



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

(21) BR 112014013597-5 A2



(22) Data do Depósito: 05/12/2012

(43) Data da Publicação Nacional: 18/08/2020

(54) Título: CASSETE DE EXPRESSÃO, VETOR DE EXPRESSÃO, CÉLULA HOSPEDEIRA, MÉTODO IN VITRO PARA EXPRESSÃO DE UM POLIPEPTÍDEO E USO DE UM CASSETE DE EXPRESSÃO OU VETOR DE EXPRESSÃO

(51) Int. Cl.: C12N 15/09.

(30) Prioridade Unionista: 07/12/2011 US 61/567,675.

(71) Depositante(es): GLENMARK PHARMACEUTICALS S.A..

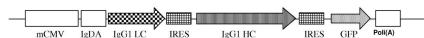
(72) Inventor(es): DANIEL LUESCHER; CHRISTEL AEBISCHER-GUMY; PIERRE MORETTI; MARTIN BERTSCHINGER.

(86) Pedido PCT: PCT IB2012056977 de 05/12/2012

(87) Publicação PCT: WO 2013/084157 de 13/06/2013

(85) Data da Fase Nacional: 05/06/2014

(57) Resumo: CASSETE DE EXPRESSÃO, VETOR DE EXPRESSÃO, CÉLULA HOSPEDEIRA, MÉTODO IN VITRO PARA EXPRESSÃO DE UM POLIPEPTÍDEO E USO DE UM CASSETE DE EXPRESSÃO OU VETOR DE EXPRESSÃO. A presente invenção refere-se a um cassete de expressão útil para a expressão de uma sequência de polinucleotídios codificantes de um polipeptídeo.



"CASSETE DE EXPRESSÃO, VETOR DE EXPRESSÃO, CÉLULA HOSPEDEIRA, MÉTODO *IN VITRO* PARA EXPRESSÃO DE UM POLIPEPTÍDEO E USO DE UM CASSETE DE EXPRESSÃO OU VETOR DE EXPRESSÃO"

Campo técnico

[0001] A presente invenção se refere a um cassete de expressão útil para expressão de uma sequência de polinucleotídio codificante de um polipeptídeo. A presente invenção também é dirigida aos vetores e células hospedeiras que compreendem o cassete de expressão e usam o cassete de expressão para a produção de um polipeptídeo a partir de uma célula hospedeira.

Antecedentes da invenção

[0002] Os sistemas de expressão para produção dos peptídeos recombinantes são bem conhecidos do estado da técnica e são descritos em, por exemplo, Marino MH (1989) "Biopharm.", 2:18-33; Goeddel DV et al. (1990) "Methods Enzymol", 185: 3-7; Wurm F & Bernard A (1999) "Curr Opin Biotechnol", 10: 156-159. Os polipeptídios para uso nas aplicações farmacêuticas são preferivelmente produzidos em células mamíferas tais como células CHO, células NOS, células SP2/0, células COS, células HEK, células BHK, ou do gênero. Os elementos essenciais de um vetor de expressão utilizados para este propósito são normalmente selecionados a partir de uma unidade de propagação de plasmídeo procariótico, por exemplo, *E. coli*, compreendendo uma origem procariótica de replicação e um marcador de seleção procariótico, opcionalmente, um marcador de seleção eucariótico, e um ou mais cassetes de expressão para a expressão de genes estruturais de interesse, cada um compreendendo um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo e, opcionalmente, um terminal

de transcrição incluindo um sinal de poliadenilação. Para a expressão transitória em células de mamífero, uma origem de mamífero de replicação, tal como SV40 Ori ou OriP, podem ser incluídos. Como promotor, um promotor constitutivo ou indutível pode ser selecionado. Para transcrição otimizada, uma sequência Kozak pode ser incluída na região não traduzida 5'. Para o processamento de mRNA, em particular na combinação de mRNA e finalização da transcrição, os sinais de combinação do mRNA, dependendo da organização do gene estrutural (organização de exon/ítron), pode ser incluídos bem como um sinal de poliadenilação. A expressão de um gene é realizada ou em uma linha de célula estável ou em uma transitória. O nível de expressão estável e alto de um polipeptídeo em uma produção de linha celular é crucial para o processo completo de produção de polipeptídios recombinantes. A ordem para moléculas biológicas tal como proteínas e, especificamente, anticorpos ou fragmentos de antícorpo tem aumentado significativamente em relação aos últimos anos. O alto custo e o pouco rendimento têm sido fatores limitantes na disponibilidade de moléculas biológicas e tem sido um desafio maior desenvolver processos robustos que aumentam o rendimento das moléculas biológicas desejáveis em uma escala industrial. Assim, ainda existe uma necessidade para melhorar a eficiência dos vetores de expressão para obter expressão elevada na produção do polipeptídeo recombinante.

#### Sumário da invenção

[0003] A presente invenção se refere genericamente aos sistemas de expressão, tal como cassetes de expressão e vetores de expressão que podem ser utilizados para obter expressão aumentada na produção de polipeptídeo recombinante.

Em um aspecto, a presente descrição provê um cassete de expressão que comprehende um promotor, uma sequência de polinucleotídio codificante de um polipeptídio, e uma sequência de DNA genômica não traduzida à jusante de um promotor da desidrogenase de gliceraldeído-3-fosfato (GAPDH) eucariótico, sendo que o polipeptídeo codificado pela sequência de polinucleotídio não é GAPDH, e onde a sequência de DNA genômico não traduzido à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo da posição do nucleotídeo +1 para a posição e nucleotídeo em torno de +7000, onde a posição do nucleotídeo é relativa ao inicio da transcrição do mRNA do GAPDH, e sendo que o comprimento da sequência do DNA genômico não traduzida à jusante do promotor GAPDH eucariótico está em torno de 100 para 15000 nucleotídeos.

[0004] Em um aspecto adicional, a presente descrição provê um cassete de expressão que comprehende um promotor, uma sequência de polinucleotídio codificante e um polipeptídio, e uma sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótico, sendo que o polipeptídio codificado pela sequência do polinucleotídio não é GAPDH, e sendo que a sequência do DNA genômico não traduzido à montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição do nucleotídeo em torno de -3500, onde a posição do nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, sendo que o comprimento da sequência do DNA genômico não traduzido a montante do promotor GAPDH eucariótico é de 100 a 15000 nucleotídeos, com a provisão de o cassete de expressão não compreender um

promotor GAPDH eucariótico ou fragmentos dos mesmos. Em um aspecto adicional, a presente descrição provê um vetor de expressão compreendendo um cassete de expressão e uma célula hospedeira compreendendo um cassete de expressão ou um vetor de expressão contendo um cassete de expressão.

[0005] Ainda em um aspecto adicional, a presente descrição provê um método *in vitro* para expressão de um polipeptídeo, compreendendo a transfecção de uma célula hospedeira com um cassete de expressão ou um vetor de expressão e recuperação do polipeptídeo e o uso de um cassete de expressão ou de um vetor de expressão para a expressão de um polipeptídeo heterólogo a partir de uma célula hospedeira de mamíferos.

#### Breve descrição dos desenhos

[0006] A figura 1 ilustra o constructo de expressão repórter (REP) consistindo do promotor citomegalovírus de camundongo (mCMV), o fragmento receptor do doador IG (IgDA) contendo um primeiro ítron, a cadeia leve do anticorpo IgG1 (IgG1 LC), sítios de entrada ribossomal interno derivado do vírus da encefalomiocardite (IRES), a cadeia pesada do anticorpo IgG1 (IgG1 HC), proteína fluorescente verde (GFP) e o sinal de poliadenilação do vírus de símio 40 (poli(A));

[0007] A figura 2 ilustra a expressão transitória do anticorpo IgG1 em células CHO-S no dia 5 após a transfecção (Média de titulação da IgG são pontuadas para duas transfecções independentes). As células foram transfectadas usando os vetores GAPDH\_A e GAPDH\_B (GAPDH\_A e GAPDH\_B), os mesmos vetores sem os elementos a jusante e a montante GAPDH (A e B) e o vetor pGLEX41 como um controle (Pplex41). A concentração do anticorpo IgG1 acumulado no sobrenadante foi determinado usando o instrumento Octet (ForteBio, Menlo, CA,

USA);

