soro fetal bovino e que contém os corantes TMRE e MTG em uma concentração final de 30 nM e 200 nM, respectivamente. Um conjunto paralelo de amostras de fusossoma foi deixado sem marcação como controle negativo. As amostras foram incubadas por 45 minutos a 37 °C. Após a incubação, as amostras foram peletizadas por centrifugação e suspensas novamente em 400 µl de meio DMEM sem vermelho de fenol que contém TMRE a 30 nm. Um conjunto de duplicatas foi tratado com FCCP 20 µM durante 5 minutos antes da avaliação por citometria de fluxo.

[001452] As amostras foram, então, medidas para a marcação com anexina V por análise de citometria de fluxo com o uso de um citômetro de focalização acústica Invitrogen Attune NxT. O MTG foi submetido à excitação com um laser de 488 nm e teve a emissão capturada a 530 ± 30 nm. TMRE foi submetido à excitação com um laser de 561 nm e teve a emissão capturada a 585 ± 16 nm. As portas de dispersão frontal e lateral foram inicialmente usadas para capturar eventos do tamanho de fusossoma e descartar pequenos detritos. Os eventos positivos para marcação com MTG e TMRE foram determinados por gating no nível máximo para o qual a amostra de controle não marcada mostrou <0,5% de eventos positivos para marcação com MTG ou TMRE. Os eventos fechados positivos para marcação de MTG e TMRE foram, então, avaliados quanto à F.I. média de MTG e TMRE.

[001453] Os valores do potencial da membrana (em milivolts, mV) são calculados com base na intensidade de TMRE após normalizar os valores da F.I. de TMRE para valores de F.I. de MTG. Este valor da relação TMRE/MTG permite a normalização da intensidade do TMRE para a quantidade de mitocôndrias na amostra. O valor da razão TMRE/MTG para as amostras não tratadas e tratadas com FCCP é calculado e usado para determinar o potencial de membrana em milivolts com o uso de uma equação de Nernst modificada (consultar a

seguir) que pode determinar o potencial de membrana mitocondrial com base na fluorescência de TMRE (como TMRE se acumula em mitocôndrias à maneira Nernstiana). O potencial de membrana de fusossoma é calculado com a seguinte fórmula: (mV) = -61,5 * log(FI(não tratado)/FI(FCCP-tratado)). Usando essa equação, o potencial de membrana mitocondrial calculado da amostra de fusossoma da VSV-G foi de -29,6 ± 1,5 milivolts.

EXEMPLO 80: MEDIÇÃO DO POTENCIAL DE SEGMENTAÇÃO EM UM INDIVÍDUO (GESÍCULAS BIVS-CRE)

[001454] Esse exemplo avalia a capacidade de um fusossoma de atingir um sítio específico do corpo. Os fusossomas foram derivados com o uso de métodos, conforme descritos no presente documento e foram carregados com a proteína cre-recombinase.

[001455] Duas doses de fusossomas (1x e 3x) foram entregues em Loxp Luciferase (Jackson Laboratory, 005125). Os camundongos foram injetados por via intravenosa (IV) através da veia da cauda. Os camundongos foram colocados sob uma lâmpada de calor (com o uso de uma lâmpada de calor de 250 W (infravermelho)) durante ~ 5 minutos (ou até que os camundongos comecem a limpar seus bigodes excessivamente) para dilatar a veia da cauda. Os camundongos foram colocados em um limitador, e a cauda foi limpa com etanol 70% para melhor visualizar a veia.

[001456] Com o uso de uma seringa de tuberculina, 200 µl de solução a 1x (8,5e8±1,4e8 partículas/ µl, média (SEM)) ou solução 3x (2,55e9±1,4e8 partículas/ µl, média (SEM)) de fusossoma foi injetada por via IV. Após a conclusão da injeção, a seringa foi removida e a pressão aplicada no sítio da injeção.

[001457] Após a fusão, a proteína CRE translocou para o núcleo para realizar a recombinação, que resultou na expressão constitutiva da luciferase. Três dias após o tratamento, a região ventral dos indivíduos

foi preparada depilando-se a área (Creme Removedor Nair Hair durante 45 segundos, seguido de limpeza da área com etanol 70%). Os indivíduos foram, então, tratados com D-luciferina (Perkin Elmer, 150 mg/kg) por via intraperitoneal. Isso permitiu a detecção da expressão da luciferase por meio de imagens bioluminescentes *in vivo*. O animal foi colocado em uma câmara de imageamento por bioluminescência *in vivo* (Perkin Elmer), que abriga um anestesiante cone (isoflurano) para evitar o movimento do animal. A coleta de fótons foi realizada entre 3 a 15 minutos após a injeção para observar o sinal bioluminescente máximo devido à depuração farmacocinética da D-luciferina. A radiância máxima foi registrada em fótons/s/cm2/radianos. O fluxo total, que integra a radiância sobre a área, foi quantificado com o uso de uma ferramenta de região de interesse (ROI) dentro do Software de Living Image (Perkin Elmer) e relatado em fótons/s.

[001458] A evidência de entrega de proteína (Cre recombinase) por fusossomas foi detectada por imageamento por bioluminescência no tecido receptor do animal, conforme mostrado Figuras 9A a 9B. O sinal foi visto principalmente no baço e no fígado, com o grupo 3x que mostra o sinal mais alto.

[001459] Após a imagem de corpo inteiro, os camundongos tiveram a cervical deslocada, o e fígado, coração, pulmões, rim, intestino delgado, pâncreas e baço foram coletados e fotografados em 5 minutos após a eutanásia. A Evidência da entrega de proteína (Cre recombinase) ao fígado e baço por fusossomas foi detectada por imageamento por bioluminescência no tecido receptor extraído dos animais. Isso pode ser visto nas Figuras 10A a 10B. O sinal foi mais alto no baço e o mais baixo no coração, com o grupo 3x mostrando o sinal significativo mais alto ($p = 0,0004$ em comparação ao coração).

EXEMPLO 81: ENTREGA DE FUSOSSOMAS POR MEIO DE UMA VIA QUE É INDEPENDENTE DA ACIDIFICAÇÃO DE LISOSSOMAS

[001460] Muitas vezes, a entrada de cargas biológicas complexas nas células-alvo é realizada por endocitose. A endocitose exige que a carga entre em um endossomo, que amadurece em um lisossoma acidificado. Desvantajosamente, a carga que entra em uma célula por meio de endocitose pode ficar presa em um endossomo ou lisossoma e ser incapaz de alcançar o citoplasma. A carga também pode ser danificada por condições ácidas no lisossoma. Alguns vírus têm capacidade para entrada não endocítica nas células-alvo; no entanto, esse processo não é completamente entendido. Esse exemplo demonstra que um fusogênio viral pode ser isolado do resto do vírus e conferir entrada não endocítica em um fusossoma que carece de outras proteínas virais.

[001461] Os fusossomas de células HEK-293T que expressa a proteína G de ligação ao receptor do vírus Nipah e a proteína F de fusão (NivG+F) na superfície celular e que expressa a proteína Cre recombinase foram gerados de acordo com o procedimento padrão de ultracentrifugação através de um gradiente de Ficoll para obter pequenas partículas fusossomas, conforme descrito no presente documento. A fim de demonstrar a entrega do fusossoma a uma célula receptora por meio de uma via não endocítica, os fusossomas NivG+F foram usados para tratar as células receptoras HEK-293T projetadas para expressar um cassete “Loxp-GFP-stop-Loxp-RFP” sob o promotor CMV. A proteína NivF é uma glicoproteína de envelope independente do pH que demonstrou não precisar de acidificação ambiental para ativação e subsequente atividade de fusão (Tamin, 2002).

[001462] As células receptoras foram semeadas com 30.000 células/poço em uma placa de 96 poços de fundo transparente preto. Quatro a seis horas após as células receptoras serem colocadas em placas, os fusossomas NivG+F que expressam a proteína Cre recombinase foram aplicados às células-alvo receptoras ou não alvo em meio

DMEM. A amostra de fusossoma foi medida primeiramente quanto ao teor total da proteína por ensaio de ácido bicinconínico (BCA, Thermo-Fisher, n° de catálogo 23225) de acordo com as instruções do fabricante. As células receptoras foram tratadas com 10 µg de fusossomas e incubadas por 24 horas a 37 °C e 5% de CO_2. Para demonstrar que a entrega de Cre por meio dos fusossomas NivG+F ocorreu por meio de uma via não endocítica, poços paralelos de células receptoras que recebem o tratamento com o fusossoma NivG+F foram cotratados com um inibidor de acidificação de endossomo/lisossoma, bafilomicina A1 (Baf; 100 nM; Sigma, n° de catálogo B1793).

[001463] As placas celulares foram fotografadas com o uso de um microscópio automatizado (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). A população total de células em um determinado poço foi determinada marcando-se primeiramente as células com Hoechst 33342 em meio DMEM durante 10 minutos. O Hoechst 33342 marca os núcleos de células intercalando-se em DNA e, portanto, foi usado para identificar células individuais. A marcação com Hoechst foi obtida com o uso de LED de 405 nm e cubo de filtro DAPI. A imagem da GFP foi obtida com o uso do LED de 465 nm e cubo de filtro GFP, ao passo que a imagem de RFP foi obtida com o uso do LED de 523 nm e do cubo de filtro de RFP. Imagens de poços de células-alvo e não alvo foram adquiridas estabelecendo-se primeiramente a intensidade do LED e os tempos de integração em um poço de controle positivo que contém células receptoras tratadas com codificação de adenovírus para Cre recombinase em vez de fusossomas.

[001464] As configurações de aquisição foram definidas de modo que as intensidades Hoescht, RFP e GFP estivessem nos valores máximos de intensidade de pixel, porém não saturados. Em seguida, os poços de interesse foram fotografados com o uso das configurações

estabelecidas. O foco foi definido em cada poço por autofoco no canal Hoescht e, em seguida, com o uso do plano focal estabelecido para os canais GFP e RFP. A análise de células positivas para GFP e RFP foi realizada com software Gen5 fornecido com microscópio fluorescente automatizado (https://www.biotek.com/products/software-robotics-software/gen5-microplate-reader-and-imager-software/).

[001465] As imagens foram pré-processadas com o uso de um algoritmo de subtração em segundo plano com esfera rolante com uma largura de 60 μm. As células com intensidade de GFP significativamente acima das intensidades de fundo foram submetidas a um limiar, e áreas muito pequenas ou grandes para serem células GFP-positivas foram excluídas. As mesmas etapas de análise foram aplicadas ao canal RFP. O número de células positivas para RFP (células receptoras que recebem Cre) foi, então, dividido pela soma das células positivas para GFP (células receptoras que não mostraram entrega) e células positivas para RFP para quantificar a conversão de RFP percentual, que indica a quantidade de fusão de fusossoma com as células receptoras.

[001466] Com esse ensaio, o fusossoma derivado de uma célula HEK-293T que expressa NivG+F em sua superfície e que contém a proteína Cre recombinase mostrou entrega significativa por meio de uma via independente de lisossoma, que é consistente com a entrada por uma via não endocítica, conforme evidenciado por uma entrega significativa de carga Cre por fusossomas NivG+F, até mesmo quando as células receptoras foram cotratadas com Baf para inibir a absorção mediada por endocitose (Figura 11). Nesse caso, a inibição da entrega de carga pelo cotratamento Baf foi de 23,4%.

EXEMPLO 82: MEDIÇÃO DA CAPACIDADE DE POLIMERIZAR ACTINA PARA MOBILIDADE

[001467] Os fusossomas foram gerados pelo procedimento padrão

de colheita e preparação de fusossomas produzidos a partir de células HEK-293T que expressam a glicoproteína G do envelope do vírus da estomatite vesicular (VSV-G) na superfície celular, conforme descrito no presente documento. As partículas de controle (fusossomas não fusogênicos) foram produzidas a partir de células HEK-293T transitoriamente submetidas à transfecção reversamente com vetor vazio pcDNA3.1. Os fusossomas e as células parentais foram, então, testados quanto à sua capacidade de polimerizar actina (ao longo do tempo) com o uso de um ensaio do citometria de fluxo de rodamina-faloidina e ELISA Tubulina. Resumidamente, aproximadamente $1x10^6$ fusossomas que correspondem a 60 µl de uma preparação de fusossoma VSV-G padrão e $1x10^5$ células parentais usadas para gerar os fusossomas foram colocados em 1 ml de meio completo em uma placa de 96 poços de fixação baixa de múltiplos poços em incubada a 37 °C e 5% de CO_2. As amostras foram retiradas periodicamente, às 3 horas, 5 horas e 24 horas após o plaqueamento. As amostras foram centrifugadas a 21.000 xg durante 10 minutos, suspensas novamente em 200 µl 4% (v/v) de PFA em solução salina tamponada com fosfato por 10 minutos, lavadas com 1 ml de solução salina tamponada com fosfato, centrifugadas a 21.000 xg durante 10 minutos, lavadas novamente e armazenadas a 4 ℃ até uso posterior.

[001468] Para a marcação com roamina-faloidina, as amostras foram centrifugadas a 21.000 xg durante 10 minutos e incubadas em 100 µl de Triton X-100 0,1% (v/v) em solução salina tamponada com fosfato durante 20 minutos. Após a incubação de 20 min, um adicional de 100 µl de 0,1% (v/v) Triton X-100 em solução salina tamponada com fosfato contendo 165 µRodamina-faloidina M foi adicionada à amostra e misturada com a pipeta, o controle negativo foi recebido e 100 µl adicionais de 100 µl de Triton X-100 0,1% (v/v) em solução salina tamponada com fosfato apenas. As amostras foram incubadas por 45 minu-

tos antes de serem lavadas com 1 ml de solução salina tamponada com fosfato, centrifugadas a 21.000 xg durante 10 minutos, lavadas novamente e suspensas novamente em 300 µl de solução salina tamponada com fosfato e analisadas por citometria de fluxo (Attune, ThermoFisher) com o uso de um laser de 561 nm para excitação e emissão do filtro 585 +/- 16 nm, conforme mostrado na tabela abaixo:

CONFIGURAÇÕES DO CITÔMETRO DE FLUXO

Corante	Laser/Filtro Attune	Comprimento de onda do Laser	Filtro de Emissão (nm)
AF47	YL1	585	585/16

[001469] O software Attune NxT foi usado para a aquisição e o FlowJo usou a análise. Para aquisição de dados, os canais FSC e SSC foram definidos no eixo linear para determinar uma população representativa das células ou fusossomas. Em seguida, essa população foi submetida a gating e os eventos apenas no interior dessa porta foram usados para exibir eventos no canal de emissão 585 +/- 16 nm em uma escala logarítmica. Um mínimo de 10.000 eventos dentro das células ou porta de fusossomas foi coletado para cada afecção. Para análise de dados, os canais FSC e SSC foram definidos no eixo linear para determinar uma população representativa das células ou fusossomas. Em seguida, essa população foi submetida a gating e os eventos apenas no interior dessa porta foram usados para exibir eventos no canal de emissão 585 +/- 16 nm em uma escala logarítmica. A emissão de controle negativo 585 +/- 16 nm foi usada para determinar onde a porta deve ser colocada no histograma de modo que fosse inferior ao que a porta inclui, menos que 1% positivo. Com o uso dos critérios de análise listados acima, as células parentais demonstraram 19,9%, 24,8% e 82,5% de eventos positivos para rodamina-faloidina, nos instantes do tempo de 3 horas, 5 horas e 24 horas, respectivamente. Os fusossomas eram 44,6%, 41,9% e 34,9% rodamina-faloidina nos ins-

tantes do tempo de 3 horas, 5 horas e 24 horas, respectivamente (Figura 2). Esse exemplo demonstra que os fusossomas não aumentam em quantidade de actina ao longo do tempo, ao passo que as células-mãe aumentam.

EXEMPLO 83: MEDIÇÃO DE GAPDH EM FUSOSSOMAS

[001470] Esse exemplo descreve a quantificação do nível de gliceraldeído 3-fosfato desidrogenase (GAPDH) nos fusossomas e o nível relativo de GAPDH nos fusossomas em comparação às células parentais. Os fusossomas foram preparados conforme descrito nos Exemplos 68 e 87.

[001471] A GAPDH foi medida nas células parentais e nos fusossomas com o uso de um ELISA padrão comercialmente disponível para GAPDH (ab176642, Abcam) de acordo com as instruções do fabricante. Os níveis de total de proteína foram medidos de forma similar através do ensaio de ácido bicinconínico. Os níveis medidos de GAPDH e proteína são mostrados na tabela abaixo:

	[Proteína] (mg/ml)	[GAPDH] (ng/ml)	GAPDH:Proteína (μg/g)
Fusossomas	0,82	37,2	45,3
Células	0,45	50,4	112,0

[001472] As razões entre o total de proteína e GAPDH são mostradas na Figura 12.

EXEMPLO 84: RAZÃO ENTRE LIPÍDIOS E PROTEÍNAS EM FUSOSSOMAS

[001473] Esse exemplo descreve a quantificação da razão entre massa de lipídios e massa de proteína em fusossomas. Contempla-se que os fusossomas podem ter uma razão entre massa de lipídios e massa de proteína que é semelhante àquela de células nucleadas. Os fusossomas e células parentais foram preparados conforme descrito no presente documento nos Exemplos 68 e 87.

[001474] O teor de lipídios foi calculado com o uso de fosfolipídios

que contêm colina como um subconjunto de lipídios totais com o uso de um kit de ensaio de fosfolipídios comercialmente disponível (MAK122 Sigma St. Louis, MO) de acordo com as instruções do fabricante. O teor total da proteína dos fusossomas foi medido por meio de ensaio de ácido bicinconínico, conforme descrito no presente documento. Os níveis de fosfolipídios medidos, os níveis de proteína e a razão entre fosfolipídios e proteína são mostrados na Figura 13 e a tabela abaixo:

	Fosfolipídios (μM)	Proteína (g/l)	Fosfolipídios:Proteína (μmol/g)
Fusossomas	115,6	0,82	141,0
Células	47,9	0,45	106,4

EXEMPLO 85: RAZÃO ENTRE PROTEÍNAS E DNA EM FUSOSSOMAS

[001475] Esse exemplo descreve a quantificação da razão entre massa de proteína e massa de DNA em fusossomas. Contempla-se que os fusossomas podem ter uma razão entre massa de proteína e massa de DNA que é muito maior que a das células. Os fusossomas foram preparados conforme descrito nos Exemplos 68 e 87.

[001476] O teor total da proteína dos fusossomas e das células foi medido por meio do ácido bicinchonínico, conforme descrito no presente documento. A massa de DNA dos fusossomas e células foi medida por absorção a 280 nm após a extração do DNA total com o uso de um kit de isolamento disponível comercialmente (n° 69504 Qiagen Hilden, Alemanha) de acordo com as instruções do fabricante. A razão entre proteínas e ácidos nucleicos totais foi determinada dividindo-se o teor total da proteína pelo teor total do DNA para produzir uma razão dentro de um determinado intervalo para uma preparação típica de fusossoma. Os níveis de proteína medidos, os níveis de DNA e a razão de proteína para DNA são mostrados na Figura 14 e a tabela abaixo:

	[Proteína] (mg/ml)	[DNA] (ng/μl)	Proteína:DNA (g/g)
Fusossomas	0,82	29,5	27,8
Células	0,45	15,9	28,3

EXEMPLO 86: RAZÃO ENTRE LIPÍDIOS E DNA NOS FUSOSSOMAS

[001477] Esse exemplo descreve a quantificação da razão de lipídios para DNA em fusossomas em comparação às células parentais. Em uma modalidade, os fusossomas terão maior razão entre lipídios e DNA em comparação às células parentais. Os fusossomas foram preparados, conforme descrito anteriormente nos Exemplos 68 e 87.