[0008] A figura 3 ilustra a expressão do anticorpo IgG1 em células HEK293 EBNA. As células foram transfectadas usando os vetores GAPDH\_A e GAPDH\_B (GAPDH\_A e GAPDH\_B) e o vetor pGLEX41 como um controle (Pplex41). O sobrenadante foi colhido e analisado em 10 dias após a transfecção usando o instrumento Octet. Os dados representam N=3 independente das transfecções em tubos de centrífugas por vetor;

[0009] A figura 4 ilustra um estudo do nível de expressão em uma produção em batelada usando agrupamentos celulares. As células foram transfectadas e os agrupamentos de células estáveis foram criados usando os vetores GAPDH\_A e GAPDH\_B (GAPDH\_A (1), GAPDH\_A (2), e GAPDH\_B(1) e GAPDH\_B (2)), os mesmos vetores sem os elementos a jusante e a montante de GAPDH (A(1), e A(2)) e o vetor pGLEX41 como o controle (pGLEX41). Após 7 dias de cultura, o sobrenadante foi analisado usando o instrumento Octet para o anticorpo acumulado no sobrenadante. A média de titulação da IgG são dadas ( $\mu$ g/ml) para cada agrupamento. Os dados representam N=2 bateladas por grupo;

[0010] A figura 5 ilustra um estudo do nível de expressão sobre as populações geradas pelas transfecções estáveis e diluições limitantes. As células foram transfectadas usando os vetores GAPDH\_A e GAPDH\_B (GAPDH\_A (1), GAPDH\_A (2), e GAPDH\_B(1) e GAPDH\_B (2)), os mesmos vetores sem os elementos a jusante e a montante de GAPDH (A e B) e o vetor pGLEX41 como o controle (pGLEX41). O valor médio de GFP fluorescente expresso pelos clones e mini-agrupamentos a partir das transfecções estáveis foi lido em 14 dias após transfecções. As células foram cultivadas sob a pressão da seleção em

placas de 96 poços. Os dados representam N=48 clones ou mini-agrupamentos por vetor;

[0011] A figura 6 ilustra o efeito de aditivos ao meio, insulina e PMA (forbol 12-miristato-13-acetato, um éster forbol) sobre a expressão do anticorpo IgG1 no sobrenadante. Após a transfecção com o vetor GAPDH\_A (GAPDH\_A) e o vetor Pglex41 como um controle (pGLEX41) às células foram também diluídos no meio PowerCHO2, 4mM Gln, +/- insulina e PowerCHO2, 4mM Gln, insulina PMA +/- . Nenhuma diferença na expressão poderia ser observada comparada ao meio padrão para pGLEX41 (barras cheias) ou GAPDH\_A (barras abertas);

[0012] A figura 7 mostra uma visão geral do GAPDH *in locu*. O gene GAPDH é flanqueado pelos genes NCAPD2 e IFFO1;

[0013] A figura 8 mostra detalhes do gene GAPDH humano, os elementos GAPDH a jusante e a montante e os fragmentos criados para análise do estudo da fragmentação a montante GAPDH. O sítio de restrição NruI foi introduzido para facilitar as etapas de clonagem e não é parte da sequência a montante da GAPDH 5' genômico (foi, portanto, destacado usando um asterisco). Os tamanhos dos fragmentos são: Fragmento 1 (SEQ ID NO:9): 511 bps, o fragmento 2 (SEQ ID NO:10): 2653 bps, fragmento 3 (SEQ ID NO:11): 1966 bps, fragmento 4 (SEQ ID NO:12): 1198 bps, fragmento 8 (SEQ ID NO:13): 259 bps, Fragmento 9 (SEQ ID NO:14): 1947 bps, fragmento 11 (SEQ ID NO:15): 1436 bps, fragmento 17 (SEQ ID NO:16): 1177 bps;

[0014] A figura 9 mostra os resultados da expressão da fragmentação dos elementos GAPDH a montante e a jusante. Os resultados de expressão foram obtidos na transfecção transitória nas células CHO no dia 10 após a transfecção. A

quantificação foi feita usando o instrumento Octet. O vetor pGLEX41 serve como controle negativo. pGLEX41-ampiA também é um controle negativo mostrando a expressão basal do vetor sem os elementos de flancos GAPDH. O pGLEX41 a jusante e a montante contém o flanco de extensão total das regiões (a montante e a jusante) e serve como controle positivo. O pGLEX41 a montante contém apenas a região do flanco a montante e pGLEX41 a jusante contém apenas a região de flanco a jusante. Todos os outros constructos contêm os fragmentos descritos na figura 8. Os fragmentos 2 e 3 foram também clonados na mesma direção como IgG1 LC e IgG1 HC ou em direção oposta na relação a IgG1 LC e IgG1 HC (AS); e [0015] A figura 10 mostra uma expressão transitória do anticorpo IgG1 nas células CHO-S no dia 8 após a transfeção (títulos da IgG média são organizados por três transfeções independentes; barras de erro; SD +/-). As células foram transfetadas usando vetores com os elementos GAPDH de Hamsters Chineses a montante em combinação com o camundongo CMV (A\_GAPDH\_UP) ou o promotor GAPDH de Hamsters Chineses (A\_GAPDH\_UP\_PR). Os plasmídeos tendo apenas o CMV de camundongo (A) ou o promotor GAPDH de Hamsters Chineses (A\_PR) foram transfetados como um controle. A concentração do anticorpo IgG1 acumulado no sobrenadante foi determinada usando o instrumento Octet QK (ForteBio, Menlo, CA, USA).

#### Descrição detalhada da invenção

[0016] A presente descrição refere aos cassetes de expressão e vetores de expressão que compreendem um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômica não traduzida a jusante de um promotor (GAPDH) da desidrogenase

gliceraldeído-3-fosfato, onde o polipeptídeo codificado pela sequência do polinucleotídio não é GAPDH, e sendo que a sequência de DNA genômico não traduzido a jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo da posição do nucleotídeo em torno de +1 à posição de nucleotídeo +7000, onde a posição de nucleotídeo é relativa ao inicio da transcrição do mRNA do GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida a jusante do promotor GAPDH eucariótico está em torno de 100 para 15000 nucleotídeos.

[0017] A presente descrição refere-se ainda a um cassete de expressão que compreende um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômico não traduzida a montante de um promotor GAPDH eucariótico, onde o polipeptídio codificado pela sequência de polinucleotídio não é GAPDH, e sendo que a sequência de DNA genômico não-traduzido a montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição do nucleotídeo em torno de -3500, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA GAPDH, onde o comprimento da sequência de DNA genômico não traduzido a montante do promotor GAPDH eucariótico é de 100 em torno de 15000 nucleotídeos, com a provisão de que o cassete de expressão não compreende um promotor GAPDH eucariótico ou fragmento do mesmo.

[0018] O termo "cassete de expressão" como utilizado aqui inclui uma sequência de polinucleotídio codificante de um polipeptídio a ser expresso e sequências controlando sua expressão tal como um promotor e, opcionalmente, uma

sequência melhoradora, incluindo qualquer combinação dos elementos de controle da transcrição de ação *cis*. A sequência controlando a expressão do gene, ou seja, sua transcrição e a tradução do produto de transcrição são comumente referidas como unidade regulatória. Muitas partes da unidade regulatória estão localizadas a montante da sequência codificante do gene e são operativamente ligadas ao mesmo. O cassete de expressão pode também contém uma região não traduzida em 3' a jusante compreendendo um sítio de poliadenilação. A unidade regulatória da invenção é também operativamente ligada ao gene a ser expressão, ou seja, a unidade de transcrição, ou é separado da mesma através da intervenção no DNA tal como, por exemplo, pela região 5' não traduzida do gene heterólogo. Preferivelmente, o cassete de expressão é flanqueado por um ou mais sítios de restrição de modo a capacitar a inserção do cassete de expressão em um vetor e/ou sua retirada a partir do vetor. Assim, o cassete de expressão de acordo com a presente invenção pode ser utilizado para a construção de um vetor de expressão, em particular um vetor de expressão de mamífero. O cassete de expressão da presente invenção pode compreender um ou mais, por exemplo, dois, três ou mesmo mais sequências de DNA genômicas não traduzida a jusante de um promotor GAPDH eucariótico ou fragmentos dos mesmos, e/ou um ou mais, por exemplo, dois, três ou mesmo mais sequências de DNA genômico não traduzido a montante de um promotor GAPDH eucariótico ou fragmentos dos mesmos. Se o cassete de expressão a presente invenção compreende mais do que uma sequência de DNA a jusante e/ou a montante de um promotor GAPDH eucariótico ou fragmentos dos mesmos, estas sequências de DNA podem ser

diretamente ligadas, ou seja, podem compreender sequências ligantes, por exemplo, sequências ligantes contendo sítios de restrição que são liadas às extremidades 5' e 3' e que permitem a clonagem sequencial confortável da sequência ou fragmentos dos mesmos. Alternativamente, as sequências de DNA a jusante e/ou a montante de um promotor GAPDH eucariótico ou fragmentos dos mesmos, podem não ser diretamente ligados, ou seja, podem ser clonados com as sequências de intervenção do DNA.

[0019] O termo “sequência de polinucleotídio codificante de um polipeptídeo” como utilizado aqui inclui o DNA codificante para a um gene, preferivelmente, um gene heterólogo expressando o polipeptídeo.

[0020] Os termos “sequência codificante heteróloga”, “sequência de gene heterólogo”, “gene heterólogo”, “gene recombinante”, ou “gene” são utilizados intercaladamente. Estes termos refere-se a uma sequência de DNA que codifica para um recombinante, em particular, um produto da proteína heteróloga recombinante que é exigido para ser expresso em uma célula hospedeira, preferivelmente, em célula de mamífero e colhido. O produto do gene pode ser um polipeptídeo. A sequência de gene não está naturalmente presente na célula hospedeira e é derivado de um organismo do mesmo ou uma espécie diferente e pode ser geneticamente modificada.

[0021] Os termos “proteína” e “polipeptídeo” são utilizados intercaladamente para incluir uma série de resíduos de aminoácidos conectados a outros por meio de ligações de peptídeos entre os grupos alfa-amino e carboxi de resíduos adjacentes.

[0022] O termo “sequência de DNA genômico não traduzido”

como utilizado aqui inclui DNA que constitui informação genética de um organismo. O genoma de quase todos os organismos é DNA, a única exceção sendo algumas viroses que tem um RNA genômico. As moléculas DNA genômica em muitos organismos são organizados dentro dos complexos de DNA-proteína chamados cromossomos. O tamanho, número de cromossomos e a natureza de DNA genômico variam entre organismos diferentes. Os genomas de DNA viral podem ser de fita única, ou fita dupla, linear ou circular. Todos os outros organismos têm genomas de DNA de fita dupla. A bactéria tem um cromossomo circular simples. Nos eucarióticos, muito DNA genômico está localizado dentro dos núcleos (DNA nuclear) como cromossomos lineares múltiplos de tamanhos diferentes. As células eucarióticas contêm, adicionalmente, o DNA genômico na mitocôndria e, em plantas e eucarióticos inferiores, cloroplasto. Este DNA é usualmente uma molécula circular e está presente como cópias múltiplas dentro destas organelas. Uma sequência de DNA genômica não traduzida é, normalmente, não operativamente ligada para um promotor e, assim, não é traduzido. Ela pode conter gene(s) que não é(são) traduzido, assim, o(s) gene(s) que codifica, por exemplo, uma proteína que não é expressa.

[0023] O termo "sequência de DNA genômico não traduzido a jusante de um promotor GAPDH eucariótico" como utilizado aqui corresponde ao DNA genômico eucariótico não traduzido na extremidade 3' de um promotor GAPDH eucariótico. A sequência de DNA genômico não traduzido a jusante de um promotor GAPDH eucariótico normalmente inicia na posição de nucleotídeo em torno de +1, preferivelmente, na posição de nucleotídeo +1, onde a posição de nucleotídeo é relativa ao início de

transcrição de um mRNA de GAPDH, ou seja, é relativa à origem do início da transcrição do gene eucariótico para GAPDH. A sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico é usualmente da mesma origem que a do promotor GAPDH eucariótico, por exemplo, se o promotor GAPDH é de origem humana ao da sequência de DNA genômico não traduzido a jusante do promotor GAPDH humano é de origem bem humana e corresponde a sequência de DNA genômico humano de ocorrência natural a jusante do promotor GAPDH humano.

[0024] O termo "sequência de DNA genômico não traduzido a montante de um promotor GAPDH eucariótico" como utilizado aqui corresponde ao DNA genômico eucariótico não traduzido na 5' de um promotor GAPDH eucariótico. A sequência de DNA genômico não traduzida de um promotor GAPDH eucariótico normalmente inicia em uma posição de nucleotídeo imediatamente após da extremidade 5' do promotor GAPDH eucariótico. A sequência de DNA genômico não traduzida a montante de um promotor GAPDH eucariótico é usualmente da mesma origem que a do promotor GAPDH eucariótico, por exemplo, se o promotor GAPDH é de origem humana, a sequência de DNA genômico não traduzida a montante do promotor GAPDH humano é bem de origem humana e corresponde à sequência de DNA genômico de ocorrência naturalmente humana do promotor GAPDH humano.

[0025] As posições do promotor GAPDH eucariótico, a sequência de DNA genômico não traduzido à jusante ou à montante do promotor GAPDH eucariótico e outras sequências de DNA como indicada aqui são relativas ao início da transcrição do mRNA de GAPDH, por exemplo, são relativas à origem do início da transcrição do GAPDH eucariótico se não

especificamente de outro modo indicado.

[0026] O termo "sequência de DNA genômico não traduzido à montante de um promotor GAPDH eucariótico se estende a" ou "sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico se estende a" é utilizada para definir a extensão do comprimento de sequência DNA genômico não traduzido a montante e/ou a jusante de um promotor de GAPDH eucariótico a partir do início para um elemento genético particular, por exemplo, extensão para um ítron ou uma parte do mesmo.

[0027] O promotor de GAPDH eucariótico e o DNA genômico eucariótico a montante e/ou a jusante do promotor GAPDH pode ser encontrado para humanos, ratos e camundongos no banco de dados públicos NCBI (Gene GAOPDH de entrada para humanos, camundongos, ratos e Hamsters Chineses são genes Ids 2597 (mRNA: NM\_002046,3), 14433 (mRNA: NM\_008084,2), 24383 (mRNA: NM\_017008.3) e 100736557 (mRNA: NM\_001244854.2), respectivamente; "National Center for Biotechnology Information" (NCBI): <http://www.ncbi.nlm.nih.gov/>) e são exemplificativamente mostrados na Figura 7 e 8 para o gene GAPDH humano.

[0028] O promotor GAPDH eucariótico é usualmente considerado para alcançar em torno de -500 bps para +50 em relação ao início da transcrição do mRNA de GAPDH. O promotor GAPDH humano está localizado no cromossomo 12. O promotor GAPDH humano é considerado por Graven et al. (Graven et al., 1999) "Biochimica et Biophysics Acta", 147:203-218) para alcançar de -488 bps a +20 em relação ao início da transcrição do mRNA GAPDH com base em um estudo de fragmentação. De acordo com o banco de dados públicos NCBI do

promotor GAPDH humano alcança de -462 bps a +46 em relação ao início da transcrição de mRNA GAPDH como definido pelo banco de dados públicos NCBI. Se não especificamente de outro modo indicado, o promotor GAPDH humano como referido aqui alcança de -462 para a posição +46 em relação ao inicio da transcrição do mRNA GAPDH que corresponde à sequência alcançando de 4071 bps a 4578 da SEQ ID NO:17.

[0029] A numeração utilizada para o DNA do gene GAPDH, o gene IFF01 e o gene NCAPD2 de origem humana, de camundongo e de rato como referido aqui corresponde ao número utilizado para estes genes no banco de dados públicos NCBI (<http://www.ncbi.nlm.nih.gov/>).

[0030] O termo "promotor" como utilizado aqui define uma sequência de DNA regulatório geralmente localizado a montante de um gene que media o inicio da transcrição por direção da polimerase RNA para ligação ao DNA e inibe a síntese de RNA.

[0031] O termo "melhorador" como utilizado aqui define uma sequência de nucleotídeo que age para potencializar a transcrição dos genes independentes da identidade do gene, a posição da sequência na relação ao gene, ou a orientação da sequência. Os vetores da presente invenção incluem, opcionalmente, melhoradores.