[001478] Essa razão é definida como o teor de lipídios descrito no Exemplo 41, e o teor de ácido nucleico é determinado conforme descrito no Exemplo 42. Os níveis de lipídios medidos, os níveis de DNA e a razão entre lipídios e DNA são mostrados na Figura 15 e a tabela abaixo:

	[Lipídios] (μM)	[DNA] (ng/μl)	Lipídios:DNA (μmol/mg)
Fusossomas	115,6	29,5	3,92
Células	47,9	15,9	3,01

EXEMPLO 87: MEDIÇÃO DA COMPOSIÇÃO LIPÍDICA EM FUSOSSOMAS

[001479] Esse exemplo descreve a quantificação da composição lipídica de fusossomas. Contempla-se que a composição lipídica dos fusossomas pode ser semelhante às células das quais eles são derivados. A composição lipídica afeta parâmetros biofísicos importantes de fusossomas e células, tais como tamanho, interações eletrostáticas e comportamento coloidal.

[001480] As medições de lipídios se basearam em espectrometria de massa. Os fusossomas foram preparados conforme descrito no presente documento por transfecção transitória de VSV-G e GFP em pra-

tos de 10 cm, seguida por filtração e ultracentrifugação do meio condicionado 48 horas após a transfecção para obter fusossomas. As células submetidas à transfecção foram colhidas em paralelo com o meio condicionado e submetidas à análise. Os exossomas também foram colhidos de células que não foram submetidas à transfecção com VSV-G ou GFP.

[001481] A análise de lipídios baseada em espectrometria de massa foi realizada por Lipotype GmbH (Dresden, Alemanha) conforme descrito (Sampaio et al. 2011). Os lipídios foram extraídos com o uso de um procedimento com clorofórmio/metanol em duas etapas (Ejsing et al. 2009). As amostras foram marcadas com mistura padrão de lipídio interna que contém: cardiolipina 16:1/15:0/15:0/15:0 (CL), ceramida 18:1;2/17:0 (Cer), diacilglicerol 17:0/17:0 (DAG), hexosilceramida 18:1;2/12:0 (HexCer), liso-fosfatidato 17:0 (LPA), liso-fosfatidilcolina 12:0 (LPC), liso-fosfatidiletanolamina 17:1 (LPE), liso-fosfatidilglicerol 17:1 (LPG), liso-fosfatidilinositol 17:1 (LPI), liso-fosfatidilserina 17:1 (LPS), fosfatidato 17:0/17:0 (PA), fosfatidilcolina 17:0/17:0 (PC), fosfatidiletanolamina 17:0/17:0 (PE), fosfatidilglicerol 17:0/17:0 (PG), fosfatidilinositol 16:0/16:0 (PI), fosfatidilserina 17:0/17:0 (PS), éster colesterílico 20:0 (CE), esfingomielina 18:1;2/12:0;0 (SM), triacilglicerol 17:0/17:0/17:0 (TAG) e colesterol D6 (Chol).

[001482] Após a extração, a fase orgânica foi transferida para uma placa de infusão e seca em um concentrador a vácuo. O extrato seco da 1ª etapa foi suspenso novamente em acetato de amônio a 7,5 mM em clorofórmio/metanol/propanol (1:2:4, V:V:V) e extrato seco de 2ª etapa em solução de etanol a 33% de metilamina em clorofórmio/metanol (0,003:5:1; V:V:V). Todas as etapas de manipulação de líquidos foram realizadas com o uso da plataforma robótica STARlet da Hamilton Robotics com o recurso de Controle Antigotículas para pipetagem de solventes.

[001483] As amostras foram analisadas por infusão direta em um espectrômetro de massa QExactive (Thermo Scientific) equipado com uma fonte de íons TriVersa NanoMate (Advion Biosciences). As amostras foram analisadas em ambos os modos de íon positivo e negativo com uma resolução de Rm/z = 200 = 28.0000 para MS e Rm/z = 200 = 17.500 para experimentos de MSMS, em uma única aquisição. MSMS foi acionado por uma lista de inclusão que engloba os intervalos de massa de MS correspondentes submetidos à varredura em incrementos de 1 Da (Surma et al. 2015). Os dados de MS e MSMS foram combinados para monitorar íons de CE, DAG e TAG como adutos de amônio; PC, PC O-, como adutos de acetato; e CL, PA, PE, PE O-, PG, PI e PS como ânions desprotonados. MS foi usado apenas para monitorar LPA, LPE, LPE O-, LPI e LPS como ânions desprotonados; Cer, HexCer, SM, LPC e LPC O- como adutos de acetato e colesterol como aduto de amônio de um derivado acetilado (Liebisch et al. 2006).

[001484] Os dados foram analisados com software de identificação de lipídios desenvolvido internamente com base no LipidXplorer (Herzog et al. 2011; Herzog et al. 2012). O pós-processamento e a normalização dos dados foram realizados com o uso de um sistema de gerenciamento de dados desenvolvido internamente. Apenas as identificações de lipídios com uma relação sinal-ruído > 5 e uma intensidade de sinal 5 vezes maior que nas amostras em branco correspondentes foram consideradas para análise de dados posterior.

[001485] A composição lipídica dos fusossomas foi comparada às composições lipídicas das células parentais, com as espécies lipídicas não detectadas atribuídas a um valor zero. As espécies de lipídios identificadas em fusossomas e células parentais são mostradas na tabela abaixo:

Petição 870210022257, de 09/03/2021, pág. 692/731

	Total de Espécies de Lipídios identificadas	Espécies lipídicas compartilhadas (identificadas nas células parentais e fusossomas)	Espécies lipídicas compartilhadas com 25% de expressão parental em fusossomas	Fração de Espécies de Lipídios Compartilhados em Lipídios Totais
Fusossomas	679	569	548	0,700
Células Parentais	783			

[001486] Contempla-se que fusossomas e células parentais podem ter uma composição lipídica semelhante caso ≥70% das espécies de lipídios identificadas em qualquer amostra replicada das células parentais estão presentes em qualquer amostra replicada dos fusossomas e, dentre esses lipídios identificados, o nível médio no fusossoma pode ser > 25% do nível médio de espécies lipídicas correspondente na célula parental.

EXEMPLO 88: MEDIÇÃO DA COMPOSIÇÃO PROTEÔMICA EM FUSOSSOMAS

[001487] Esse exemplo descreve a quantificação da composição de proteína de fusossomas. Contempla-se que a composição de proteína dos fusossomas pode ser semelhante às células parentais das quais são derivados.

[001488] Os fusossomas e células parentais foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87.

[001489] Cada amostra foi suspensa novamente em tampão de lise (ureia 6 M, tioureia 2 M, CHAPS a 4%, Tris 50 mM pH 8,0), submetida à sonicação em um banho de gelo e passada por uma seringa de pequeno calibre. As proteínas foram reduzidas com DTT a 10 mM durante 15 minutos a 65 °C e alquilado com 15 mM de iodoacetamida (IAA) durante 30 minutos no escuro à temperatura ambiente. O excesso de IAA foi extinto com um DTT adicional de 10 mM. As proteínas foram, então, precipitadas com a adição de 8 volumes de acetona gelada + 1 volume de metanol gelado e colocadas a -80 °C de um dia para o outro. As proteínas precipitadas foram peletizadas por centrifugação. O tampão de lise restante foi lavado com 200 µl de metanol gelado 3 vezes. As proteínas foram suspensas novamente em ureia 0,75 M + Tris a 50 mM pH 8,0 + 1 µg de Tripsina/LysC e pré-digeridas por 4 horas a 37 °C°C com agitação. Mais 1 µg de tripsina/LysC foi adicionado às proteínas, e a digestão foi continuada de um dia para o outro. Os peptídeos foram purificados por SPE de fase reversa e analisados por LC-MS.

[001490] Uma amostra replicada para cada afecção foi lisada e combinada em um tubo. Esse pool foi, então, submetido ao mesmo protocolo de preparação que as amostras e analisados por LC-MS nas na aquisição dependente de informações ou separadas em um gel, con-

forme descrito a seguir.

[001491] Um total de 100 µg de proteínas reunidas foi colocado em tampão de carga Laemmli 2x e separado em 12,5% SDS PAGE. As proteínas foram brevemente marcadas com azul de Coomassie e as faixas de proteína foram separadas em 12 frações. Cada fração foi, então, desidratada com 50% de acetonitrila e reidratada com DTT a 10 mM para a redução. Pedaços de gel foram colocados em 65 °C durante 15 minutos e alquilado durante 30 minutos em temperatura ambiente com IAA 15 mM no escuro. Os géis foram desidratados ainda mais com 50% de acetonitrila e reidratados em Tris a 50 mM pH 8 com 1 µg de tripsina/LysC de um dia para o outro a 37 °C. Os peptídeos foram extraídos do gel por desidratação e sonicação. Os peptídeos foram purificados por SPE de fase reversa e analisados por LC-MS/MS (1xIDA por fração).

[001492] A aquisição foi realizada com um ABSciex TripleTOF 5600 (ABSciex, Foster City, CA, EUA) equipado com uma interface de eletroaspersão com um capilar iD de 25 µm e acoplado a um µUHPLC Eksigent (Eksigent, Redwood City, CA, EUA). O software Analyst TF 1.7 foi usado para controlar o instrumento e para processamento e aquisição de dados. A aquisição foi realizada no modo Aquisição Dependente de Informações (IDA) para as 12 frações do gel ou do pool não fracionado. As amostras foram analisadas no modo de aquisição SWATH. Para o modo IDA, a tensão da fonte foi definida para 5,2 kV e mantida a 225 °C, o gás de cortina foi ajustado para 0,19 MPa (27 psi), o gás um a 0,08 MPa (12 psi) e o gás dois a 0,07 MPa (10 psi). Para o modo SWATH, a tensão da fonte foi definida para 5,5 kV e mantida em 225 °C, o gás de cortina foi ajustado para 25 psi, o gás um para 0,11 MPa (16 psi) e o gás dois para 0,1 MPa (15 psi). A separação foi realizada em uma coluna HALO C18-ES de fase reversa de 0,3 mm id, partículas de 2,7 µm, 150 mm de comprimento (Advance Materials

Technology, Wilmington, DE) que foi mantida a 60 °C. As amostras foram injetadas por transbordamento de alça em uma alça de 5 µl. Para o gradiente de LC de 60 minutos, a fase móvel consistia no seguinte solvente A (ácido fórmico a 0,2% v/v e DMSO a 3% em v/v em água) e solvente B (ácido fórmico a 0,2% em v/v e DMSO a 3% em EtOH) a uma taxa de fluxo de 3 µl/minuto.

[001493] A fim de gerar a biblioteca de íons para a análise das amostras, o software ProteinPilot foi executado nos arquivos wiff que foram gerados pelas execuções IDA. Esse banco de dados foi utilizado no software Peakview (ABSciex) para quantificar as proteínas em cada uma das amostras, utilizando 3 transições/peptídeos e 15 peptídeos/proteínas. A fim de maximizar o número de proteínas quantificadas, as amostras foram quantificadas em um banco de dados SWATH humano (Atlas) disponível publicamente com os mesmos parâmetros. Um peptídeo será considerado adequadamente medido quando a pontuação calculada pelo Peakview for superior a 1,5 e tivesse um FDR <1%. A quantificação de cada banco de dados foi combinada em uma quantificação final com o uso do nome da proteína de ambos os bancos de dados. Um fator de correção foi computado para cada amostra considerando-se o sinal total de cada proteína naquela amostra, quando comparado à média do sinal total para cada amostra.

[001494] A composição proteômica do fusossoma foi comparada à composição proteômica da célula parental. Uma composição proteômica semelhante entre fusossomas e células parentais foi observada quando >33% das proteínas identificadas estavam presentes no fusossoma, e dessas proteínas identificadas, o nível era > 25% do nível de proteína correspondente na célula parental, conforme mostrado na Tabela A a seguir.

	Total de Proteínas Identificadas	Proteínas compartilhadas (identificadas nas células parentais e fusossomas)	Proteínas compartilhadas com 25% de expressão parental em fusossomas	Fração de proteínas compartilhadas em proteínas totais
Fusossomas	1.926	1.487	957	0,333
Células	2.870			

EXEMPLO 89: QUANTIFICANDO UM NÍVEL DE PROTEÍNA ENDÓGENA OU SINTÉTICA POR FUSOSSOMA

[001495] Esse exemplo descreve a quantificação de uma carga de proteína endógena ou sintética em fusossomas. Os fusossomas podem, em alguns casos, compreender uma carga de proteína endógena ou sintética. O fusossoma ou célula parental descrita nesse Exemplo foi projetado para alterar a expressão de uma proteína endógena ou expressar uma carga sintética que medeia uma função celular terapêutica ou nova.

[001496] Fusossomas e células parentais que expressam GFP foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. A quantificação de GFP em fusossomas foi realizada com o uso de um kit ELISA disponível comercialmente (ab171581 Abcam Cambridge, Reino Unido) de acordo com as instruções do fabricante. A quantificação de fusossoma foi realizada por Análise de Rastreamento de Nanopartículas com o uso de um NanoSight NS300 (Malvern Instruments, Malvern, Worcestershire, Reino Unido). Os resultados são mostrados na tabela abaixo.

	Concentração (n°/ml)
Proteína GFP	$4,41 \times 10^{13}$
Fusossomas	$2,66 \times 10^{11}$
GFP:Fusossoma	165,8

[001497] Contempla-se que os fusossomas podem ter pelo menos 1, 2, 3, 4, 5, 10, 20, 50, 100 ou mais moléculas de agente de proteína por fusossoma. Em uma modalidade, os fusossomas terão 166 moléculas de agente de proteína por fusossoma.

EXEMPLO 90: MARCADORES DE MEDIÇÃO DE PROTEÍNAS EXOSSÔMICAS EM FUSOSSOMAS

[001498] Esse ensaio descreve a quantificação da razão entre proteínas que são conhecidas como marcadores específicos de exossomas.

[001499] Os fusossomas foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. Os exossomas foram preparados conforme descrito no presente documento para fusossomas pelo método dos Exemplos 68 e 87, com a exceção de que as células parentais não foram submetidas à transfecção com VSV-G ou GFP. A quantificação de proteínas por espectrometria de massa para fusossomas e exossomas foi realizada, conforme descrito no presente documento no Exemplo 36.

[001500] Os dados de quantificação de proteína resultante foram analisados para determinar os níveis de proteína e as proporções do marcador exossômico CD63 conhecido. As intensidades logarítmicas médias por grupo foram calculadas somando-se 1 aos valores de intensidade da espectrometria de massa, transformando por log10 e computando-se a média entre as repetições. Os resultados são mostrados na Figura 16.

EXEMPLO 91: MEDIÇÃO DE CALNEXINA EM FUSOSSOMAS

[001501] Esse ensaio descreve a quantificação do nível de calnexina (CNX) nos fusossomas e do nível relativo de CNX nos fusossomas em comparação às células parentais.

[001502] Os fusossomas e células parentais foram preparados conforme descrito no presente documento nos Exemplos 68 e 87. Calne-

xina e o total de proteína foram medidos com o uso de espectrometria de massa conduzida de acordo com o método do Exemplo 36. A intensidade do sinal de calnexina determinada para células parentais e fusossomas é mostrada na Figura 17.

[001503] Nas modalidades, com o uso desse ensaio, o teor fracionário médio (calculado conforme descrito no presente documento no Exemplo 36) de CNX nos fusossomas será $<2,43 \times 10^{-4}$.

[001504] Em uma modalidade, a diminuição da calnexina por total de proteína em ng/µg da célula parental para a preparação será superior a 88%.

EXEMPLO 92: RAZÃO ENTRE LIPÍDIOS E DNA EM FUSOSSOMAS

[001505] Esse exemplo descreve a quantificação da razão de lipídios para DNA em fusossomas em comparação às células parentais. Em uma modalidade, os fusossomas terão maior razão entre lipídios e DNA em comparação às células parentais. Os fusossomas foram preparados, conforme descrito anteriormente nos Exemplos 68 e 87.

[001506] Essa razão é definida como o teor de lipídios descrito no Exemplo 41, e o teor de ácido nucleico é determinado conforme descrito no Exemplo 42. Como mostrado na Figura 18 e na tabela abaixo, verificou-se que os fusossomas exibem uma razão entre lípido:DNA maior que as células parentais.

	[Lipídios] (µM)	[DNA] (ng/µl)	Lipídios:DNA (µmol/mg)
Fusossomas	115,6	29,5	3,92
Células	47,9	15,9	3,01

EXEMPLO 93: ANÁLISE DE MARCADORES DE SUPERFÍCIE NOS FUSOSSOMAS

[001507] Esse ensaio descreve a identificação de marcadores de superfície nos fusossomas.

[001508] Os fusossomas foram preparados, conforme descrito no

presente documento, nos Exemplos 68 e 87. A fosfatidilserina foi medida por espectrometria de massa, conforme descrito no presente documento nos Exemplos 68 e 87. A quantidade de fosfatidilserina em relação aos lipídios totais em fusossomas foi determinada como sendo 121% maior que a quantidade de fosfatidilserina em relação ao lipídio total nas células parentais, conforme mostrado na tabela abaixo.

	Fosfatidilserina (% molar)	Alteração percentual de fosfatidilserina
Fusossomas	14,6	121%
Células Parentais	6,6	

EXEMPLO 94: ANÁLISE DE PROTEÍNAS DO CAPSÍDEO VIRAL EM FUSOSSOMAS

[001509] Nesse exemplo, a composição da preparação da amostra foi analisada, e a razão entre proteínas que são derivadas de fontes de capsídeo viral foi avaliada.

[001510] Os fusossomas foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. A quantificação de proteína por espectrometria de massa para fusossomas foi realizada, conforme descrito no presente documento no Exemplo 36. O teor fracionário das proteínas do capsídeo viral foi calculado conforme descrito no presente documento no Exemplo 36, calculado sobre amostras de fusossoma e expresso como um percentual.

[001511] Com o uso dessa abordagem, descobriu-se que a amostra continha 0,05% de proteína capsidial viral, conforme mostrado na tabela abaixo. A única proteína capsidial viral detectada foi o complexo do capsídeo endógeno de coelho (RELIK) com ciclofilina A (PDB 2XGY|B).

	Intensidade Bruta de MS	Viral:Proteína Total (%)
Proteínas de capsídeo viral	$5,10 \times 10^5$	0,05
Total de Proteínas	$9,46 \times 10^8$	

EXEMPLO 95: QUANTIFICAÇÃO DE RAZÕES DE PROTEÍNA DE FUSOGÊNIO EM FUSOSSOMAS

[001512] Esse exemplo descreve a quantificação da razão entre proteína de fusogênio e o total de proteína ou outras proteínas de interesse em fusossomas. Outras proteínas de interesse podem incluir, porém sem limitação, EGFP, CD63, ARRDC1, GAPDH, Calnexina (CNX) e TSG101. Os fusossomas foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. A quantificação de proteína por espectrometria de massa para fusossomas foi realizada, conforme descrito no presente documento no Exemplo 36. A quantificação de todas as proteínas foi calculada, conforme descrito no presente documento no Exemplo 36, calculada em média sobre as amostras de fusossoma e expressa como uma fração.