[0032] Os termos "ligados funcionalmente" e "operativamente ligados" são utilizados intercaladamente e se refere a uma relação funcional entre dois ou mais segmentos de DNA, em particular as sequências do gene para ser expressos e aquelas sequências controlando sua expressão. Por exemplo, um promotor e/ou sequência melhoradora, incluindo qualquer combinação de elementos de controle de transcrição com ação cis é, operativamente ligado a uma sequência codificante se

ela estimula ou modula a transcrição da sequência codificante em uma célula hospedeira apropriada ou outro sistema de expressão. As sequências regulatórias promotoras que são operativamente ligadas à sequência de gene transcrita são fisicamente contínuas a sequência transcrita.

[0033] A "orientação" se refere à ordem de nucleotídeos em uma determinada sequência de DNA. Por exemplo, uma orientação de uma sequência de DNA em direção oposta em relação a uma outra sequência de DNA é uma na qual a extremidade 5' para 3' de modo que a sequência em relação a uma outra sequência é revertida quando comparada a um ponto da referência no DNA a partir do qual a sequência foi obtida. Os referidos pontos de referência pode incluir a direção da transcrição de outras sequências de DNA especificadas na fonte de DNA e/ou a origem de replicação dos vetores replicáveis contendo a sequência.

[0034] O termo "vetor de expressão" como utilizado aqui inclui uma molécula de DNA isolada ou purificada que durante a transfecção dentro de uma célula hospedeira apropriada provê um nível alto de expressão de um produto de gene recombinante dentro da célula hospedeira. Em adição, a sequência de DNA codificante para o recombinante ou produto de gene do vetor de expressão compreende as sequências de DNA regulatória que são requeridas para uma transcrição eficiente da sequência codificante do DNA no mRNA e para uma tradução eficiente do mRNAs dentro das proteínas nas linhas de célula hospedeiras.

[0035] Os termos "célula hospedeira" ou "linha de célula hospedeira" como utilizada aqui incluem quaisquer células, em particular células mamíferas, que são capazes de crescimento em cultura e expressão de uma proteína de produto

recombinante desejado.

[0036] O termo "fragmento" como utilizado aqui inclui uma porção da respectiva sequência de nucleotídeo, por exemplo, uma porção da sequência de DNA genômico não traduzida a jusante e/ou a montante de um promotor GAPDH eucariótico ou uma porção da sequência de nucleotídeo codificante de um elemento genético particular tal como um promotor. Os fragmentos de uma sequência de DNA genômico não-traduzida à jusante e/ou a montante de um promotor GAPDH eucariótico pode reter a atividade biológica e assim alterar, por exemplo, o aumento do padrão de expressão das sequências codificantes operativamente ligadas a um promotor. Os fragmentos de uma sequência de DNA genômico não traduzida a jusante e/ou a montante de um promotor GAPDH eucariótico pode variar de pelo menos cerca de 100 a cerca de 3000 bp, preferivelmente, de cerca de 200 a cerca de 2800 bp, mais preferivelmente, de cerca de 300 a cerca de 2000 bp nucleotídeos, em particular de cerca de 500 a cerca de 1500 bp nucleotídeos. De modo a clonar os fragmentos da sequência de DNA genômico não traduzida a jusante e/ou a montante de um promotor GAPDH eucariótico no cassete de expressão da presente invenção, usualmente as sequências ligantes contendo sítios de restrição que permite a clonagem confortável são ligados nas extremidades 5' - e 3' - dos fragmentos.

[0037] O termo "identidade da sequência de nucleotídeos" ou "sequência de nucleotídeos idêntica" como aqui utilizada incluem a porcentagem de nucleotídeos na sequência candidato que são idênticas com a sequência de nucleotídeo de, por exemplo, na sequência de DNA genômico não traduzida a jusante e/ou a montante de um promotor GAPDH eucariótico, após

alinhamento das sequências e introdução de lacunas, se necessário, para conseguir a identidade da sequência da porcentagem máxima. Assim, a identidade de sequência pode ser determinada pelos métodos padrões que são comumente utilizados para comparar a similaridade na posição dos nucleotídeos de duas sequências de nucleotídeos. Usualmente, a sequência de nucleotídeo identifica da sequência de candidato de duas sequências de nucleotídeo. Usualmente, a identidade da sequência de nucleotídeo da sequência candidato para a sequência de DNA genômico não traduzido à jusante e/ou a jusante de um promotor GAPDH eucariótica é pelo menos 80%, preferivelmente, pelo menos 85%, mais preferivelmente pelo menos 90%, e mais preferivelmente, pelo menos 95%, em particular 96%, mais particularmente 97%, ainda mais particularmente 98%, mais particularmente 99%, incluindo, por exemplo, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 94%, 95%, 96%, 97%, 98%, 99%, e 100%.

[0038] O termo "sítio CpG" como utilizado aqui inclui regiões de DNA onde um nucleotídeo citosina ocorre próximo a um nucleotídeo guanina na sequência linear das bases ao longo de seu comprimento. "CpG" é taquigrafado para "-C-fosfato-G-", ou seja, citosina e guanina separado apenas por um fosfato; o fosfato liga qualquer dois nucleosídeos juntos no DNA. A observação "CpG" é utilizado para distinguir esta sequência linear a partir o par de base CG da citosina e guanina.

[0039] O termo "uso do códon alternativo" como utilizado aqui inclui o uso de códons alternativos codificantes para o mesmo aminoácido de modo a evitar a sequência da porção CpG.

Isto inclui usando códons preferivelmente não tendo um sítio de CpG interno (por exemplo, GCG codificante para Alanina e contendo um sítio CpG, pode ser substituído tanto por GCT, GCC ou GCA) bem como evitando a união de dois códons que conduz a um novo sítio CpG.

[0040] O termo "em torno" como utilizado aqui em relação ao comprimento de uma sequência de DNA e em relação a uma posição de nucleotídeo que é relativa ao início da transcrição do mRNA GAPDH por exemplo, é relativa à origem do início da transcrição do GAPDH eucariótico inclui valores com desvio de um máximo de  $\pm 50\%$ , usualmente de um máximo de  $\pm 10\%$  dos valores declarados, por exemplo, "nucleotídeos em torno de 3000" inclui valores de 2700 a 3300 nucleotídeos, preferivelmente de 2900 a 3100 nucleotídeos, mais preferivelmente, 2995 a 3005 nucleotídeos, "em torno de 100 nucleotídeos" inclui valores de 50 a 150 nucleotídeos, preferivelmente, de 90 a 110 nucleotídeos, mais preferivelmente, de 95 a 105 nucleotídeos, "em torno de 15000 nucleotídeos" incluem valores de 13500 para 16500 nucleotídeos, preferivelmente, 14500 a 15500 nucleotídeos, mais preferivelmente, de 14990 a 15010 nucleotídeos, mais preferivelmente, de 14995 a 15505 nucleotídeos, "em torno de 200 nucleotídeos" incluem valores de 150 para 250 nucleotídeos, preferivelmente, 190 a 210 nucleotídeos, mais preferivelmente, de 195 a 205 nucleotídeos; "em torno de 8000 nucleotídeos" incluem valores de 7200 para 8800 nucleotídeos, preferivelmente, 7500 a 8500 nucleotídeos, mais preferivelmente, de 7990 a 8010 nucleotídeos, mais preferivelmente, de 7995 a 8005 nucleotídeos; "em torno de 500 nucleotídeos" incluem valores de 450 para 550