[001513] Conforme mostrado na tabela abaixo, constatou-se que o fusogênio tem uma razão para EGFP de 156,9, uma razão para CD63 de 2.912,0, uma razão para ARRDC1 de 664,9, uma razão para GAPDH de 69,0, uma razão para CNX de 558,4 e um razão para TSG101 de 3.064,1.

Proteínas	Intensidade Bruta de MS	Razão de fusogênio:proteína(s)
VSV-G	$1,29 \times 10^8$	N/A
Total de Proteínas	$9,46 \times 10^8$	0,136
EGFP	$8,22 \times 10^5$	156,9
CD63	$4,43 \times 10^4$	2.912,0
ARRDC1	$1,94 \times 10^5$	664,9
GAPDH	$1,87 \times 10^6$	69,0
CNX	$2,31 \times 10^5$	558,4
TSG101	$4,21 \times 10^4$	3.064,1

EXEMPLO 96: QUANTIFICAÇÃO DE RAZÕES ENTRE PROTEÍNAS ENDÓGENAS E SINTÉTICAS EM FUSOSSOMAS

[001514] Esse exemplo descreve a quantificação de uma carga de proteína endógena ou sintética em relação à total de proteína ou outras proteínas de interesse em fusossomas. Outras proteínas de interesse podem incluir, porém sem limitação, VSV-G, CD63, ARRDC1, GAPDH, Calnexina (CNX) ou TSG101. Os fusossomas foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. A quantificação de proteína por espectrometria de massa para fusossomas foi realizada, conforme descrito no presente documento no Exemplo 36. A quantificação de todas as proteínas foi calculada, conforme descrito no presente documento no Exemplo 36, calculada em média sobre as amostras de fusossoma e expressa como uma fração.

[001515] Conforme mostrado na tabela abaixo, a carga de proteína sintética apresentou uma razão para VSV-G de $6,37 \times 10^{-3}$, uma razão para CD63 de 18,6, uma razão para ARRDC1 de 4,24, uma razão para GAPDH de 0,44, a razão para CNX de 3,56 e uma razão para TSG101 de 19,52.

Proteínas	Intensidade Bruta de MS	Razão Proteína de Carga: proteína(s)
EGFP	$8,22 \times 10^{5}$	N/A
Total de Proteínas	$9,46 \times 10^{8}$	$8,69 \times 10^{-4}$
VSV-G	$1,29 \times 10^{8}$	$6,37 \times 10^{-3}$
CD63	$4,43 \times 10^{4}$	18,6
ARRDC1	$1,94 \times 10^{5}$	4,24
GAPDH	$1,87 \times 10^{6}$	0,44
CNX	$2,31 \times 10^{5}$	3,56
TSG101	$4,21 \times 10^{4}$	19,52

EXEMPLO 97: COMPOSIÇÃO LIPÍDICA ENRIQUECIDA EM FUSOSSOMAS

[001516] Esse exemplo descreve a quantificação da composição lipídica de fusossomas, células parentais e exossomas. Contempla-se que a composição lipídica dos fusossomas pode ser enriquecida e/ou depletada para lipídios específicos em relação às células das quais são derivados. A composição lipídica afeta parâmetros biofísicos importantes de fusossomas e células, tais como tamanho, interações eletrostáticas e comportamento coloidal.

[001517] A composição lipídica foi medida conforme descrito em Exemplos 68 e 87. Os fusossomas foram preparados conforme descrito no presente documento por transfecção transitória de VSV-G e GFP em pratos de 10 cm, seguida por filtração e ultracentrifugação do meio condicionado 48 horas após a transfecção para obter fusossomas. As células submetidas à transfecção foram colhidas em paralelo com o meio condicionado e submetidas à análise. Os exossomas foram preparados conforme descrito no presente documento para fusossomas com a exceção de que as células parentais não foram submetidas à transfecção com VSV-G ou GFP.

[001518] A composição lipídica para fusossomas, exossomas e células parentais é mostrada nas Figuras 19A a 19B. Em comparação às células parentais, os fusossomas foram enriquecidos em éster colesterílico, colesterol livre, liso-fosfatidiletanolamina ligada a éter, liso-fosfatidilserina, fosfatidato, fosfatidiletanolamina ligada a éter, fosfatidilserina e esfingomielina. Em comparação às células parentais, fusossomas estão depletados para ceramida, cardiolipina, liso-fosfatidilcolina, liso-fosfatidiletanolamina, liso-fosfatidilglicerol, liso-fosfatidilinositol, ligados a éter fosfatidilcolina, fosfatidiletanolamina, fosfatidilglicerol, fosfatidilinositol e triacilglicerol. Em comparação aos exossomas, os fusossomas foram enriquecidos com éster colesterílico, ceramida,

diacilglicerol, liso-fosfatidato e fosfatidiletanolamina, triacilglicerol. Em comparação aos exossomas, os fusossomas são depletados para colesterol livre, hexosil ceramida, liso-fosfatidilcolina, liso-fosfatidilcolina ligada a éter, liso-fosfatidiletanolamina, liso-fosfatidiletanolamina ligada a éter e liso-fosfatidilserina.

EXEMPLO 98: MEDIÇÃO DO CONTEÚDO PROTEÔMICO ESPECÍFICO DO COMPARTIMENTO DE FUSOSSOMAS

[001519] Esse Exemplo descreve a quantificação da razão entre proteínas que são conhecidas por serem derivadas de compartimentos celulares específicos em fusossomas, células parentais de fusossomas e exossomas.

[001520] Os fusossomas e células parentais foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. Os exossomas foram preparados conforme descrito no presente documento para fusossomas pelo método dos Exemplos 68 e 87, com a exceção de que as células parentais não foram submetidas à transfecção com VSV-G ou GFP. A quantificação de proteínas por espectrometria de massa para fusossomas e exossomas foi realizada, conforme descrito no presente documento no Exemplo 36. Os dados de quantificação de proteínas resultante foram analisados para determinar os níveis de proteínas e proporções de proteínas exossômicas, retículo endoplasmático, ribossoma, nuclear e proteínas mitocondriais conhecidas, conforme anotado pelos termos de anotação do Compartimento Celular da Ontologia Gene (exossoma: GO:0070062, retículo endoplasmático: GO:0005783, ribossoma: GO:0005840, GO:0022625, GO:0022626, GO:0022627, GO:0044391, GO:0042788, GO:0000313) com código de evidência IDA (inferido por ensaio direto). A fração de proteínas específicas do compartimento em relação ao total de proteína em cada amostra foi determinada para amostras de fusossoma, amostras de exossoma e células parentais.

[001521] Conforme mostrado na Figura 20, verificou-se que os fusossomas estão depletados com a proteína do retículo endoplasmático em comparação às células parentais e exossomas. Constatou-se, também, que os fusossomas foram depletados de à proteína exossomal em comparação aos exossomas. Os fusossomas foram depletados de proteína mitocondrial em comparação às células parentais. Fusossomas foram enriquecidos em proteína nuclear em comparação a células parentais. Os fusossomas foram enriquecidos em proteínas ribossômicas em comparação a células parentais e exossomas.

EXEMPLO 99: MEDIÇÃO DO CONTEÚDO DE TSG101 E ARRDC1 EM FUSOSSOMAS

[001522] Esse exemplo descreve a quantificação da proporção de proteínas que são conhecidas por serem importantes na liberação de fusossoma das células.

[001523] Os fusossomas e células parentais foram preparados conforme descrito no presente documento pelo método dos Exemplos 68 e 87. Os exossomas foram preparados conforme descrito no presente documento para fusossomas pelo método dos Exemplos 68 e 87, com a exceção de que as células parentais não foram submetidas à transfecção com VSV-G ou GFP. A quantificação de proteínas por espectrometria de massa para fusossomas e exossomas foi realizada, conforme descrito no presente documento no Exemplo 36. Os dados de quantificação de proteína resultantes foram analisados para determinar os níveis de proteína e as proporções da proteína TSG101 e ARRDC1. As intensidades logarítmicas médias por grupo foram calculadas somando-se 1 aos valores de intensidade da espectrometria de massa, transformando por log10 e computando-se a média entre as repetições. A porcentagem do teor total da proteína de TSG101 ou ARRDC1 em fusossomas em relação aos exossomas ou células parentais foi determinada como a intensidade logarítmica média de

TSG101 ou ARRDC1 para cada amostra, dividida pela soma das intensidades de todas as proteínas na mesma amostra, em média sobre as repetições e expresso como uma porcentagem.

[001524] Conforme mostrado na Figura 21, verificou-se que ARRDC1 está presente em níveis maiores como uma porcentagem do teor total da proteína em fusossomas do que em células parentais ou exossomas. O nível de ARRDC1 como um percentual do teor total da proteína era de pelo menos 0,02% em fusossomas. Constatou-se que TSG101 está presente em níveis maiores como uma porcentagem do teor total da proteína em fusossomas do que em células parentais ou exossomas. O nível de TSG101 como um percentual do teor total da proteína era de pelo menos 0,004% em fusossomas.

EXEMPLO 100: MEDIÇÃO DA INATIVAÇÃO SÉRICA DE FUSOSSOMAS APÓS MÚLTIPLAS ADMINISTRAÇÕES

[001525] Esse Exemplo descreve a quantificação de inativação do soro de fusossomas com o uso de um ensaio de entrega *in vitro* seguindo múltiplas administrações do fusossoma. Contempla-se que um fusossoma modificado, por exemplo, modificado por um método descrito no presente documento, pode ter uma inativação de soro reduzida (por exemplo, reduzido em comparação à administração de um fusossoma não modificado) seguindo múltiplas (por exemplo, mais de um, por exemplo, 2 ou mais), administrações do fusossoma modificado. Em alguns casos, um fusossoma descrito no presente documento não será inativado por soro após múltiplas administrações.

[001526] Uma medida de imunogenicidade para fusossomas é a inativação sérica. Em uma modalidade, as injeções repetidas de um fusossoma podem causar o desenvolvimento de anticorpos antifusossomas, por exemplo, anticorpos que reconhecem fusossomas. Em uma modalidade, anticorpos que reconhecem fusossomas podem se ligar de maneira que limitem a atividade ou longevidade do fusossoma

e medeiem degradação complementar.

[001527] Nesse Exemplo, a inativação do soro é examinada após uma ou mais administrações de fusossomas. Os fusossomas são produzidos por qualquer um dentre os Exemplos anteriores. Nesse exemplo, os fusossomas são gerados a partir de: células HEK293 modificadas com uma expressão mediada por lentivírus de HLA-G (doravante HEK293-HLA-G) e HEK293 modificadas com uma expressão mediada por lentivírus de um vetor vazio (doravante, HEK293). Em algumas modalidades, fusossomas são derivados de células que expressam outras proteínas imunorreguladoras.

[001528] O soro é retirado das diferentes coortes: camundongos injetados sistemicamente e/ou localmente com 1, 2, 3, 5, 10 injeções de veículo (grupo fusossoma naïve), fusossomas HEK293-HLA-G ou fusossomas HEK293. Os soros são coletados de camundongos coletando-se sangue total fresco e permitindo que coagule completamente por várias horas. Os coágulos são peletizados por centrifugação, e os sobrenadantes do soro são removidos. Um controle negativo é soro de camundongo inativado por calor. As amostras de controle negativo são aquecidas a 56 graus Celsius durante 1 hora. O soro pode ser congelado em alíquotas.

[001529] Os fusossomas são testados para a dose em que 50% das células em uma população receptora recebem a carga útil nos fusossomas. Os fusossomas podem ser produzidos por meio de qualquer um dos outros exemplos descritos no presente documento e podem conter qualquer uma das cargas úteis no presente documento descritas. Muitos métodos para testar a entrega de fusossoma de uma carga útil para células receptoras também são descritos no presente documento. Nesse exemplo particular, a carga útil é a proteína Cre e as células recipientes são célula RPMI8226 que expressa de maneira estável o cassete “LoxP-GFP-stop-LoxP-RFP” sob um promotor de CMV,

que mediante recombinação por Cre, comuta da expressão GFP para a expressão RFP, indicando a fusão Cre, como um marcador de entrega. A dose identificada na qual 50% das células receptoras são RFP-positivas é usada para outros experimentos. Em outras modalidades, a dose identificada na qual 50% das células recipientes recebem a carga útil são usadas para experimentos adicionais.

[001530] A fim de avaliar a inativação do soro de fusossomas, fusossomas são diluídos a 1:5 em soro normal ou inativado por calor (ou meio que contém 10% de FBS inativado por calor como o controle não sérico), e a mistura é incubada a 37 °C durante 1 hora. Após a incubação, o meio é adicionado à reação para uma diluição adicional de 1:5 e, em seguida, diluído em série duas vezes na razão de 1:10. Seguindo essa etapa, os fusossomas devem estar presentes na dose identificada anteriormente na qual 50% das células receptores receberam a carga útil (por exemplo, são RFP-positivas). Contempla-se que a dose identificada na qual 50% das células receptoras recebem a carga útil pode ser semelhante entre fusossomas.

[001531] Os fusossomas que foram expostas a soros são, então, incubados com células recipientes. A porcentagem de células que recebem a carga útil, e são, portanto, RFP-positivas, é calculada. A porcentagem de células que recebem a carga útil pode não ser diferentes entre amostras de fusossoma que foram incubadas com soro e o soro inativado por calor de camundongos tratados com fusossomas HEK293-HLA-G, o que indica que não há inativação do soro de fusossomas ou uma resposta imune adaptativa. A porcentagem de células que recebem a carga útil pode não ser diferente entre as amostras de fusossomas que foram incubadas de camundongos tratados 1, 2, 3, 5 ou 10 vezes com fusossomas HEK293-HLA-G, o que indicaria que não houve inativação sérica de fusossomas ou uma resposta imunológica adaptativa. Em alguns exemplos, a porcentagem de células que rece-

bem a carga útil não é diferente entre amostras de fusossoma que foram incubadas com soro dos camundongos tratados com veículo e de camundongos tratados com fusossomas HEK293-HLA-G, indicando que não haja inativação do soro de fusossomas ou uma resposta imune adaptativa. Em alguns casos, a porcentagem de células que recebem a carga útil é menor para fusossomas derivados de HEK293 do que para fusossomas HEK293-HLA-G, indicando que não há inativação sérica de fusossomas HEK293-HLA-G ou uma resposta imunológica adaptativa.

EXEMPLO 101: MEDIÇÃO DO ALVEJAMENTO DO COMPLEMENTO DE FUSOSSOMAS

[001532] Esse Exemplo descreve a quantificação de atividade de complemento contra fusossomas com o uso de um ensaio *in vitro*. Contempla-se que um fusossoma modificado descrito no presente documento pode induzir atividade de complemento reduzida em comparação a um fusossoma não modificado correspondente.

[001533] Nesse Exemplo, o soro de um camundongo é avaliado quanto à atividade do complemento em relação a um fusossoma. O exemplo mede o nível de complemento C3a, que é um nó central em todas as vias do complemento. Notavelmente, os métodos descritos no presente documento podem ser igualmente aplicáveis a seres humanos, ratos, macacos com otimização do protocolo.

[001534] Nesse Exemplo, os fusossomas são produzidos por qualquer um dos Exemplos anteriores. Os fusossomas são gerados a partir de células HEK293 modificadas com uma expressão mediada por lentivírus de uma proteína regulatória do complemento DAF (fusossomas HEK293-DAF) ou células HEK 293 que não expressam uma proteína regulatória do complemento (fusossomas HEK293). Em outras modalidades, outras proteínas regulatórias do complemento podem ser utilizadas, tais como proteínas que se ligam ao fator acelerador de decai-

mento (DAF, CD55), por exemplo, proteína-1 semelhante ao fator H (FH) (FHL-1), por exemplo, proteína de ligação a C4b (C4BP), por exemplo, receptor de complemento 1 (CD35), por exemplo, proteína cofator de membrana (MCP, CD46), por exemplo, Profectina (CD59), por exemplo, proteínas que inibem as enzimas CD/C5 convertase da via clássica e alternativa do complemento, por exemplo, proteínas que regulam a montagem do MAC.

[001535] O soro é recuperado dos camundongos naïve, dos camundongos que recebem os fusossomas HEK293-DAF ou dos camundongos que recebem fusossomas HEK293. Os soros são coletados de camundongos coletando-se sangue total fresco e permitindo-se que coagule completamente por várias horas. Os coágulos são peletizados por centrifugação, e os sobrenadantes do soro são removidos. Um controle negativo é soro de camundongo inativado por calor. As amostras de controle negativo são aquecidas a 56 graus Celsius durante 1 hora. O soro pode ser congelado em alíquotas.

[001536] Os diferentes fusossomas são testados para a dose em que 50% das células em uma população receptora recebem a carga útil nos fusossomas. Os fusossomas podem ser produzidos por meio de qualquer um dos outros exemplos descritos no presente documento e podem conter qualquer uma das cargas úteis no presente documento descritas. Muitos métodos para testar a entrega de fusossoma de uma carga útil para células receptoras também são descritos no presente documento. Nesse exemplo particular, a carga útil é a proteína Cre e as células recipientes são célula RPMI8226 que expressa de maneira estável o cassete “LoxP-GFP-stop-LoxP-RFP” sob um promotor de CMV, que mediante recombinação por Cre, comuta da expressão GFP para a expressão RFP, indicando a fusão Cre, como um marcador de entrega. A dose identificada na qual 50% das células receptoras são RFP-positivas é usada para outros experimentos. Em outras modali-

dades, a dose identificada na qual 50% das células recipientes recebem a carga útil são usadas para experimentos adicionais. Em modalidades preferidas, a dose identificada na qual 50% das células receptoras recebem a carga útil é semelhante entre os fusossomas.

[001537] Diluições duplas dos fusossomas que começam na dose de fusossomas na qual 50% das células recipientes recebem a carga útil em solução salina tamponada com fosfato (PBS, pH 7,4) são misturadas com uma diluição 1:10 dos soros dos camundongos tratados com os mesmos fusossomas ou camundongos naïve (volume de ensaio, 20 µl) e incubados durante 1 hora a 37 °C. As amostras são posteriormente diluídas 1:500 e utilizadas em um ensaio de ensaio de imunoadsorção ligado à enzima (ELISA) específico para C3a. O ELISA é um produto do kit LS-F4210 de complemento de camundongo ELISA para C3a vendido pela LifeSpan BioSciences Inc, que mede a concentração de C3a em uma amostra. A dose de fusossomas na qual 200 pg/ml de C3a está presente é comparada ao longo dos soros isolados de camundongos.

[001538] Em alguns casos, a dose de fusossomas em que 200 pg/ml de C3a está presente é maior para fusossomas HEK293-DAF incubados com soros de camundongos HEK-293 DAF do que para fusossomas HEK293 incubados com soros de camundongos HEK293, indicando que a atividade do complemento que alveja fusossomas é maior em camundongos tratados com fusossomas HEK293 do que fusossomas HEK293-DAF. Em alguns casos, a dose de fusossomas em que 200 pg/ml de C3a está presente é maior para fusossomas HEK293-DAF incubados com soros de camundongos ingênuos do que para fusossomas HEK293 incubados com soros de camundongos ingênuos, indicando que a atividade do complemento que alveja fusossomas é maior em camundongos tratado com fusossomas HEK293 do que fusossomas HEK293-DAF.

EXEMPLO 102: AVALIAÇÃO DA ESPECIFICIDADE DA EXPRESSÃO DO TRANSGENE COM O USO DE PROMOTORES ESPECÍFICOS DE TECIDO E SILENCIAMENTO GÊNICO MEDIADO POR miRNA.

[001539] Esse Exemplo descreve a quantificação de um agente exógeno na linhagem celular de hepatoma humano alvo (HepG2) e em comparação com linhagens de células não alvo (não hepáticas). As linhagens celulares foram submetidas à transdução com lentivírus (LV) que contém TCSREs positivos (por exemplo, promotor específico de tecido) ou uma combinação de TCSREs e NTCSREs positivos (por exemplo, silenciamento de gene mediado por miRNA com uma sequência de reconhecimento de miRNA específica de tecido). Linhagens de células-alvo e não alvo foram submetidas à transdução com partículas lentivirais geradas que contêm os elementos regulatórios positivos e negativos e o efeito da expressão do transgene nas linhagens de células foi avaliado.

A. EFEITO DA REGULAÇÃO GÊNICA MEDIADA POR miRNA NA ESPECIFICIDADE DA EXPRESSÃO DO TRANSGENE.

[001540] Além dos hepatócitos, grandes populações celulares que alinham sinusoides hepáticos incluem células endoteriais e células de Kupffer (macrófagos derivados de linhagem hematopoiética), que expressa mir-126-3p e mir-142-3p, respectivamente. Esses miRNAs não são expressos substancialmente em hepatócitos. Vetores lentivirais foram construídos para conter um cassete de expressão de proteína fluorescente verde intensificada (eGFP) sob o controle da fosfoglicerato quinase de promotor constitutivamente ativo (hPGK, TCSRE positivo; consultar, por exemplo, Tabela 3, com ou sem quatro cópias em tandem de cada uma das sequências complementares a mir-142-3p (por exemplo, Tabela 4) e miR-126-3p (por exemplo, Tabela 4) como um NTCSRE. Construtos do vetor lentiviral (LV) com o NTCSRE são

denominados de hPGK-eGFP+miRT, e os construtos sem o NTCSRE são denominados de hPGK-eGFP.

[001541] Os lentivírus (LVs) gerados a partir desses vetores hPGK-eGFP+miRT e hPGK-eGFP, respectivamente, foram usados para transduzir uma linhagem de células-alvo de hepatoma humano (HepG2) ou linhagem de células renais embrionárias humanas (293LX), linhagem de células T humanas de origem hematopoiéticas (Molt4.8) e linhagem celular endotelial derivada de cérebro de camundongo (bEND.3). Sete dias após a transdução, a expressão de GFP foi medida por citometria de fluxo.

[001542] Conforme mostrado na **Figura 22A**, 18 a 30% de todos os tipos de células submetidas à transdução com hPGK-eGFP LV expressaram GFP. Após a transdução de LVs que contêm sequências mirT (hPGK-eGFP+ miRT) em células não alvo, apenas 0,6% de expressão de GFP foi observada em células Molt4.8 (expressa mir-142-3p) e nenhuma expressão foi observada em células bEND.3 (expresso mir-126-3p). A redução leve da expressão de GFP foi observada em células 293LX, que podem ter um nível muito baixo de expressão de um ou de ambos os miRNAs. Nenhum efeito na expressão de GFP foi observado em células-alvo HepG2 que haviam sido submetidas à transdução com LVs que contêm sequências mirT (hPGK-eGFP+ miRT). Esses resultados demonstraram que a incorporação de sequências de miRT em vetores lentivirais resultou em uma redução de pelo menos 50 vezes na expressão do transgene em células das linhagens hematopoiéticas e endoteliais ao mesmo tempo que manteve a expressão robusta em células hepáticas.

B. EFEITO COMBINADO DA REGULAÇÃO GÊNICA MEDIADA POR miRNA E PROMOTORES ESPECÍFICOS DE TECIDO NA EXPRESSÃO DO TRANSGENE.

[001543] Vetores lentivirais adicionais (LVs) foram gerados substan-

cialmente conforme descrito acima, porém ou com um eGFP ou com um transgene que codifica a enzima fenilalanina amônia-liase (*PAL*) (consultar, por exemplo, Tabela 5) com uma tag de sinalização N-terminal, sob o controle de um promotor humano hepatócito-específico (ApoE.HCR-hAAT (hApoE) (por exemplo, Tabela 3) como um promotor regulatório tecido-específico como um TCSRE positivo ou um promotor constutivamente ativo de Vírus Formador de Foco no Baço (SFFV) (por exemplo, Tabela 3). A expressão de PAL em células do fígado usando os construtos de ácido nucleico fornecidos é representativa da expressão de um agente exógeno desejado em células-alvo. Por exemplo, a deficiência endógena de PAH em células hepáticas humanas pode resultar no acúmulo tóxico de Phe no sangue causando fenilcetonúria (PKU), uma afecção clínica caracterizada por distúrbios neurológicos graves e crescimento atrofiado. Em alguns aspectos, a administração precoce de PAL a pacientes com PKU demonstrou diminuir com sucesso os níveis de Phe no sangue e aliviar os sintomas.

[001544] Conforme mostrado na **Figura 22B**, a transdução com LVs que contêm hPGK-eGFP ou LVs que contêm sequências de mirT e GFP sob o controle do promotor hepatócito específico (hApoE-eGFP+miRT) resultou na repressão >250 vezes da expressão de GFP em células 293LX, sem efeito substancial em células HepG2, conforme medido por citometria de fluxo.

[001545] As células HepG2 e 293LX foram submetidas à transdução com LVs que contêm o transgene PAL sob o controle do promotor SFFV (SFFV-PAL), ou LVs que contêm transgene PAL juntamente com sequências mirT sob o controle do promotor hApoE (hApoE-PAL+ miRT). PAL catalisa a conversão de fenilalanina (Phe) em amônia e ácido cinâmico e tem sido usado na terapia de reposição enzimática em pacientes com deficiência congênita de fenilalanina hidroxilase (PAH). A especificidade da expressão do transgene PAL foi medida

pela redução nos níveis de Phe no sobrenadante de cultura (SN) em relação ao meio fresco. Conforme mostrado na **Figura 22C**, a expressão de um promotor constitutivo (SFFV) resultou em redução substancial nos níveis de Phe em SN coletados de ambos os tipos de células. No entanto, o construto hApoE-PAL+miRT levou a uma redução substancial de Phe apenas em células HepG2; os níveis de Phe em SN coletados a partir de células 293LX submetidas à transdução com os LVs hApoE-PAL+miRT foram indistinguíveis dos controles não transduzidos. Esse resultado é consistente com a alta expressão do transgene PAL exemplificativo em células-alvo HepG2 quando transduzido com construtos LV que contém um TSCRE positivo e um NTSCRE, porém não em células 293LX não alvo.

[001546] Vetores adicionais foram produzidos usando: i) o promotor ApoE.HCR-hAAT, que é um promotor quimérico formado pela região de controle hepático do gene da Apolipoproteína E humana e o promotor da alfa-antitripsina humana (Miao et al, Mol Ther, 2000; PMID: 10933977), ii) o promotor da Transtirretina Reforçada (ET), um promotor sintético gerado pela montagem aleatória de locais de ligação de fator de transcrição específico de hepatócitos ao promotor de transtirretina (Vigna et al, MolTher, 2005; PMID: 15851015 e Brown et al, Blood, 2007; PMID: 17726165), e iii) o promotor TBG, que é um promotor híbrido com base na globulina de ligação do hormônio tireoidiano humano e do intensificador α1-Microglobulina/Bikunina (Yan et al, Gene, 2012; PMID: 22820390). Esses três promotores foram clonados em um plasmídeo de estrutura lentiviral (pSF) a montante do gene eGFP. Esses construtos, juntamente com um cassete SFFV ubíquo, foram empacotados em VSVg-LVs pseudotipados e transduzidos em HepG2 (uma linhagem celular de hepatócitos humanos), hepatócitos humanos primários e SupT1 (uma linhagem de células T humanas, para avaliar a especificidade). As análises de imagem das células submetidas à

transdução demonstraram que os três promotores específicos dos hepatócitos produziram uma expressão potente e específica em ambas as linhagens celulares dos hepatócitos, porém não nas células não hepatócitos (consultar a tabela a seguir). Embora ApoE e TBG não tenham medido nenhuma expressão detectável nas células T, um dos promotores, ET, ainda produziu baixos níveis de GFP na linhagem de células T. A fim de aumentar ainda mais a especificidade dos hepatócitos, foram incluídos 4x locais alvos em tandem para miR142-3p (um miRNA expresso em células de origem hematopoiética) (Brown et al, Blood, 2007; PMID: 17726165), bem como para miR126-3p (um miRNA expresso na maioria das células endoteliais) (Chiriaco et al, Mol-Ther, 2014; PMID: 24869932) a jusante da sequência expressa. A análise de imagem de linhagens de células hepatocíticas e não hepatocíticas submetidas à transdução com esses novos construtos demonstraram claramente que a adição de locais alvo de miRNA suprime fortemente a expressão do transgene em células imunes, causando a supressão completa da expressão de GFP em células não hepatócitos, ao passo que mantém expressão significativa em hepatócitos humanos primários e imortalizados.

	Hepatócitos humanos primários	HepG2	SupT1
pSF-SFFV-eGFP	++++	++++	+++++
pSF-SFFV-eGFP+1xmiRT	+++	++++	++
pSF-SFFV-eGFP+4xmiRT	+++	++	+/-
pSF-ApoE-eGFP	++	++	-
pSF-ET-eGFP	+++	++++	+/-
pSF-TBG-eGFP	++	+++	-
pSF-ApoE-eGFP+4xmiRT	++	+	-
pSF-ET-eGFP+4xmiRT	++	++	-
pSF-TBG-eGFP+4xmiRT	+	++	-

C. CONCLUSÃO

[001547] Juntos, esses dados apoiam a constatação de que o uso de um promotor específico de tecido em combinação com locais-alvo de miRNA pode conferir especificidade substancial à expressão do transgene.

REIVINDICAÇÕES

1. Fusossoma, caracterizado pelo fato de que compreende:

a) uma bicamada lipídica que compreende um fusogênio; e

b) um ácido nucleico que compreende:

(i) um gene de carga útil que codifica um agente exógeno; e

(ii) um elemento regulatório específico de célula do fígado positivo operativamente ligado ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula do fígado em relação a um fusossoma de outra forma similar sem o elemento regulatório específico de célula do fígado positivo.

2. Fusossoma, de acordo com a reivindicação 1, caracterizado pelo fato de que o ácido nucleico compreende ainda um elemento regulatório específico de célula não alvo (NTCSRE) operacionalmente ligado ao gene de carga útil, em que o NTCSRE diminui a expressão do gene de carga útil em uma célula não hepática em relação a um fusossoma de outra forma similar sem o NTCSRE.

3. Fusossoma, caracterizado pelo fato de que compreende:

a) uma bicamada lipídica que compreende um fusogênio; e

b) um ácido nucleico que compreende:

(i) um gene de carga útil que codifica um agente exógeno; e

(ii) um promotor operativamente ligado ao gene de carga útil, em que o promotor é escolhido dentre um promotor Apoa2, Cyp3a4, LP1B, MIR122, hemopexina, SERPINA1 ou HLP.

4. Fusossoma, caracterizado pelo fato de que compreende:

a) uma bicamada lipídica que compreende um fusogênio; e

b) um ácido nucleico que compreende:

(i) um gene de carga útil que codifica um agente exógeno; e

(ii) um elemento regulatório específico de célula não alvo (NTCSRE), operacionalmente ligado ao gene de carga útil, em que:

o NTCSRE é uma sequência de reconhecimento de miRNA específico de célula não hepática, um local de reconhecimento de protease específico de célula não hepática, um local de ubiquitina ligase específico de célula não hepática, um local de repressão transcricional específico de célula não hepática ou um local de repressão epigenética específico de célula não hepática, e o NTCSRE diminui a expressão do gene de carga útil em uma célula ou tecido não hepático em relação a um fusossoma similar sem o NTCSRE.

5. Fusossoma, de acordo com a reivindicação 4, caracterizado pelo fato de que o ácido nucleico compreende ainda um elemento regulatório específico de célula do fígado positivo operativamente ligado ao gene de carga útil, em que o elemento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula do fígado em relação a um fusossoma de outra forma similar sem o elemento regulatório específico de célula do fígado positivo.

6. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que o fusossoma compreende ainda um ou ambos dentre:

(i) uma primeira proteína imunossupressora exógena ou superexpressa na bicamada lipídica; ou

(ii) uma proteína imunoestimuladora que está ausente ou presente em níveis reduzidos, opcionalmente em que o nível reduzido é reduzido em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, em comparação com um fusossoma gerado a partir de uma célula de origem não modificada de outra forma similar.

7. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 6, caracterizado pelo fato de que o gene de carga útil é um gene que trata uma deficiência genética.

8. Fusossoma, caracterizado pelo fato de que compreende:

a) uma bicamada lipídica que compreende um fusogênio;

b) um ácido nucleico que compreende um gene de carga útil que codifica um agente exógeno para o tratamento de uma deficiência genética; e

c) um ou ambos dentre:

(i) uma primeira proteína imunossupressora exógena ou superexpressa na bicamada lipídica; ou

(ii) uma proteína imunoestimuladora que está ausente ou presente em níveis reduzidos (por exemplo, reduzida em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%) em comparação com um fusossoma gerado a partir de uma célula de origem não modificada, de outra forma, similar.

9. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 8, caracterizado pelo fato de que compreende (i) e (ii).

10. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 9, caracterizado pelo fato de que compreende (i) e compreende ainda uma segunda proteína imunossupressora exógena ou superexpressa na bicamada lipídica.

11. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 10, caracterizado pelo fato de que compreende (ii) e compreende ainda uma segunda proteína imunoestimuladora que está ausente ou presente em níveis reduzidos, opcionalmente em que o nível reduzido é reduzido em pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, em comparação com um fusossoma gerado a partir de uma célula de origem não modificada, de outra forma, similar.

12. Fusossoma, de acordo com qualquer uma das reivindicações 8 a 11, caracterizado pelo fato de que o ácido nucleico compreende ainda um elemento regulatório específico de célula do fígado positivo operativamente ligado ao gene de carga útil, em que o ele-

mento regulatório específico de célula do fígado positivo aumenta a expressão do gene de carga útil em uma célula do fígado em relação a um fusossoma de outra forma similar sem o elemento regulatório específico de célula do fígado positivo.

13. Fusossoma, de acordo com qualquer uma das reivindicações 8 a 12, caracterizado pelo fato de que o ácido nucleico compreende ainda um elemento regulatório específico de célula não alvo (NTCSRE) operacionalmente ligado ao gene de carga útil, em que o NTCSRE diminui a expressão do gene de carga útil em uma célula não hepática ou tecido em relação a um fusossoma de outra forma similar sem o NTCSRE.

14. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 13, caracterizado pelo fato de que, quando administrado a um sujeito, um ou mais dentre:

i) o fusossoma não produz uma resposta de anticorpo detectável, ou anticorpos contra o fusossoma estão presentes em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima de um nível precedente;

ii) o fusossoma não produz uma resposta imunológica celular detectável, ou uma resposta imunológica celular contra o fusossoma está presente em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima do nível precedente;

iii) o fusossoma não produz uma resposta imunológica inata detectável, ou a resposta imunológica inata contra o fusossoma está presente em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima do nível precedente;

iv) menos de 10%, 5%, 4%, 3%, 2% ou 1% dos fusossomas são inativados pelo soro;

v) uma célula-alvo que recebeu o agente exógeno do fusossoma não produz uma resposta de anticorpo detectável, ou anti-

corpos contra a célula-alvo estão presentes em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima do nível precedente; ou

vi) uma célula-alvo que recebeu o agente exógeno do fusossoma não produz uma resposta imunológica celular detectável, ou uma resposta celular contra a célula-alvo está presente em um nível inferior a 10%, 5%, 4%, 3%, 2% ou 1% acima do nível precedente.

15. Fusossoma, de acordo com a reivindicação 14, caracterizado pelo fato de que o nível precedente é o nível correspondente no mesmo sujeito antes da administração do fusossoma.

16. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 15, caracterizado pelo fato de que a proteína imunossupressora é uma proteína regulatória do complemento ou CD47.

17. Fusossoma, de acordo com qualquer uma das reivindicações 6 a 16, caracterizado pelo fato de que a proteína imunoestimuladora é uma proteína MHC I ou MHC II.

18. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 17, caracterizado pelo fato de que o gene de carga útil é selecionado dentre OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAH, PAL, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LNBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTD, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1, LDLR, ACAD8, ACADSB, ACAT1, ACSF3, ASPA, AUH, DNAJC19, ETHE1, FBP1, FTCD, GSS, HIBCH, IDH2, L2HGDH, MLYCD, OPA3, OPLAH, OXCT1, POLG, PPM1K,

SERAC1, SLC25A1, SUCLA2, SUCLG1, TAZ, AGK, CLPB, TMEM70, ALDH18A1, OAT, CA5A, GLUD1, GLUL, UMPS, SLC22A5, CPT1A, HADHA, HADH, SLC52A1, SLC52A2, SLC52A3, HADHB, GYS2, PYGL, SLC2A2, ALG1, ALG2, ALG3, ALG6, ALG8, ALG9, ALG11, ALG12, ALG13, ATP6V0A2, B3GLCT, CHST14, COG1, COG2, COG4, COG5, COG6, COG7, COG8, DOLK, DHDDS, DPAGT1, DPM1, DPM2, DPM3, G6PC3, GFPT1, GMPPA, GMPPB, MAGT1, MAN1B1, MGAT2, MOGS, MPDU1, MPI, NGLY1, PGM1, PGM3, RFT1, SEC23B, SLC35A1, SLC35A2, SLC35C1, SSR4, SRD5A3, TMEM165, TRIP11, TUSC3, ALG14, B4GALT1, DDOST, NUS1, RPN2, SEC23A, SLC35A3, ST3GAL3, STT3A, STT3B, AGA, ARSA, ARSB, ASAH1, ATP13A2, CLN3, CLN5, CLN6, CLN8, CTNS, CTSA, CTSD, CTSF, CTSK, DNAJC5, FUCA1, GAA, GALC, GALNS, GLA, GLB1, GM2A, GNPTAB, GNPTG, GNS, GRN, GUSB, HEXA, HEXB, HGSNAT, HYAL1, IDS, IDUA, KCTD7, LAMP2, MAN2B1, MANBA, MCOLN1, MFSD8, NAGA, NAGLU, NEU1, NPC1, NPC2, SGSH, PPT1, PSAP, SLC17A5, SMPD1, SUMF1, TPP1, AHCY, GNMT, MAT1A, GCH1, PCBD1, PTS, QDPR, SPR, DNAJC12, ALDH4A1, PRODH, HPD, GBA, HGD, AMN, CD320, CUBN, GIF, TCN1, TCN2, PREPL, PHGDH, PSAT1, PSPH, AMT, GCSH, GLDC, LIAS, NFU1, SLC6A9, SLC2A1, ATP7A, AP1S1, CP, SLC33A1, PEX7, PHYH, AGPS, GNPAT, ABCD1, ACOX1, PEX1, PEX2, PEX3, PEX5, PEX6, PEX10, PEX12, PEX13, PEX14, PEX16, PEX19, PEX26, AMACR, ADA, ADSL, AMPD1, GPHN, MOCOS, MOCS1, PNP, XDH, SUOX, OGDH, SLC25A19, DHTKD1, SLC13A5, FH, DLAT, MPC1, PDHA1, PDHB, PDHX, PDP1, ABCC2, SLCO1B1, SLCO1B3, HFE2, ADAMTS13, PYGM, COL1A2, TNFRSF11B, TSC1, TSC2, DHCR7, PGK1, VLDLR, KYNU, F5, C3, COL4A1, CFH, SLC12A2, GK, SFTPC, CRTAP, P3H1, COL7A1, PKLR, TALDO1, TF, EPCAM, VHL, GC, SERPINA1, ABCC6, F8, F9, ApoB, PCSK9, LDLRAP1, ABCG5, ABCG8, LCAT, SPINK5 e GNE.

19. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 18, caracterizado pelo fato de que o gene de carga útil codifica um agente exógeno que compreende a sequência estabelecida em qualquer uma das SEQ ID NOS: 161 a 518, um fragmento funcional do mesmo ou uma variante funcional do mesmo que compreende uma sequência de aminoácidos tendo pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com uma sequência de aminoácidos estabelecida em qualquer uma das SEQ ID NOS: 161 a 518.

20. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 19, caracterizado pelo fato de que um ou mais dentre:

i) o fusossoma se funde a uma taxa mais alta com uma célula-alvo do que com uma célula não alvo, opcionalmente em que a taxa mais alta é de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

ii) o fusossoma se funde a uma taxa mais alta com uma célula-alvo do que com outro fusossoma, opcionalmente em que a taxa mais alta é de pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

iii) o fusossoma se funde com células do fígado a uma taxa tal que o agente exógeno no fusossoma seja entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% de células do fígado após 24, 48 ou 72 horas;

iv) o fusossoma entrega o ácido nucleico a uma célula do fígado a uma taxa mais alta do que a uma célula não hepática, opcionalmente em que a taxa mais alta é de pelo menos 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes;

v) o fusossoma entrega o ácido nucleico a uma célula do fígado a uma taxa mais alta do que a outro fusossoma, opcionalmente em que a taxa mais alta é de pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90%, 2 vezes, 3 vezes, 4 vezes, 5 vezes, 10 vezes, 20 vezes, 50 vezes ou 100 vezes; ou

vi) o fusossoma entrega o ácido nucleico a uma célula-alvo a uma taxa tal que um agente no fusossoma seja entregue a pelo menos 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% ou 90% das células-alvo após 24, 48 ou 72 horas.

21. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 20, caracterizado pelo fato de que o gene de carga útil que codifica um agente exógeno é escolhido dentre: OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAH, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LMBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTD, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1 ou LDLR.

22. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 21, caracterizado pelo fato de que o fusogênio é uma proteína de envelope viral.

23. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 22, caracterizado pelo fato de que o fusogênio compreende VSV-G.

24. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 21, caracterizado pelo fato de que o fusogênio compreende

uma sequência escolhida dentre proteínas F e G do vírus Nipah, proteínas F e H do vírus do sarampo, proteínas F e H do paramixovírus de tupaia, proteínas F e G do paramixovírus ou proteínas F e H ou proteínas F e HN, proteínas F e G do vírus Hendra, proteínas F e G do Henipavírus, proteínas F e H do morbilivírus, proteína F e HN do respirovírus, uma proteína F e HN do vírus Sendai, proteínas F e HN do rubulavírus ou proteínas F e HN do avulavírus, ou um derivado das mesmas, ou qualquer combinação das mesmas.

25. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 21 e 24, caracterizado pelo fato de que o fusogênio compreende um domínio de pelo menos 100 aminoácidos de comprimento com pelo menos 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade de sequências para um fusogênio de paramixovírus de tipo selvagem, opcionalmente em que o fusogênio de paramixovírus de tipo selvagem é estabelecido em qualquer uma das SEQ ID NOS: 1 a 132.

26. Fusossoma, de acordo com a reivindicação 25, caracterizado pelo fato de que o paramixovírus de tipo selvagem é um vírus Nipah, opcionalmente em que o vírus Nipah é um henipavírus.

27. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 26, caracterizado pelo fato de que o fusogênio é realvejado para entrega a uma célula do fígado.

28. Fusossoma, de acordo com qualquer uma das reivindicações 1, 2, 5, 6, 7 e 14 a 27, caracterizado pelo fato de que o elemento regulatório específico do fígado positivo compreende um promotor específico do fígado, um intensificador específico do fígado, um local de splice específico do fígado, um local específico do fígado que estende a meia-vida de um RNA ou proteína, um local de promoção de exportação nuclear de mRNA específico do fígado, um local de intensificação de tradução específico do fígado ou um local de modificação pós-tradução específico do fígado.

29. Fusossoma, de acordo com a reivindicação 28, caracterizado pelo fato de que o elemento regulatório específico do fígado positivo compreende um promotor específico de hepatócito.

30. Fusossoma, de acordo com a reivindicação 28 ou 29, caracterizado pelo fato de que o elemento regulatório específico do fígado positivo compreende um promotor selecionado dentre um promotor transtirretina intensificada (ET), Alb, Apoa2, Cyp3a4, LP1B, MIR122, hemopexina, SERPINA1 ou HLP.

31. Fusossoma, de acordo com a reivindicação 3 ou 30, caracterizado pelo fato de que o promotor tem a sequência estabelecida em qualquer uma das SEQ ID NO: 133 a 136 ou 519 a 525 ou uma sequência com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com o mesmo.

32. Fusossoma, de acordo com a reivindicação 28 ou 29, caracterizado pelo fato de que o elemento regulatório específico do fígado positivo compreende um promotor ApoE.HCR-hAAT, opcionalmente em que o promotor compreende a sequência estabelecida na SEQ ID NO: 133, ou uma sequência com pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% ou 99% de identidade com a sequência estabelecida na SEQ ID NO:133.

33. Fusossoma, de acordo com qualquer uma das reivindicações 2, 6, 7 e 15 a 32, caracterizado pelo fato de que o NTCSRE compreende uma sequência de reconhecimento de miRNA específico de célula não hepática, local de reconhecimento de protease específico de célula não hepática, local de ubiquitina ligase específico de célula não hepática, local de repressão transcricional específico de célula não hepática ou local de repressão epigenética específico de células não hepática.

34. Fusossoma, de acordo com a reivindicação 33, caracterizado pelo fato de que o NTCSRE compreende uma sequência de

reconhecimento de miRNA específica de célula não hepática e a sequência de reconhecimento de miRNA é capaz de ser ligada por um ou mais dentre miR-142, mir-181a-2, mir-181b-1, mir-181c, mir-181a-1, mir-181b-2, mir-181d, miR-223 ou miR-126.

35. Fusossoma, de acordo com a reivindicação 33 ou 34, caracterizado pelo fato de que o NTCSRE está situado ou codificado dentro de uma região transcrita que codifica o agente exógeno, opcionalmente em que um RNA produzido pela região transcrita compreende a sequência de reconhecimento de miRNA específica de tecido dentro de uma UTR ou de uma região de codificação.

36. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 35, caracterizado pelo fato de que o ácido nucleico compreende um ou mais elementos isolantes.

37. Fusossoma, de acordo com a reivindicação 36, caracterizado pelo fato de que o ácido nucleico compreende dois elementos isolantes, opcionalmente em que os dois elementos isolantes compreendem um primeiro elemento isolante a montante do gene de carga útil e um segundo elemento isolante a jusante do gene de carga útil, opcionalmente em que o primeiro elemento isolante e o segundo elemento isolante compreendem as mesmas sequências ou sequências diferentes.

38. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 37, caracterizado pelo fato de que o fusossoma é uma partícula de vetor retroviral.

39. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 38, caracterizado pelo fato de que o ácido nucleico é capaz de se integrar ao genoma de uma célula do fígado.

40. Fusossoma, de acordo com qualquer uma das reivindicações 1 a 39, caracterizado pelo fato de que a célula do fígado é escolhida dentre um hepatócito, célula endotelial sinusoidal do fígado,

colangiócito, célula estrelada, célula apresentadora de antígeno residente no fígado, linfócito imune residente no fígado ou fibroblasto portal.

41. Composição farmacêutica, caracterizada pelo fato de que compreende o fusossoma, como definido em qualquer uma das reivindicações 1 a 40, e um carreador, diluente ou excipiente farmaceuticamente aceitável.

42. Método de entrega de um agente exógeno a um sujeito, caracterizado pelo fato de que compreende administrar ao sujeito o fusossoma, como definido em qualquer uma das reivindicações 1 a 40, ou a composição farmacêutica, como definida na reivindicação 41, entregando, assim, o agente exógeno ao sujeito.

43. Método de modulação de uma função, em um sujeito, fígado ou célula hepática, caracterizado pelo fato de que compreende colocar fígado ou a célula do fígado do sujeito em contato com o fusossoma, como definido em qualquer uma das reivindicações 1 a 40, ou a composição farmacêutica, como definida na reivindicação 41.

44. Método, de acordo com a reivindicação 43, caracterizado pelo fato de que o tecido-alvo ou a célula-alvo está presente no sujeito.

45. Método, de acordo com a reivindicação 43 ou 44, caracterizado pelo fato de que o contato é realizado pela administração do fusossoma ao sujeito.

46. Método de tratamento de uma deficiência genética em um sujeito, caracterizado pelo fato de que compreende administrar ao sujeito o fusossoma, como definido em qualquer uma das reivindicações 1 a 40, ou a composição farmacêutica, como definida na reivindicação 41.

47. Método, de acordo com a reivindicação 46, caracterizado pelo fato de que a deficiência genética é uma deficiência genética

capaz de ser tratada pelo gene de carga útil que codifica o agente exógeno.

48. Método, de acordo com qualquer uma das reivindicações 42 a 47, caracterizado pelo fato de que o sujeito é um sujeito humano.

49. Fusossoma, como definido em qualquer uma das reivindicações 1 a 40, ou composição farmacêutica, como definida na reivindicação 41, caracterizada pelo fato de que é para uso no tratamento de um sujeito com uma deficiência genética.

50. Uso de um fusossoma, como definido em qualquer uma das reivindicações 1 a 40, ou composição farmacêutica, como definida na reivindicação 41, caracterizado pelo fato de que é para a fabricação de um medicamento para uso no tratamento de um sujeito com uma deficiência genética.

51. Fusossoma ou composição farmacêutica para uso, de acordo com a reivindicação 49, ou uso, de acordo com a reivindicação 50, caracterizado pelo fato de que o fusossoma compreende um gene de carga útil que codifica um agente exógeno para tratar a deficiência genética.

52. Método de produção do fusossoma, de acordo com qualquer uma das reivindicações 1 a 40, caracterizado pelo fato de que compreende:

a) fornecer uma célula que compreende o ácido nucleico e o fusogênio;

b) cultivar a célula sob condições que permitem a produção do fusossoma, e

c) separar, enriquecer ou purificar o fusossoma da célula, produzindo, assim, o fusossoma.

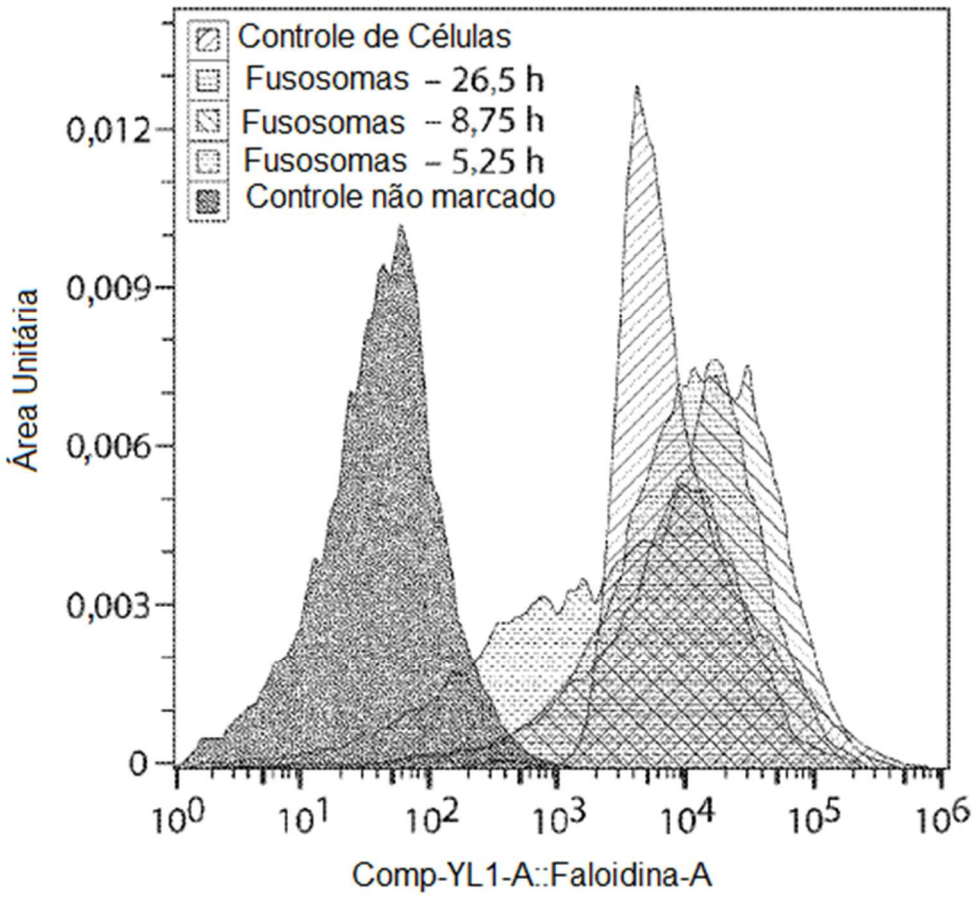

Controle de Células
Fusosomas – 26,5 h
Fusosomas – 8,75 h
Fusosomas – 5,25 h
Controle não marcado
Área Unitária
0,012
0,009
0,006
0,003
0
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Comp-YL1-A::Faloidina-A

Fig. 1

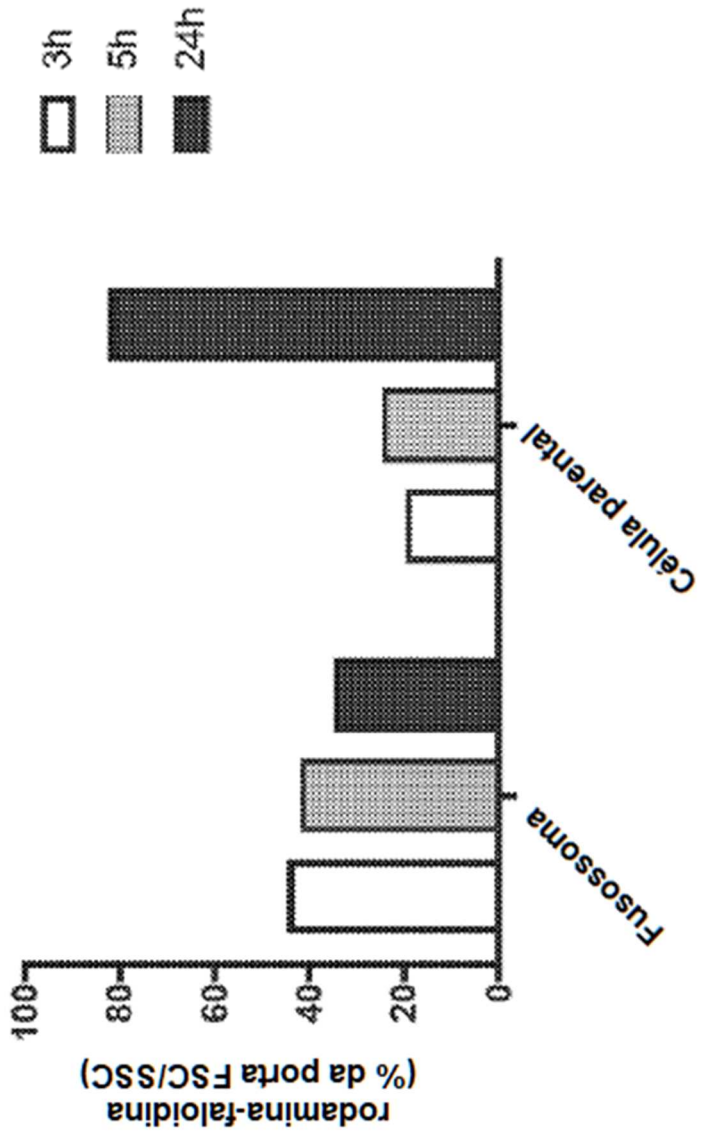

3h
5h
24h
rodamina-faloidina
(% da porta FSC/SSC)
100
80
60
40
20
0
Fusossoma
Célula parental

Fig. 2

	Fusossomas	Células parentais
Diâmetro médio (nm)	128,7	175,4
Diâmetro mínimo (nm)	14	29
Diâmetro máximo (nm)	18720	19802
Diâmetro mediano (nm)	134	99
10% quantil (nm)	53	52
25% quantil (nm)	88	66
75% quantil (nm)	226	241
90% quantil (nm)	4450	10649
Volume médio (μm^3)	0,067	7,421

Fig. 3

	Fusossomas	Células parentais
Diâmetro médio (nm)	128,7	175,4
Volume médio (μm^3)	0,067	7,421

Fig. 4

	[Proteína] (mg/ml)		Razão
	Insolúvel	Solúvel	
Fusossomas celulares	2,12	0,06	0,028
	1,28	0,72	0,563

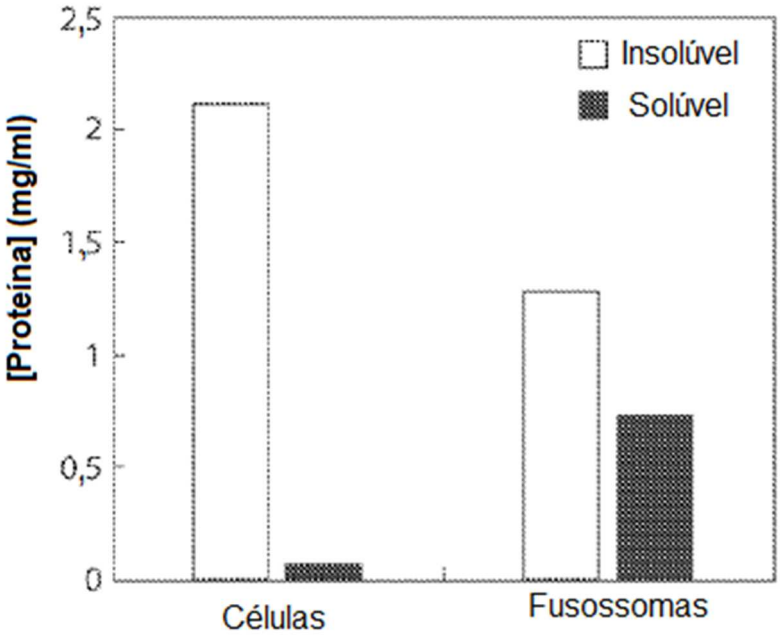
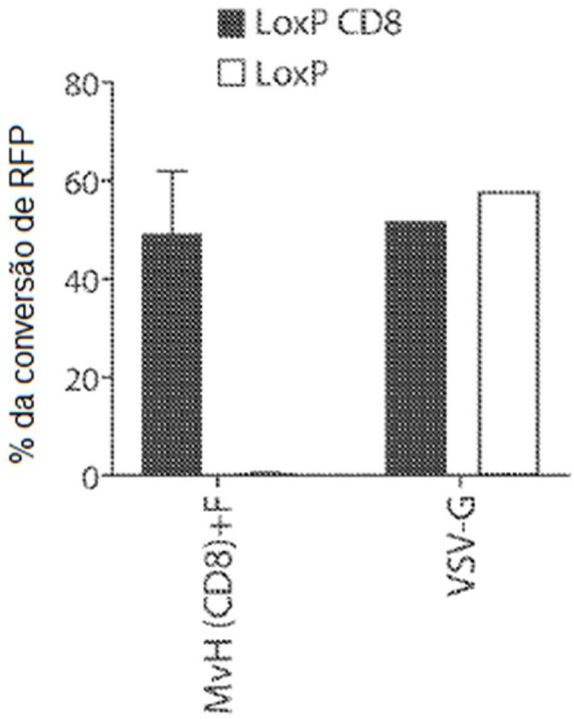