nucleotídeos, preferivelmente, 475 a 525, mais preferivelmente, de 490 a 510 nucleotídeos, mais preferivelmente, de 495 a 505 nucleotídeos; "em torno de 5000 nucleotídeos" incluem valores de 4500 para 5500 nucleotídeos, preferivelmente, 4750 a 5250, mais preferivelmente, de 4990 a 5010 nucleotídeos, mais preferivelmente de 4995 a 5005 nucleotídeos; "em torno de 1000 nucleotídeos" incluem valores de 900 para 1100 nucleotídeos, preferivelmente, 950 a 1050 nucleotídeos, mais preferivelmente, de 990 a 1010 nucleotídeos, mais preferivelmente, de 995 a 1005 nucleotídeos; "em torno de 4500 nucleotídeos" incluem valores de 4050 para 4950 nucleotídeos, preferivelmente, 4250 a 4750 nucleotídeos, mais preferivelmente, de 4490 a 4510 nucleotídeos, mais preferivelmente, de 4495 a 4505 nucleotídeos; "em torno de 1500 nucleotídeos" incluem valores de 1350 para 1650 nucleotídeos, preferivelmente, 1450 a 1550 nucleotídeos, mais preferivelmente, de 1490 a 1510 nucleotídeos, mais preferivelmente, de 1495 a 1505 nucleotídeos; "em torno de 5000 nucleotídeos" incluem valores de 3600 para 4400 nucleotídeos, preferivelmente, 3800 a 4200 nucleotídeos, mais preferivelmente, de 3990 a 4010 nucleotídeos, mais preferivelmente, de 3995 a 4005 nucleotídeos; "em torno de 2000 nucleotídeos" incluem valores de 1800 para 2200 nucleotídeos, preferivelmente, 1900 a 2100 nucleotídeos, mais preferivelmente, de 1990 a 2010 nucleotídeos, mais preferivelmente, de 1995 a 2005 nucleotídeos; "em torno de 3500 nucleotídeos" incluem valores de 3150 para 3850 nucleotídeos, preferivelmente, 3300 a 3700 nucleotídeos, mais preferivelmente, de 3490 a 3510 nucleotídeos, mais preferivelmente, de 3495 a 3505

nucleotídeos; "em torno de 2700 nucleotídeos" incluem valores de 2430 para 2970 nucleotídeos, preferivelmente, 2600 a 2800 nucleotídeos, mais preferivelmente, de 2690 a 2710 nucleotídeos, mais preferivelmente, de 2695 a 2705 nucleotídeos; "em torno de 3300 nucleotídeos" incluem valores de 2970 para 3630 nucleotídeos, preferivelmente, 3100 a 3500 nucleotídeos, mais preferivelmente, de 3290 a 3310 nucleotídeos, mais preferivelmente, de 3295 a 3305 nucleotídeos; "em torno de 3200 nucleotídeos" incluem valores de 2880 para 3520 nucleotídeos, preferivelmente, 3000 a 3400 nucleotídeos, mais preferivelmente, de 3190 a 3210 nucleotídeos, mais preferivelmente, de 3195 a 3205 nucleotídeos; em torno de +7000 ou na posição em torno de +7000 incluem posições +6300 para +7700, preferivelmente, nas posições +6700 a +7300, mais preferivelmente, posições +6990 para +7010, mais preferivelmente, nas posições +6995 para +7005, em torno de +1 ou em torno da posição +1 inclui posições -10 para +10, preferivelmente, nas posições -5 a +5, mais preferivelmente, nas posições -1 para +2, em torno de -3500 ou em torno da posição -3500 incluem posições -3150 a -3850, preferivelmente nas posições -3300 para -3700, mais preferivelmente, nas posições -3490 a -5010, mais preferivelmente nas posições -3495 a -3505. O termo "em torno" como utilizado aqui em relação à numeração utilizada par ao DNA do gene GAPDH, o gene IFF01 e o gene NCAPD2 de origem humana, de camundongo e de ratos como referido aqui ou utilizado aqui em relação á uma posição em uma sequência de um número SEQ ID inclui valores com desvio de um máximo de +500 bps, preferivelmente, +100 bps, mais preferivelmente, +10 bps, mais preferivelmente ± 5 bps.

[0041] Em uma concretização, a presente descrição provê um cassete de expressão que compreende um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico, onde o polipeptídeo codificado pela sequência de polinucleotídio não é GAPDH, e sendo que a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo da posição de nucleotídeo em torno de \_1 para a posição de nucleotídeo em torno de + 7000, sendo que a posição de nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico está em torno de 100 para 1500 nucleotídeos. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em sua máxima para o segundo dos últimos ítron do gene IFF01 ou para parte do mesmo. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em sua máxima para o último ítron do gene IFF01.

[0042] O gene IFF01 humano é localizado no DNA humano em torno de 6665249 bps para 6648694 do cromossomo 12 (NCBI ID do gene: 25900). Em uma concretização, o comprimento da sequência DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico se estendendo em seu máximo do último ítron do gene IFF01 em humanos alcança em seu máximo para em torno de 6650677 bps de cromossomo 12 codificante para o gene IFF01 em humano (posição + 7021). Em uma concretização, o

comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se em sua máxima para o penúltimo ítron do gene IFF01 em humanos alcança em seu máximo para em torno de 6657230 bps do cromossomo 12 codificante para o gene IFF01 em humano (posição +13574). As sequências de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron do gene IFF01 em humano e para o segundo dos últimos ítrons do gene IFF01 em humanos, respectivamente, estão incluídos na SEQ ID NO: 17, que mostra 6657230 bps para 6639125 bps do cromossomo 12 (NCBI ID do gene: 25900). A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao último ítron que alcança em torno de 11553 bps da sequência de nucleotídeo como mostrado pela SEQ ID NO: 17 e a sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico estendendo-se ao segundo dos últimos ítrons alcança em torno de 18106 bps da sequência de nucleotídeo como mostrado pela SEQ ID NO:17.

[0043] O gene IFF01 de camundongo (NCBI, ID do gene: 320678) está localizado no DNA de camundongo, em torno de 125095259 bps para 125111800 do cromossomo 6. Em uma concretização, no comprimento da sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico se estendendo em seu máximo para o último ítron do gene IFF01 em camundongos alcança em seu máximo para em torno de 125109211 bps do cromossomo 6 codificante para o gene IFF01 em camundongo (posição +6391). Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico se estendendo do

segundo do último ítron do gene IFF01 em camundongos alcança em seu máximo em torno de 125103521 bps do cromossomo 6 codificante para o gene IFF01 em camundongos (posição +12081). A sequência de DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico estendendo-se em seu máximo ao último ítron e para o segundo dos últimos ítrons do gene IFF01 no camundongo, respectivamente, são incluídos em SEQ ID NO:18 que mostra os 125103521 para 125119832 do cromossomo 6 (NCBI, ID do gene: 3206778). A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico se estendendo ao último ítron do gene IFF01 no camundongo alcança em torno de 10622 bps da sequência de nucleotídeo como mostrado pela SEQ ID NO:18 e a sequência de DNA genômico não traduzido à jusante do promotor GAPDH eucariótico estendendo-se ao penúltimo ítron do gene IFF01 em camundongos alcança em torno de 16312 bps da sequência de nucleotídeo como mostrado pela SEQ ID NO:18.

[0044] O gene IFF01 de rato (NCBI, ID do gene: 362437) está localizado no DNA de rato em torno de 16124966 bps para 161282150 bps do cromossomo 4. Em uma concretização, o comprimento da sequência do DNA genômico não traduzido à jusante de um promotor GAPDH eucariótico se estende em seu máximo para o último ítron do gene IFF01 em ratos alcança em seu máximo em torno de 161280937 bps do cromossomo 4 codificante para o gene IFF01 em ratos (posição +5154). Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se em sua máxima para o penúltimo ítron IFF01 em ratos alcança em seu máximo para em torno de 161279451 bps do cromossomo 4 codificante para o gene IFF01 em ratos (posição

+6640).

[0045] As sequências de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico se estendendo em seu máximo para o último ítron e para o segundo do último ítron do gene IFF01 em ratos, respectivamente, estão incluídos na SEQ ID NO:19, que mostra 161279451 bps para 161290508 do cromossomo 4 (NCBI, gene ID: 362437). A sequência do DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico se estendendo ao último ítron do gene IFF01 alcança em torno de 9572 bps das sequências de nucleotídeos como mostrado pela SEQ ID NO:19 e a sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico se estendendo ao penúltimo ítron do gene IFF01 alcança em torno de bps 11058 bps da sequência de nucleotídeos como mostrado pela SEQ ID NO:19.

[0046] O gene IFF01 do hamsters Chineses (NCBI, ID do gene: 100753382) está localizado no DNA de hamsters Chineses em torno de 3577293 bps para 3593683 bps. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor de GAPDH eucariótico em seu máximo para o último ítron do gene IFF01 em hamsters Chineses alcança em seu máximo em torno de 3579014 codificante para o gene IFF01 em hamsters Chineses (posição +6883). Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o penúltimo ítron do gene IFF01 em hamsters chineses em seu máximo, em torno de 3585061 bps codificante para o gene IFF01 em hamsters Chineses (posição +12930). A localização cromossomal não foi ainda anotada no banco de dados NCBI e a informação da sequência

atual contém muitas bases desconhecidas. Portanto, a anotação precisa dos limites pode alterar com a disponibilidade de informações de sequência mais apuradas.

[0047] As sequências de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron e para o penúltimo ítron do gene IFF01 no hamsters chinês, respectivamente estão incluídos na SEQ ID NO:29 que demonstra 3567932 bps para 3585061. A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao último ítron do gene IFF01 alcança em torno de 11083 bps da sequência de nucleotídeos como mostrado na SEQ ID NO:29 e a sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se para o penúltimo ítron do gene IFF01 para em torno de 17130 bps da sequência de nucleotídeos como mostrado pela SEQ ID NO:29.

[0048] Em uma concretização adicional, a sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico inicia no sítio de poliadenilação GAPDH eucariótico, por exemplo, inicia no primeiro nucleotídeo codificante do sítio de poliadenilação GAPDH eucariótico. Preferivelmente, a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia à jusante do sítio de poliadenilação GAPDH eucariótico, por exemplo, inicia imediatamente após o último nucleotídeo codificante do sítio de poliadenilação de GAPDH eucariótico. Ainda mais preferido, a sequência de DNA genômico não traduzido à jusante do promotor GAPDH eucariótico inicia à jusante do sítio de poliadenilação GAPDH eucariótico e o comprimento da sequência de DNA genômico não traduzida à jusante do promotor

GAPDHA eucariótica é pelo menos em torno de 100 nucleotídeos e estende-se em seu máximo ao penúltimo ítron do gene IFFO1. [0049] Em uma concretização, a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo da posição do nucleotídeo em torno de +3881 para a posição do nucleotídeo em torno de +5000, preferivelmente, dentro de uma região abrangendo da posição do nucleotídeo em torno de 3931 para a posição do nucleotídeo em torno de +5000, mais preferivelmente, dentro de uma região abrangendo da posição de nucleotídeo em torno de +4070 para a posição de nucleotídeo em torno de +5000, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA do GAPDH. Uma sequência de DNA genômica não traduzida à jusante do promotor eucariótico GAPDH que inicia, por exemplo, à jusante do sítio de poliadenilação GAPDH eucariótico utilizado na presente invenção usualmente inicia em uma posição de nucleotídeo em torno da posição +3931, preferivelmente, em uma posição de nucleotídeo em torno de +4070, onde a posição do nucleotídeo é relativa ao inicio da transcrição do mRNA de GAPDH.

[0050] Na sequência do DNA genômico não traduzida à jusante do sítio de poliadenilação GAPDH humana inicia em torno da posição de nucleotídeo +3931 (em relação ao inicio da transcrição do mRNA de GAPDH que corresponde ao 8463 bp como mostrado na SEQ ID NO:17). Preferivelmente, se a sequência de DNA genômico não traduzida à jusante do sítio de poliadenilação GAPDH é de humano, a sequência de DNA genômica não traduzida à jusante do sítio de poliadenilação GAPDH inicia em torno de +3931 (relativa ao inicio de transcrição do mRNA de GAPDH; que corresponde ao 8463 bp como mostrado na

SEQ ID NO:17) e seu comprimento é em torno de 3357 bps correspondente à sequência em torno de 8463 bps em torno de 11819 como mostrado na SEQ ID NO: 17, mais preferivelmente, ele inicia em torno de +4070 (relativo ao inicio da transcrição do mRNA de GAPDH que corresponde ao 8602 bp como mostrado na SEQ ID NO;17) e seu comprimento é em torno de 3218 bps correspondente à sequência em torno de 8602 bps para em torno de 11819 como mostrado na SEQ ID NO:17.

[0051] Em uma concretização adicional, a sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico compreende a sequência de nucleotídeo selecionado a partir do grupo consistindo das SEQ ID NOS: 8 e 21 ou fragmentos dos mesmos.

[0052] Em uma concretização adicional, a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico compreende uma sequência de nucleotídeo pelo menos 80% idêntica para a sequência de nucleotídeo selecionado a partir do grupo consistindo da SEQ ID NO: 8 e 21 ou fragmentos dos mesmos.

[0053] Em algumas concretizações, a sequência de nucleotídeo selecionado a partir do grupo consistindo das SEQ ID NOS: 8 e 21 ou fragmentos os mesmos, compreende cinco ou menos, preferivelmente, quatro ou menos, mais preferivelmente, três ou menos, mais preferivelmente, dois ou menos, em particular uma modificação no ácido nucléico, onde a modificação de ácido nucléico são preferivelmente de uma substituição de ácido nucléico.

[0054] Em uma concretização adicional, o comprimento da sequência de DNA genômica não traduzida à jusante do promotor GAPDH eucariótico é, preferivelmente, em torno de 200 a 8000

nucleotídeos, mais preferivelmente, em torno de 500 para 5000 nucleotídeos, ainda mais preferivelmente, em torno de 1000 para 4500 nucleotídeos, mais preferivelmente, em torno de 1500 em torno de 4000 nucleotídeos, em particular, em torno de 2000 em torno de 3500 nucleotídeos, mais particular em torno de 2700 em torno de 3300, ainda mais particularmente em torno de 3200, mais particularmente 3218 nucleotídeos. O comprimento da sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico como definido aqui não inclui qualquer sequência ligante adicionada à sequência de DNA genômico não traduzida.

[0055] Em uma concretização adicional, a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico é orientada na mesma direção que a da sequência de polinucleotídio codificante de um polipeptídeo.

[0056] Em uma concretização adicional, a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico é orientada na direção oposta em relação à da sequência de polinucleotídio codificante de um polipeptídeo.

[0057] Em algumas concretizações, o cassete de expressão que compreende um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômica não traduzida à jusante de um promotor GAPDH eucariótica compreende adicionalmente, uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico, onde a sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição do nucleotídeo em torno de - 3500, onde a posição de nucleotídeo é relativa ao início da

transcrição do mRNA de GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é em torno de 100 para 15000 nucleotídeos.

[0058] Em uma concretização adicional, o cassete de expressão compreende ainda uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico, sendo que a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo da posição de nucleotídeo em torno de +1 para a posição do nucleotídeo em torno de +7000, sendo que a posição de nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, e sendo que o comprimento da sequência de DNA genômico à jusante do promotor de GAPDH eucariótico é em torno de 100 para 15000 nucleotídeos. Nestas concretizações, a sequência de DNA genômica não traduzida à jusante de um promotor GAPDH eucariótico utilizado, por exemplo, como descrito *supra*.

[0059] Em algumas concretizações, o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é, preferivelmente, em torno de 200 para 8000 nucleotídeos, mais preferivelmente, em torno de 500 para 5000 nucleotídeos, ainda mais preferivelmente, em torno de 1000 para 4500 nucleotídeos, mais preferivelmente, em torno de 1500 para 4000 nucleotídeos, em particular em torno de 2000 para 3500 nucleotídeos, mais particularmente em torno de 2700 para 3300, ainda mais particularmente em torno de 3200, mais particularmente 3158 nucleotídeos de comprimento. O comprimento da sequência de DNA genômico não traduzido à montante do promotor GAPDH eucariótico como definido aqui não incluem quaisquer sequências ligantes

adicionais à sequência de DNA genômico não traduzida.

[0060] Em uma concretização adicional, o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em seu máximo para iniciar o códon do gene NCAPD2. Em uma concretização adicional, o comprimento da sequencia de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em seu máximo para antepenúltimo íntron do gene NCAPDE2. Em uma concretização adicional, o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em seu máximo para o penúltimo íntron do gene NCAPD2. Em uma concretização adicional, o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é pelo menos em torno de 100 nucleotídeos e se estende em seu máximo para o último íntron do gene NCAPD2.

[0061] O gene NCAPD2 humano (NCBI, ID do gene: 9918) está localizado em DNA humano em torno de 6603298 bps para 6641132 bps do cromossomo12. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico se estende em seu máximo para o último íntron do gene NCAPD2 em humanos alcança em seu máximo em torno de 6640243 bps do cromossomo12 codificante para o gene NCAPD2 em humanos (posição -3414 relativo ao início da transcrição do gene GAPDH que corresponde ao 1119 bp na SEQ ID NO;17).

[0062] Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH

eucariótico estendendo-se em seu máximo ao penúltimo ítron do gene NCAPD2 em humanos alcança em seu máximo, em torno de 6639984 bps do cromossomo 12 codificante para o gene NCAPD2 em humanos (posição -3673 relativa ao início da transcrição do gene GAPDH que corresponde ao 860 bp na SEQ ID NO: 17).