Fig. 5



Fusossoma	Quantidade absoluta de fusão alvejada
MvH(CD8)+F	48,8%
VSV-G	0%

Fig. 6

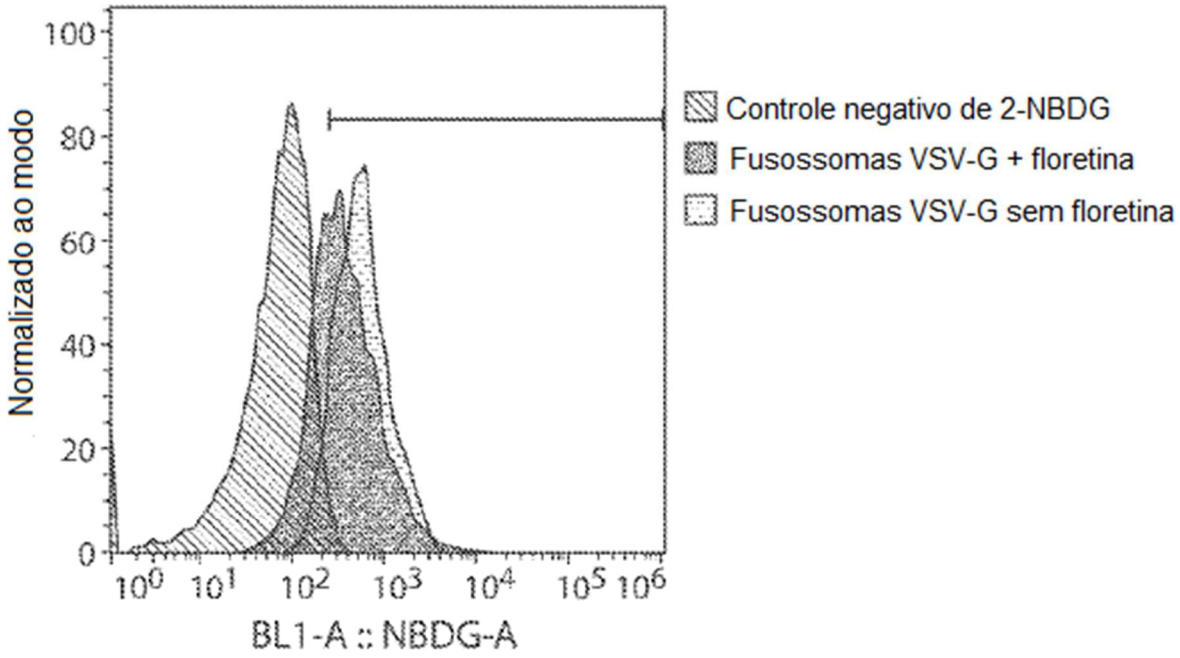

Normalizado ao modo
100
80
60
40
20
0
10^0 10^1 10^2 10^3 10^4 10^5 10^6
BL1-A :: NBDG-A
Controle negativo de 2-NBDG
Fusossomas VSV-G + floretina
Fusossomas VSV-G sem floretina

Fig. 7

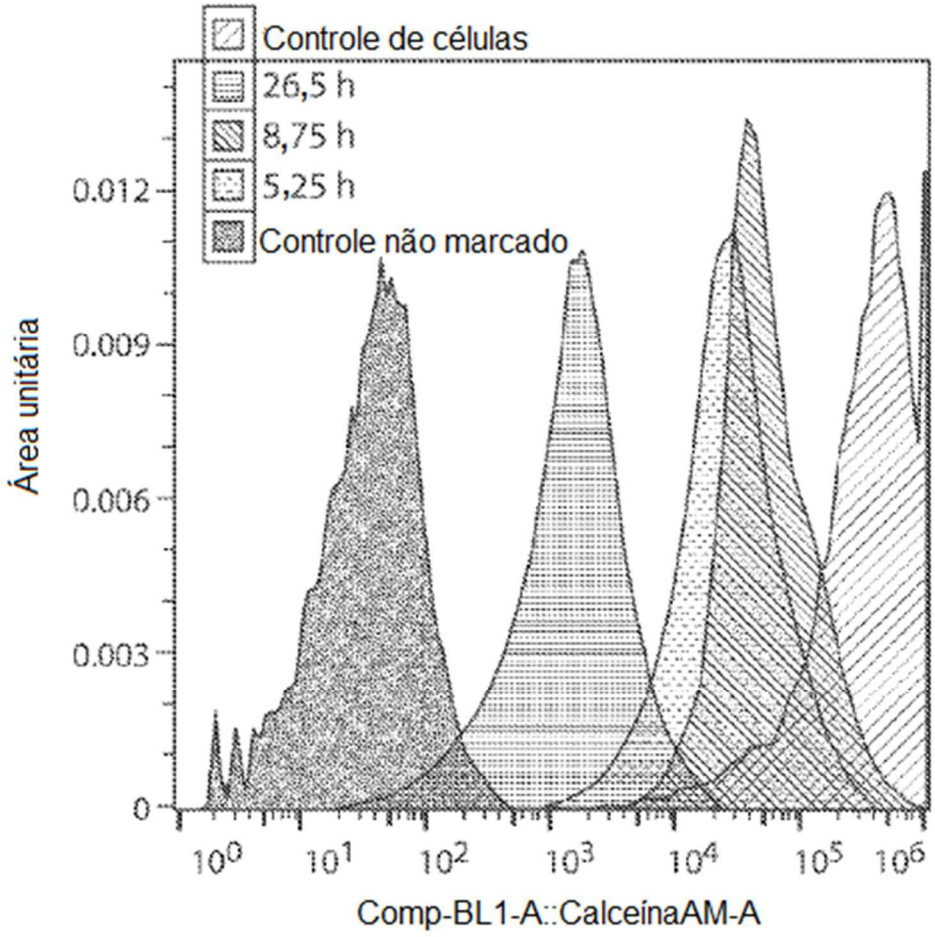

Controle de células
26,5 h
8,75 h
5,25 h
Controle não marcado
Área unitária
0.012
0.009
0.006
0.003
0
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Comp-BL1-A::CalceínaAM-A

Fig. 8

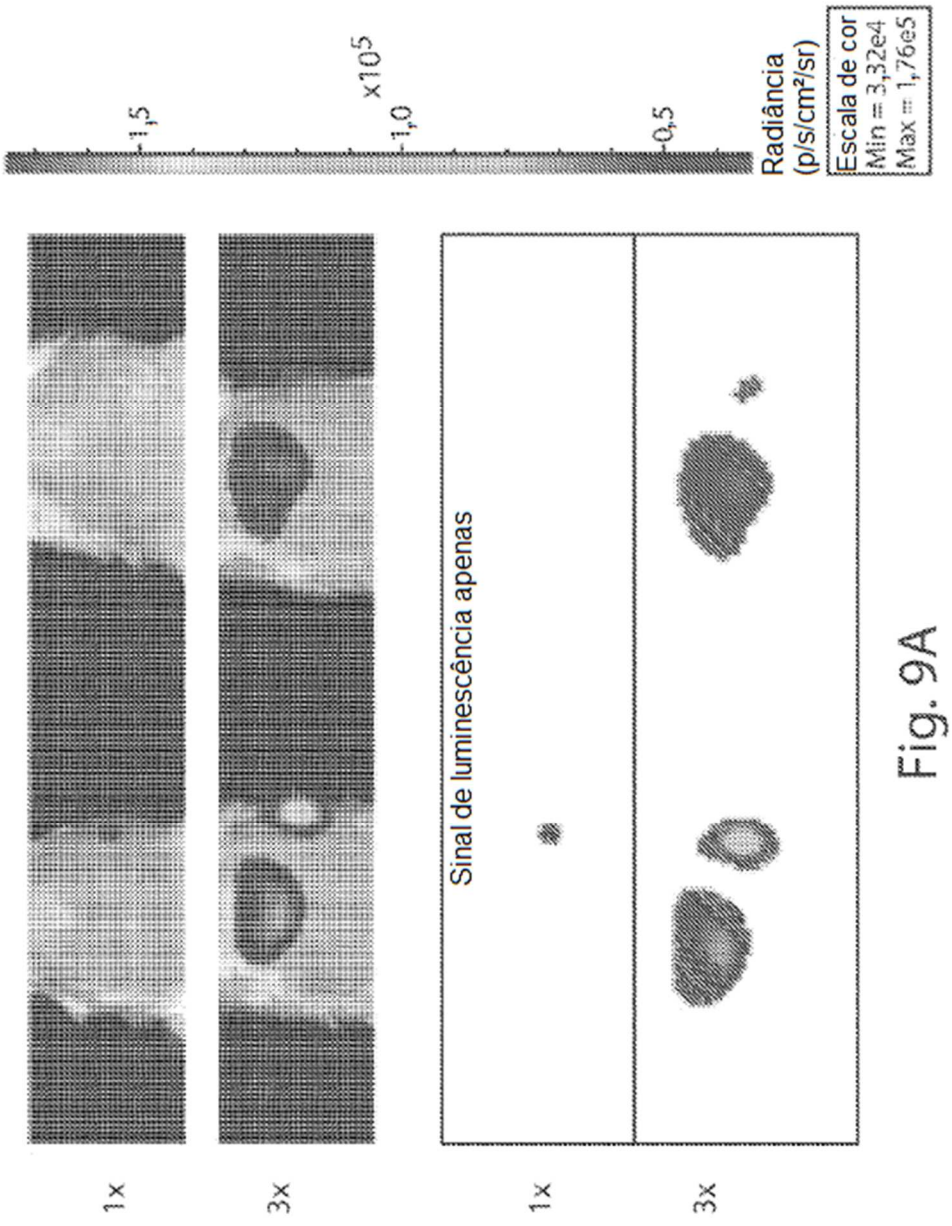

1,5
x10⁵
1,0
0,5
Radiância (p/s/cm²/sr)
Escala de cor
Min = 3,32e4
Max = 1,76e5
Sinal de luminescência apenas
1x
3x
1x
3x

Fig. 9A

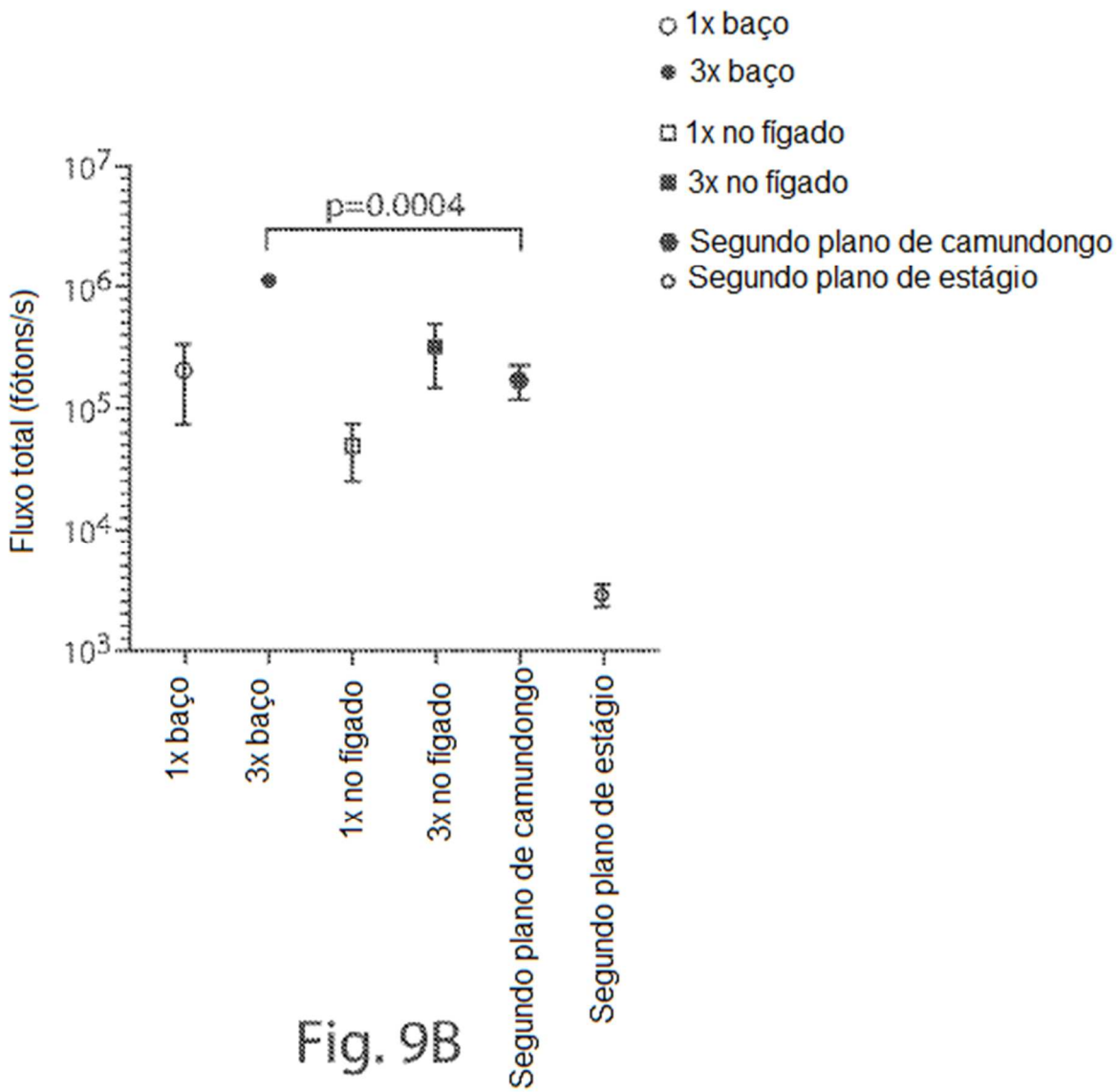

1x baço
3x baço
1x no fígado
3x no fígado
Segundo plano de camundongo
Segundo plano de estágio
p=0.0004
Fluxo total (fótons/s)
10^7
10^6
10^5
10^4
10^3
1x baço
3x baço
1x no fígado
3x no fígado
Segundo plano de camundongo
Segundo plano de estágio

Fig. 9B

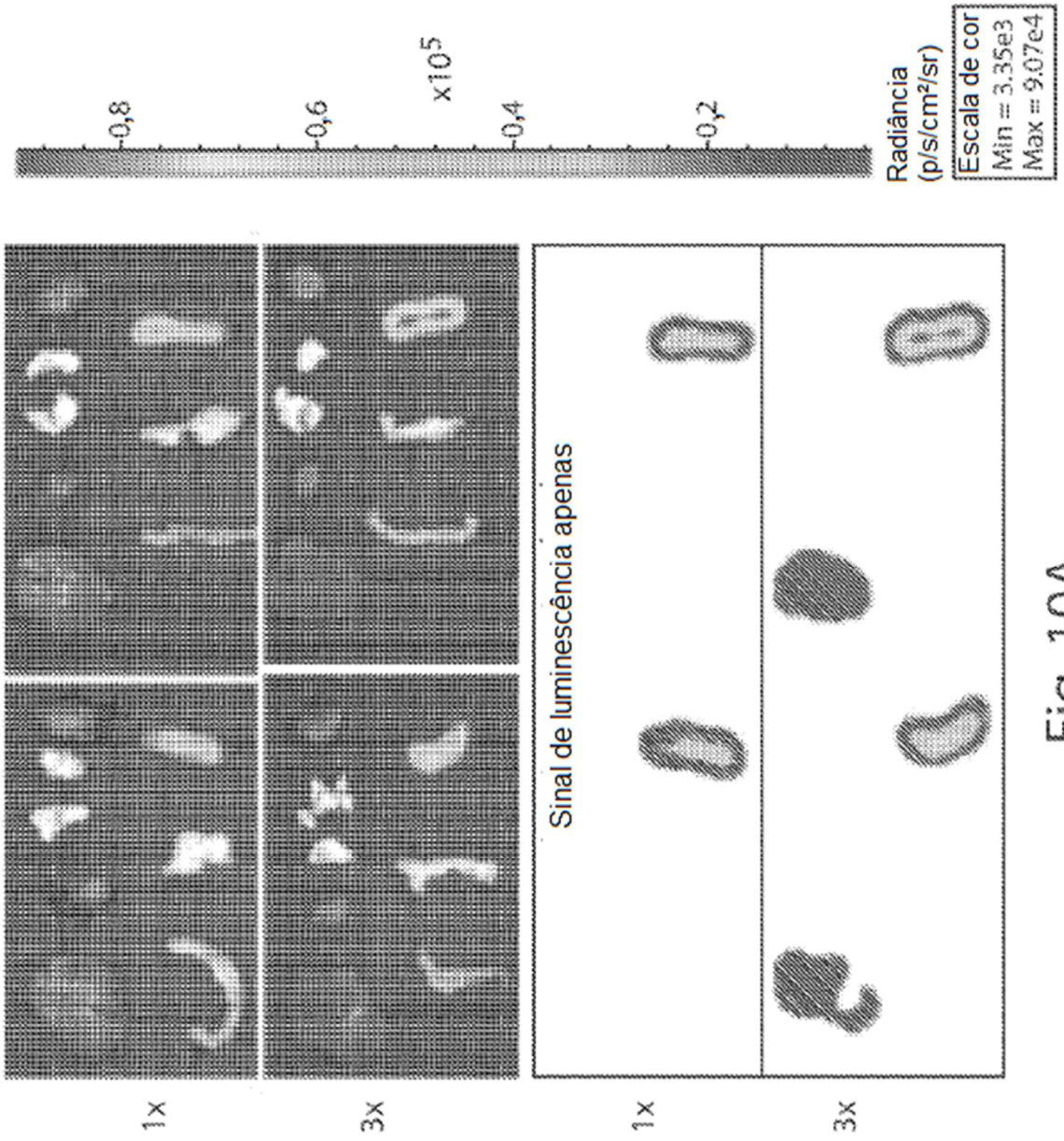

Escala de cor
Min = 3.35e3
Max = 9.07e4
Radiância
(p/s/cm²/sr)
x10⁵
0,8
0,6
0,4
0,2
Sinal de luminescência apenas
1x
3x
1x
3x

Fig. 10A

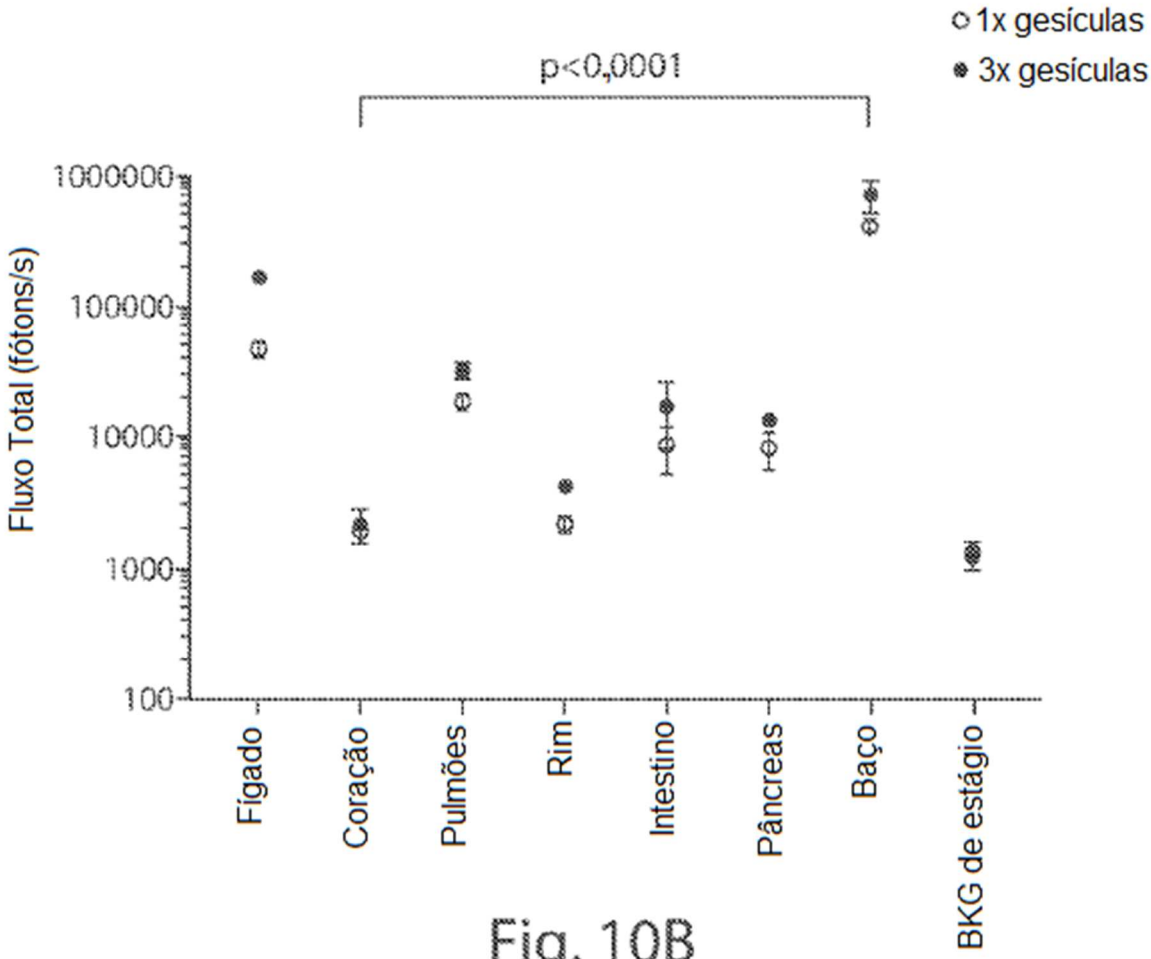

1x gesículas
3x gesículas
p<0,0001
Fluxo Total (fótons/s)
1000000
100000
10000
1000
100
Fígado
Coração
Pulmões
Rim
Intestino
Pâncreas
Baço
BKG de estágio

Fig. 10B

Petição 870210022257, de 09/03/2021, pág. 17/731

Grupo	% da conversão de RFP (±SD)
Receptor + sem fusossoma	0,4 ± 0,2%
Receptor +NivG+F fusossoma	88,9 ± 3,4%
Receptor + NivG+F fusossoma + Baf	68,1 ± 2,7%

Fig. 11

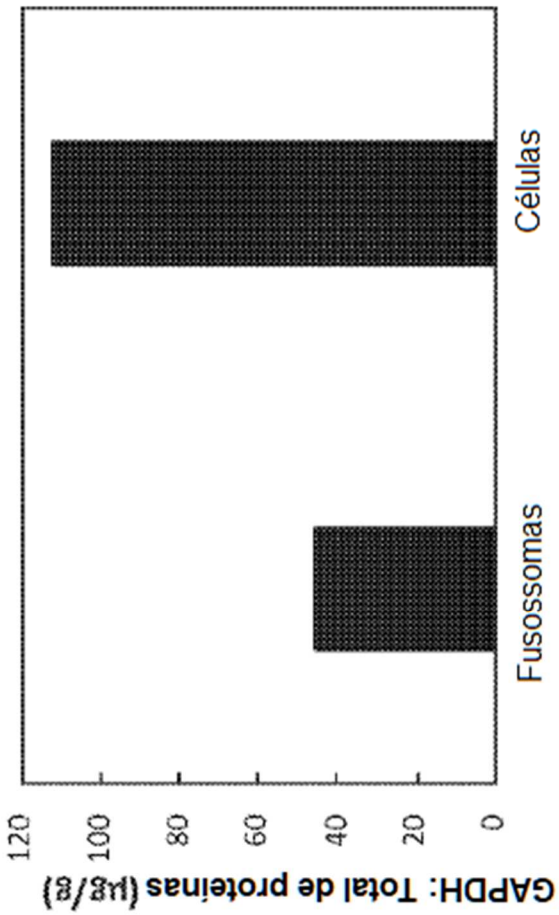

GAPDH: Total de proteínas (µg/g)
120
100
80
60
40
20
0
Fusossomas
Células

Fig. 12

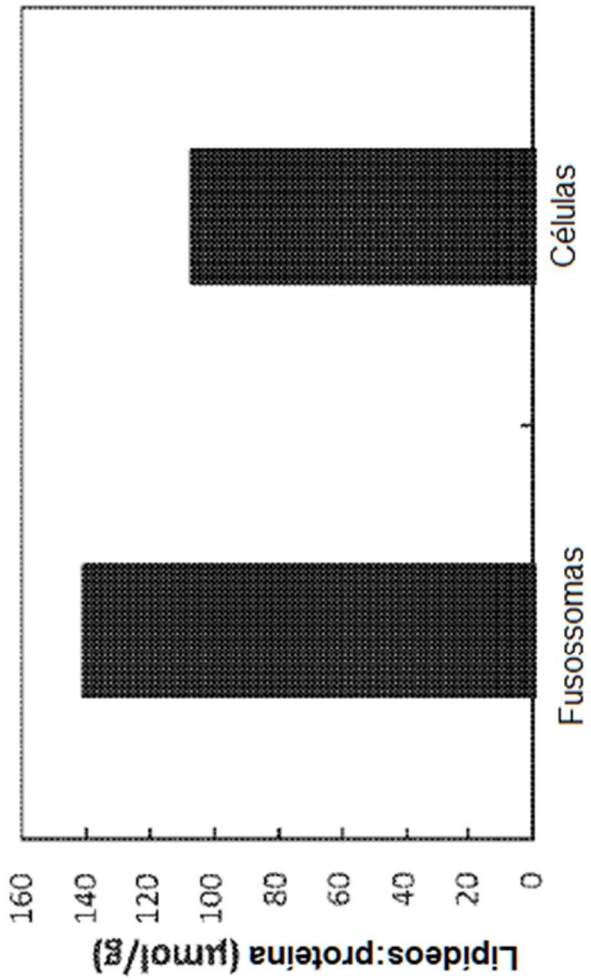

Lipídeos:proteína (µmol/g)
160
140
120
100
80
60
40
20
0
Fusossomas
Células

Fig. 13

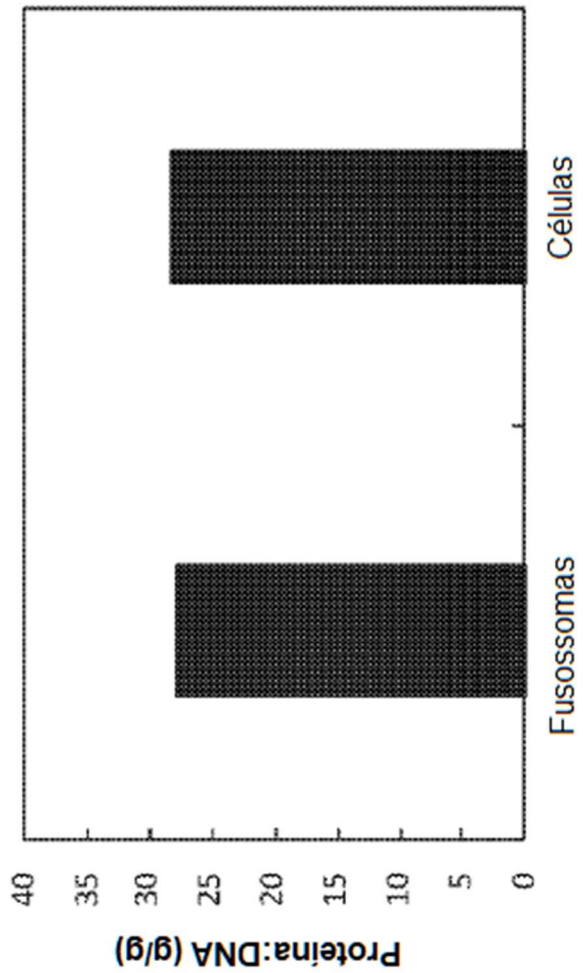


Fig. 14

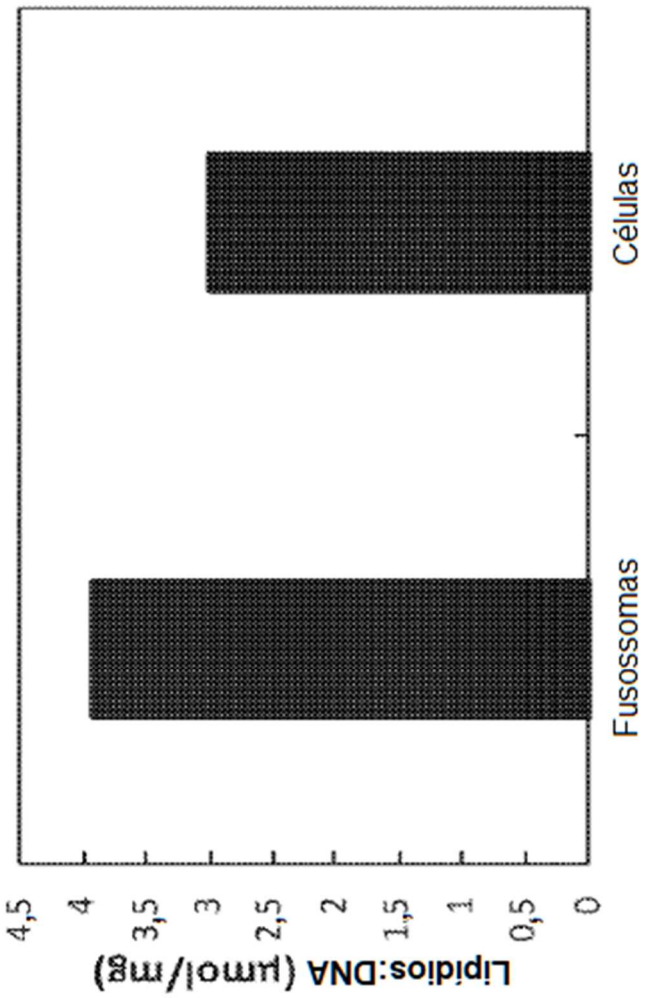


Fig. 15

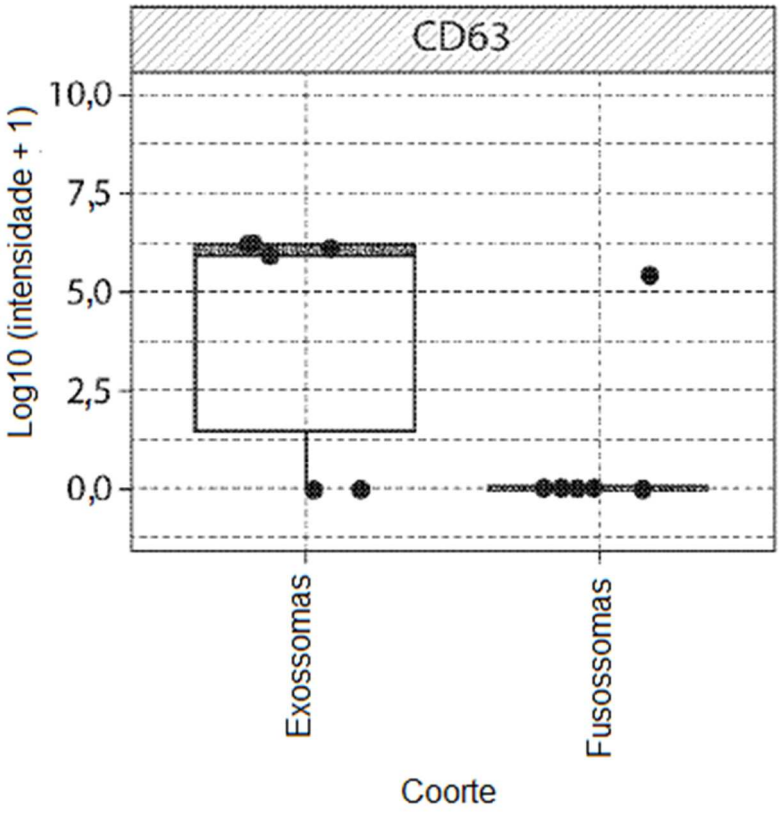

CD63
Log10 (intensidade + 1)
10,0
7,5
5,0
2,5
0,0
Exossomas
Fusossomas
Coorte

Fig. 16

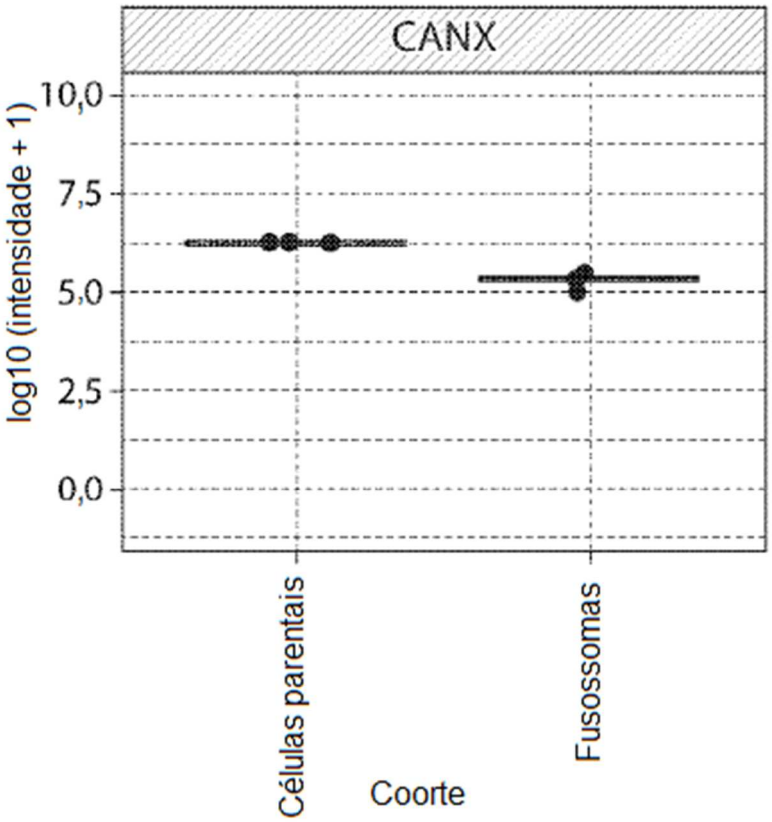

CANX
log10 (intensidade + 1)
10,0
7,5
5,0
2,5
0,0
Células parentais
Fusossomas
Coorte

Fig. 17

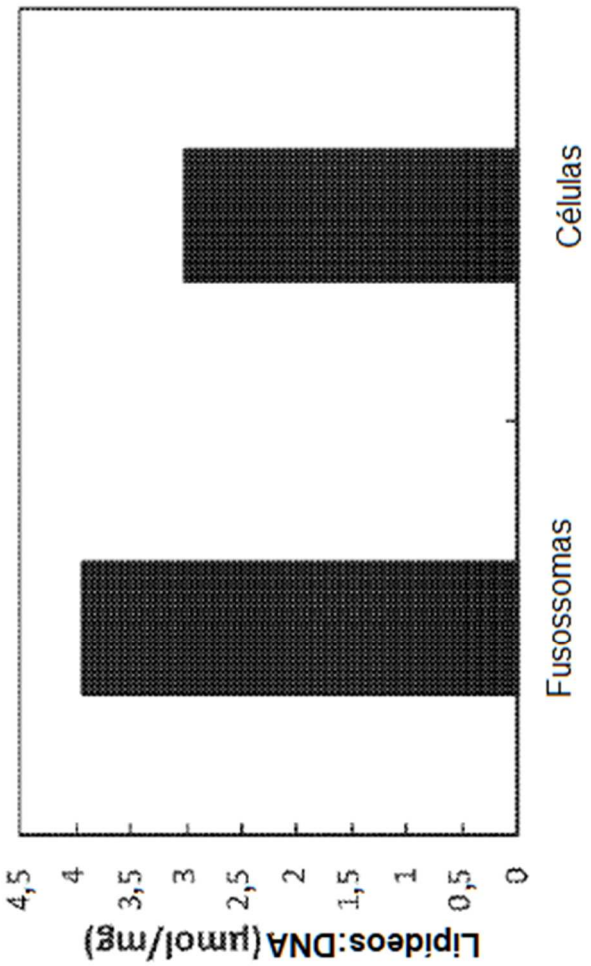


Fig. 18

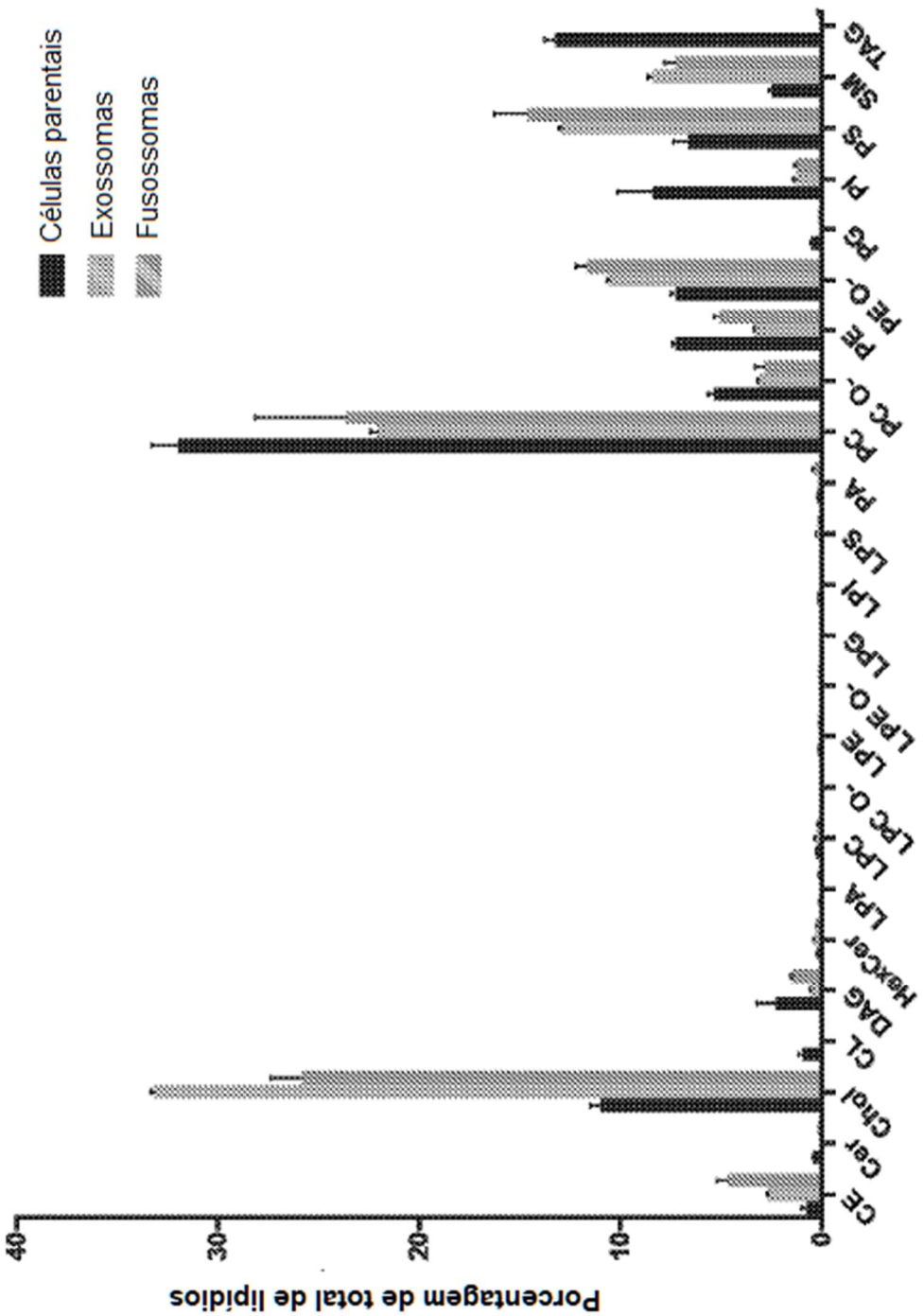

Células parentais
Exossomas
Fusossomas
Porcentagem de total de lipídios
0
10
20
30
40
CE
Cer
Chol
CL
DAG
HexCer
LPA
LPC
LPC O-
LPE
LPE O-
LPG
LPI
LPS
PA
PC
PC O-
PE
PE O-
PG
PI
PS
SM
TAG