[0063] Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo ao antepenúltimo ítron do gene NCAPD2 em humanos alcança em seu máximo, em torno de 6639125 bps do cromossomo 12 codificante para o gene NCAPD2 em humanos (posição -4532 relativa ao inicio da transcrição do gene GAPDH que corresponde a 1 bp na SEQ ID NO: 17).

[0064] A sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo do último ítron, para o penúltimo ítron e para o antepenúltimo ítron do gene NCAPD2 em humanos, respectivamente, estão incluídos SEQ ID NO: 17, na qual mostra 6657230 bps para 6639125 do cromossomo 12 (NCBI, ID do gene: 9918).

[0065] O gene NCAPD2 de camundongo (ID gene: 68298) está localizado em DNA de camundongo em torno da posição 125118025 para 125141604 do cromossomo 6. Em uma concretização, o comprimento, do comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico (estimado por ter um comprimento de 500 bps à montante do início de transcrição) estendendo-se em seu máximo para em torno de 124118607 do cromossomo 6 codificante para o gene NCAPD2 em camundongos.

[0066] Em uma concretização, o comprimento da sequência de

DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o penúltimo ítron do gene NCAPD2 em camundongos alcança em seu máximo em torno de 125118880 bps do cromossomo 6 codificante para o gene NCAPD2 em camundongos.

[0067] Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o antepenúltimo ítron do gene NCAPD2 em camundongos alcança em seu máximo em torno 125119832 bps do cromossomo 6 codificante para o gene NCAPD2 em camundongos.

[0068] A sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron, para o penúltimo ítron e para o antepenúltimo ítron do gene NCAPD2 em camundongos, respectivamente, estão incluídos na SEQ ID NO:18, que mostra 125103521 bps para 125119832 bps do cromossomo 6 (NCBI, ID do gene: 68298). A sequência de DNA genômico não traduzido à montante de um promotor GAPDH eucariótico estendendo-se para o último ítron para em torno de 1226 da sequência de nucleotídeo como mostrado pela SEQ ID NO:18 (-3006 relativa ao início da transcrição do mRNA de GAPDH de camundongo). A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se para o penúltimo ítron alcança em torno de bps 953 da sequência de nucleotídeo como mostrado pela SEQ ID NO: 18 (-3279 relativa ao inicio da transcrição do mRNA GAPDH de camundongo). A sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico estendendo-se ao antepenúltimo ítron alcança em torno de 1 bp da sequência de nucleotídeo como

mostrado pelo SEQ ID NO:18 (-4231 relativa ao início da transcrição do mRNA GAPDH de camundongo).

[0069] O gene NCAPD2 de rato (ID do gene: 362438) está localizado no DNA eucariótico em torno da posição 161288671 para 161310417 do cromossomo 4. Em uma concretização, o comprimento da sequência de DNA genômico não traduzido à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron do gene NCAPD2 em ratos alcança em seu máximo em torno de 161289191 bps do cromossomo 4 codificante para o gene NCAPD2 em ratos. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o penúltimo do gene NCAPD2 em ratos alcança em seu máximo para em torno de 161289446 bps do cromossomo 4 codificante para o gene NCAPD2 em ratos. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o antepenúltimo ítron do gene NCAPD2 em ratos alcança em seu máximo, em torno de 161290508 bps do cromossomo 4 codificante para o gene NCAPD2 em ratos.

[0070] A sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron do gene NCAPD2 em ratos, respectivamente, estão incluídos na SEQ ID NO: 19, que mostram 161279451 para 161290508 bps do cromossomo 4 (NCBI, ID do gene: 362438). A sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótico estendendo-se para o último ítron alcança em torno de 1318 da sequência de nucleotídeo como mostrado pela SEQ ID NO:19

(-3101 relativa ao início da transcrição do mRNA GAPDH de ratos). A sequência de DNA genômica não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao penúltimo ítron alcança em torno de 1063 bps da sequência de nucleotídeos como mostrado pela SEQ ID NO:19 (Posição -3356 relativa ao início da transcrição do mRNA GAPDH de ratos). A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao antepenúltimo ítron alcança em torno de 1bp das sequências de nucleotídeo como mostrado pela SEQ ID NO: 19 (posição -4418 relativa ao início da transcrição de mRNA GAPDH de ratos).

[0071] O gene NCAPD2 de hamsters chineses (ID do gene: 100753087) está localizado na posição em torno de DNA eucariótico 3544184 para 3569879. A localização do cromossomo não está disponível no banco de dados NCBI. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron do gene NCAPD2 em hamsters GAPDH estendendo-se em seu máximo em torno de 3569380 bps em hamsters chineses. Em uma concretização, o comprimento da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo em seu máximo para o penúltimo ítron no gene NCAPD2 em hamsters chineses alcança em seu máximo em torno de 3569131 bps em hamsters chineses. Em uma concretização, o comprimento da sequência de DNA genômico não traduzido à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o antepenúltimo ítron do gene NCAPD2 em hamsters chineses alcança em seu máximo em torno de 3567932 bps em hamsters chineses.

[0072] A sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se em seu máximo para o último ítron, para o penúltimo ítron e para o antepenúltimo ítron do gene NCAPDH2 em hamsters chineses, respectivamente estão incluídos na SEQ ID NO:29, que mostra 3567932 bps para 3585061. A sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico estendendo-se para o último ítron alcança em torno de 1449 bps da sequência de nucleotídeos como mostrado por SEQ ID NO:29 (-2752 relativa para início da transcrição de mRNA GAPDH de hamsters chineses). A sequência de DNA genômica não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao penúltimo ítron alcança em torno de 1200 bps da sequência de nucleotídeo como mostrado por SEQ ID NO:29 (posição -3001 relativa ao início da transcrição do mRNA GAPDH de hamsters chinês). A sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico estendendo-se ao antepenúltimo ítron alcança para em torno de 1 bp das sequências de nucleotídeos como mostrado por SEQ ID NO:29 (posição -4200 relativa ao início da transcrição de mRNA GAPDH de hamsters chineses).

[0073] Em algumas concretizações, a sequência de DNA genômico não traduzida à montante de um promotor eucariótico GAPDH que inicia usualmente dentro de uma região abrangendo da posição do nucleotídeo em torno de -500 para a posição do nucleotídeo em torno de -3500, preferivelmente, dentro de uma região abrangendo da posição de nucleotídeo em torno de -576 para a posição de nucleotídeo em torno de -3500, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA GAPDH.

[0074] Em algumas concretizações, a sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico inicia usualmente em uma posição de nucleotídeo em torno da posição -500, preferivelmente, em uma posição de nucleotídeo em torno de -576, onde a posição do nucleotídeo é relativa ao início da transcrição do mRNA GAPDH.

[0075] Em humanos, a sequência de DNA genômico não traduzida à montante do promotor GAPDH humano inicia em torno da posição do nucleotídeo -463 (relativa ao inicio da transcrição do mRNA GAPDH que corresponde ao 4533 bp como mostrado na SEQ ID NO: 17). Preferivelmente, se a sequência de DNA genômico não traduzida à montante do promotor GAPDH é de humanos, a sequência de DNA genômico não traduzida à montante do promotor GAPDH inicia em torno de -500 (relativa ao inicio da transcrição do mRNA GAPDH; que corresponde ao 453 bp como mostrado na SEQ ID NO:17). Mais preferivelmente, se a sequência de DNA genômico não traduzida à montante do promotor GAPDH é de humano, a sequência de DNA genômico não traduzida à montante do promotor GAPDH inicia em torno de -576 (relativa ao inicio da transcrição do mRNA GAPDH; que corresponde ao 4533bp como mostrado na SEQ ID NO; 17) e seu comprimento é em torno de 3158 bps correspondente a sequência em torno de 800 bps em torno de 3957 como mostrado na SEQ ID NO: 17.