Fig. 19A

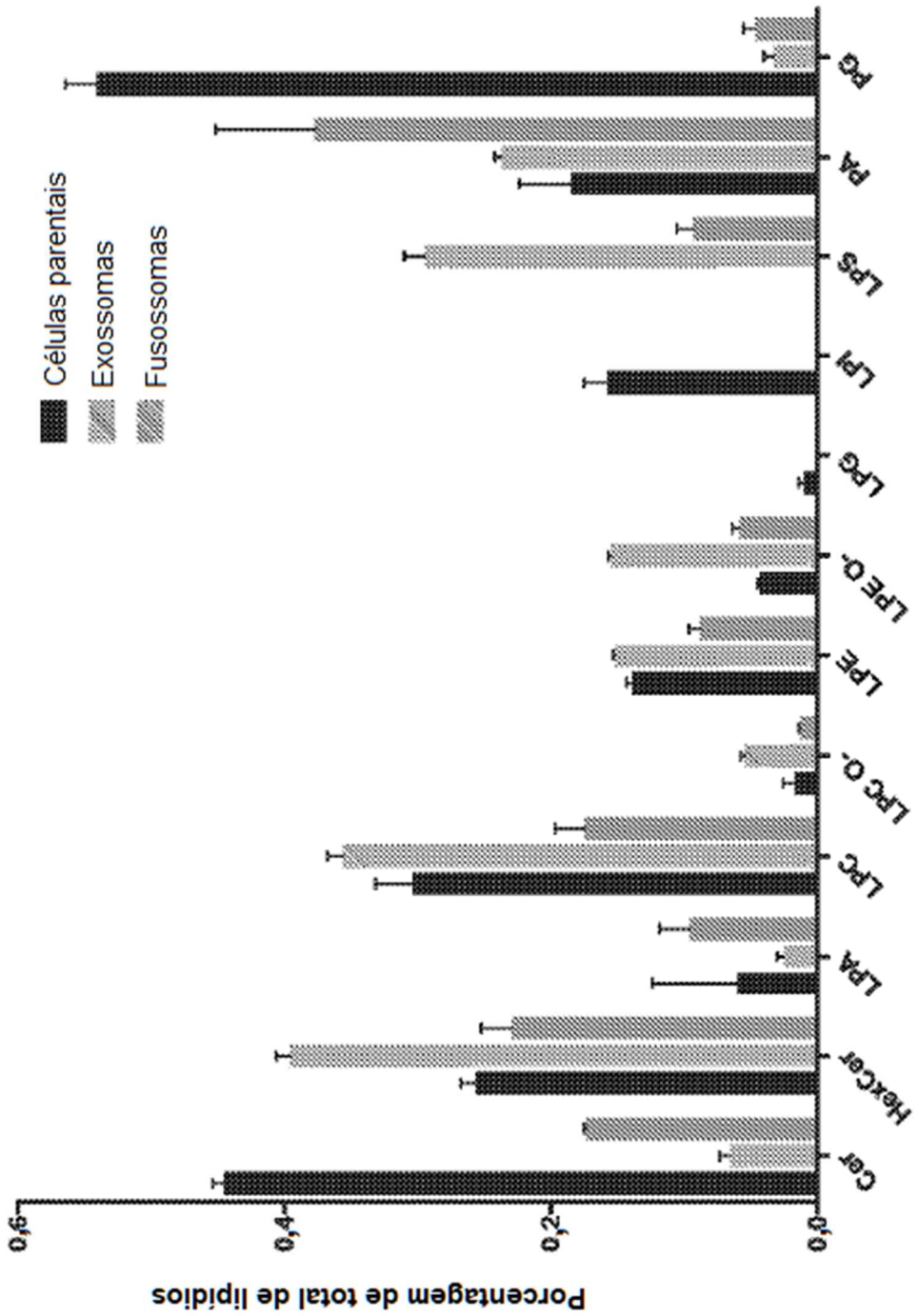

Células parentais
Exossomas
Fusossomas
Porcentagem de total de lipídios
0,0
0,2
0,4
0,6
Cer
HexCer
LPA
LPC
LPC O-
LPE
LPE O-
LPG
LPI
LPS
PA
PG

Fig. 19B

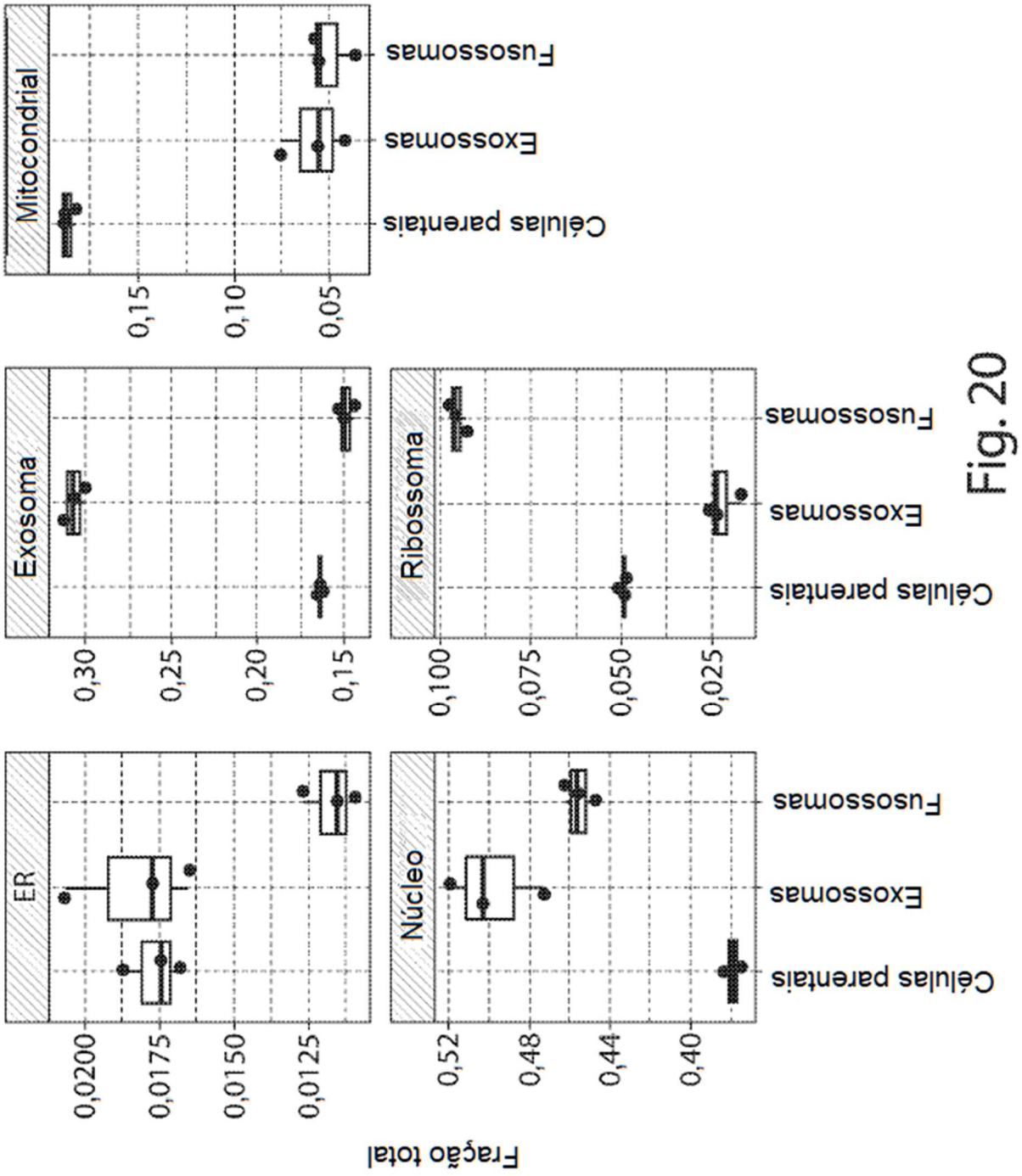

Fração total
ER
Núcleo
Exosoma
Ribossoma
Mitocondrial
0,0200
0,0175
0,0150
0,0125
0,52
0,48
0,44
0,40
0,30
0,25
0,20
0,15
0,100
0,075
0,050
0,025
0,15
0,10
0,05
Células parentais
Exossomas
Fusossomas

Fig. 20

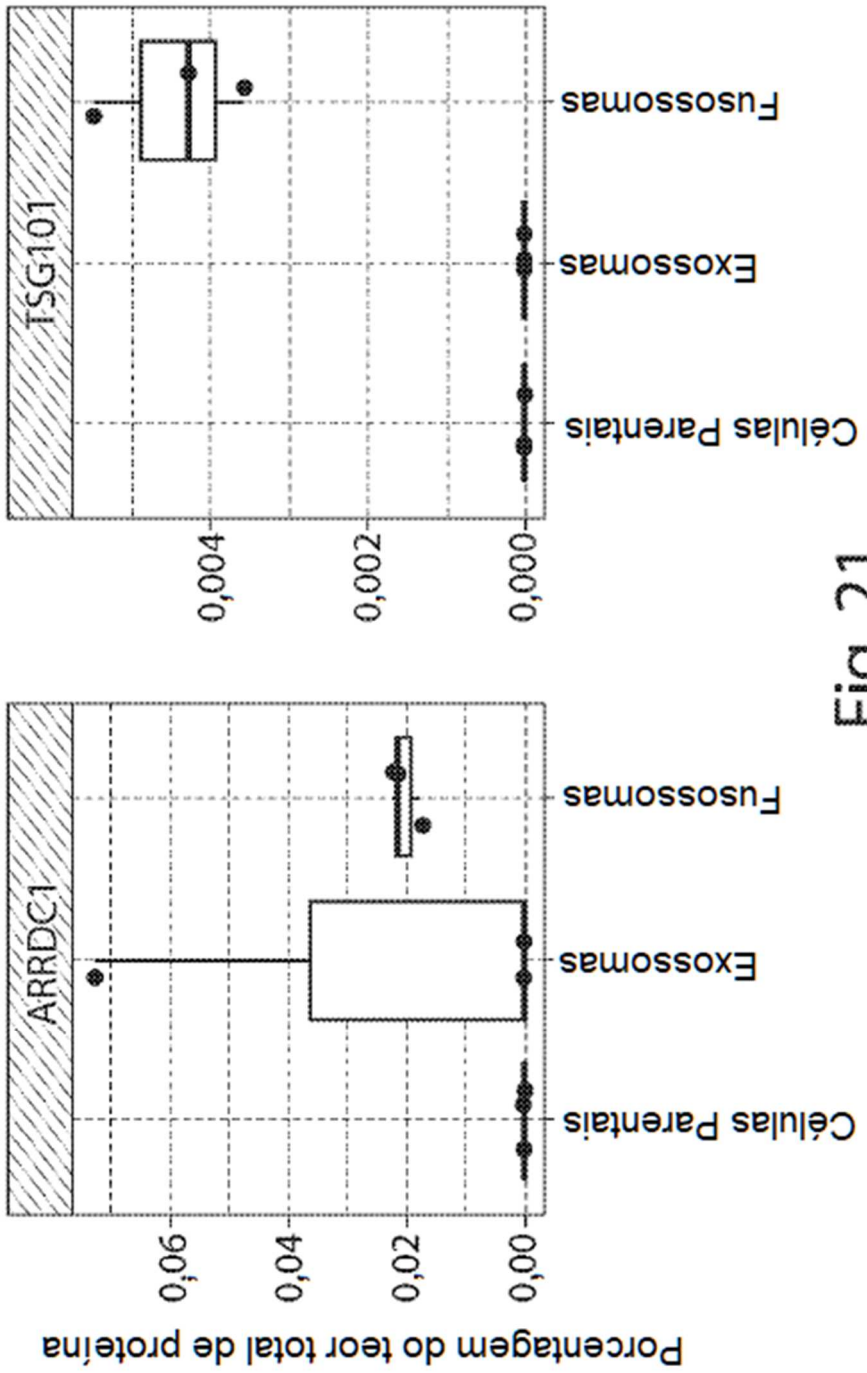

Porcentagem do teor total de proteína
0,00
0,02
0,04
0,06
ARRDC1
Células Parentais
Exossomas
Fusossomas
0,000
0,002
0,004
TSG101
Células Parentais
Exossomas
Fusossomas

Fig. 21

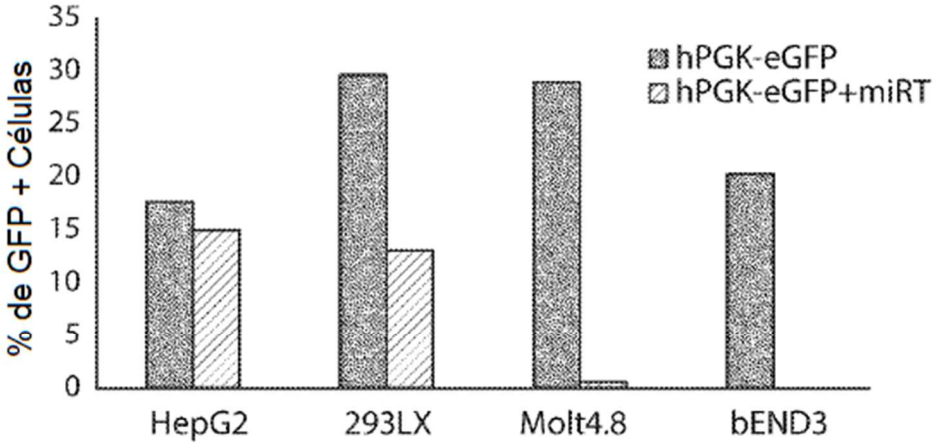


Fig. 22A

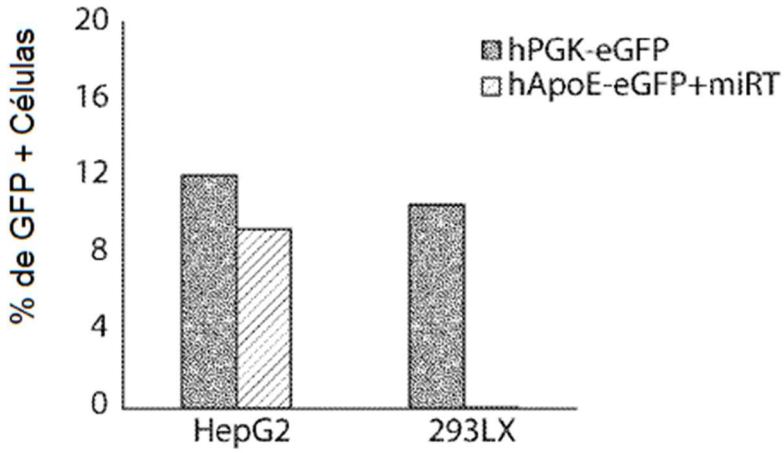


Fig. 22B

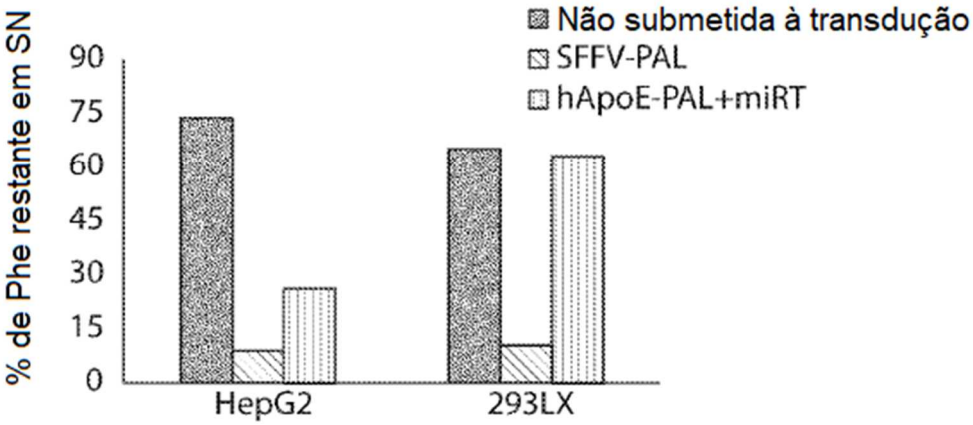


Fig. 22C

RESUMO

Patente de Invenção: “**COMPOSIÇÕES DE FUSOSSOMA E USOS DAS MESMAS”.**

A presente invenção fornece, pelo menos em parte, métodos e composições para entrega *in vivo* de fusossoma. Em algumas modalidades, o fusossoma compreende uma combinação de elementos que promovem a especificidade para células-alvo, por exemplo, um ou mais dentre um fusogênio, um elemento regulatório específico de célula-alvo positivo e um elemento regulatório específico de célula não alvo. Em algumas modalidades, o fusossoma compreende uma ou mais modificações que diminuem uma resposta imunológica contra o fusossoma.

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 JTD TRD P251312.txt
- Data de Geração do Código: 09/03/2021
- Hora de Geração do Código: 10:28:22
- Código de Controle:
  - Campo 1: 07E82F48E682E274
  - Campo 2: 73B02C8B7D090C56