[0076] Em uma concretização adicional, a sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico compreende uma sequência de nucleotídeos selecionados a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 e 28 ou fragmentos das mesmas, preferivelmente, uma sequência de

nucleotídeos selecionado a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15 e 16 ou fragmentos das mesmas, ou uma sequência de nucleotídeos selecionados a partir do grupo consistindo das SEQ ID NOS: 20, 22, 23, 24, 25, 26, 27, 28 e 16 ou fragmentos das mesmas. Mais preferido é uma sequência de nucleotídeos selecionados a partir do grupo consistindo das SEQ ID NOS: 10, 12, 15 e 16 ou fragmentos das mesmas, mais preferivelmente, uma sequência de nucleotídeos selecionados a partir do grupo consistindo de SEQ ID NOS: 10, 12, 15 e 16 ou fragmentos das mesmas, onde as sequências de nucleotídeos compreendendo as SEQ ID NO: 10 e/ou 16 são orientadas em direções opostas em relação à sequência de polinucleotídios codificantes de um polipeptídeo, e sequências de nucleotídeos compreendendo SEQ ID NOS: 12 e/ou 15 são orientadas na mesma direção que as da sequência de polinucleotídio codificante de um polipeptídeo. Igualmente mais preferida é uma sequência selecionada a partir do grupo consistindo de SEQ ID NOS: 23, 25, 28 e 16 ou fragmentos das mesmas, mais preferivelmente uma sequência de nucleotídeo selecionado a partir do grupo consistindo das SEQ ID NOS: 23, 24, 28 e 16 ou fragmentos das mesmas, onde as sequências de nucleotídeo compreendendo SEQ ID NOS: 23 e/ou 16 são orientadas em direção oposta em relação a sequência de polinucleotídio codificante de um polipeptídeo, e sequências de nucleotídeos compreendendo SEQ ID NOS: 25 e/ou 28 são orientadas na mesma direção que as da sequência de polinucleotídio codificante de um polipeptídeo.

[0077] Em uma concretização adicional, a sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico compreende uma sequência de nucleotídeos

complementares para a sequência de nucleotídeo selecionada a partir do grupo consistindo de SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27, e 28 ou fragmentos das mesmas, preferivelmente, uma sequência de nucleotídeos complementares para as sequências de nucleotídeos selecionadas a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15 e 16 ou fragmentos das mesmas, ou uma sequência de nucleotídeos complementares para sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 20, 22, 23, 24, 25, 26, 27, 28 e 16 ou fragmentos das mesmas. Mais preferida é a sequência de nucleotídeos complementares para a sequência de nucleotídeos selecionadas a partir do grupo consistindo das SEQ ID NOS: 10, 12, 15 e 16 ou fragmentos das mesmas. Igualmente mais preferida é uma sequência de nucleotídeos complementar à sequência de nucleotídeos selecionada a partir do grupo consistindo das SEQ ID NOS: 23, 25, 28 e 15 ou fragmentos das mesmas.

[0078] Em uma concretização adicional, a sequência de DNA genômica não traduzida à montante do promotor GAPDH eucariótico compreende uma sequência de nucleotídeo pelo menos 80% idêntica à sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 e 28 ou fragmentos das mesmas, preferivelmente, uma sequência de nucleotídeos pelo menos 80% idêntica para a sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15 e 16 ou fragmentos das mesmas, ou uma sequência de nucleotídeos pelo menos 80% idêntica para a sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 20, 22, 23, 24, 2, 26, 27, 28 e

16 ou fragmentos das mesmas. Mais preferida é a sequência de nucleotídeos pelo menos 80% idêntica a sequência de nucleotídeos selecionada a partir do grupo consistindo das SEQ ID NOS: 10, 11, 12, 15 e 16 ou fragmentos das mesmas, mais preferivelmente, uma sequência de nucleotídeos pelo menos 80% idêntica para a sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 10, 12, 15 e 16 ou fragmentos das mesmas, sendo que sequência de nucleotídeos compreendendo as SEQ ID NOS: 10 e/ou 16 são orientadas em direção oposta em relação à sequência de polinucleotídios codificante de um polipeptídio, e sequências de nucleotídeo compreendendo SEQ ID NOS: 12 e/ou 15 são orientadas na mesma direção que as da sequência de polinucleotídio codificante de um polipeptídeo. Igualmente, mais preferida é uma sequência de nucleotídeo pelo menos 80% idêntica à sequência de nucleotídeo selecionada a partir do grupo consistindo das SEQ ID NOS: 23, 25, 28 e 16 ou fragmentos das mesmas, mais preferivelmente, uma sequência de nucleotídeos pelo menos 80% idêntica à sequência de nucleotídeos selecionada a partir do grupo consistindo das SEQ ID NOS: 23, 25, 28 e 16 ou fragmentos das mesmas, onde as sequências de nucleotídeos compreendendo as SEQ ID NOS: 23 e/ou 16 são orientadas na direção oposta em relação a sequência de polinucleotídio codificante de um polipeptídeo, e sequências de nucleotídeos compreendendo as SEQ ID NO: 25 e/ou 28 são orientadas na mesma direção que a sequência de polinucleotídios codificantes de um polipeptídio.

[0079] Em algumas concretizações, a sequência de nucleotídeos selecionados a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24,

25, 26, 27 e 28 ou fragmentos das mesmas, compreendem cinco ou menos, preferivelmente, quatro ou menos, mais preferivelmente três ou menos, mais preferidos dois ou mesmo, em particular uma modificação de ácido nucléico, onde a modificação de ácido nucléico, sendo que a modificação de ácido nucléico é preferivelmente, uma substituição de ácido nucléico.

[0080] Em algumas concretizações, a sequência de nucleotídeo selecionado a partir do grupo consistindo das SEQ ID NOS: 7, 9, 11, 14, 20, 22, 24 e 27 ou fragmentos das mesmas, compreendem cinco ou menos, preferivelmente, quatro ou menos, mais preferivelmente, três ou menos, mais preferivelmente, dois ou menos, em particular uma modificação de ácido nucléico, onde a modificação de ácido nucléico é preferivelmente, uma substituição do ácido nucléico.

[0081] Em algumas concretizações, a sequência de ácido nucléico selecionada a partir do grupo consistindo das SEQ ID NOS: 7, 9, 11, 14, ou fragmentos das mesmas, compreende uma substituição de ácido nucléico na posição 16 relativa ao início da sequência de nucleotídeo das SEQ ID NOS: 7, 9, 11, 14. Preferivelmente, G na posição 16 relativa ao início da sequência de nucleotídeo é substituído com T.

[0082] Em algumas concretizações, a sequência de nucleotídeo selecionado a partir do grupo consistindo das SEQ ID NOS: 20, 22, 24 e 27 ou fragmentos das mesmas, compreende uma substituição de ácido nucléico na posição 13 relativa ao início da sequência de nucleotídeo das SEQ ID NOS: 20, 22, 24 e 27. Preferivelmente, G na posição 13 relativa ao início da sequência de nucleotídeo que é substituído com T.

[0083] Em uma concretização adicional, a sequência de DNA

genômica não traduzida à montante do promotor GAPDH eucariótica é orientada na mesma direção como a sequência de polinucleotídio codificante de um polipeptídeo.

[0084] Em uma concretização adicional, a sequência de DNA genômica não traduzida à montante do promotor GAPDH eucariótico é orientada na direção oposta em relação a sequência de polinucleotídio codificante de um polipeptídeo.

[0085] Em uma concretização preferida, o cassete de expressão comprehende um promotor, uma sequência de polipeptídeo codificante de um polipeptídeo, e uma sequência de DNA genômica não traduzida à jusante de um promotor GAPDH eucariótico e uma sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótico como descrita *supra*. Preferivelmente, a origem da sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótico é a mesma, ou seja, é da mesma espécie. Mais preferivelmente, a origem da sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico, a sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótico e a célula hospedeira é a mesma, ou seja, é da mesma espécie, por exemplo, a origem da sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico, a sequência de DNA genômica não traduzida à montante de um promotor GAPDH eucariótica e a célula hospedeira é do mesmo mamífero, por exemplo, a partir de um humano.

[0086] Em algumas concretizações, se a sequência de DNA genômico não traduzida à jusante e/ou à montante do promotor GAPDH eucariótica é a sequência de DNA genômica não traduzida a partir de uma espécie, o promotor do cassete de expressão não é um promotor GAPDH a partir da mesma espécie.

[0087] Em algumas concretizações, se a sequência de DNA genômico não traduzida à jusante e/ou à montante do promotor GAPDH eucariótica é a sequência de DNA genômico não traduzido à jusante e/ou a montante da origem humana, o promotor do cassete de expressão não é promotor GAPDH humano.

[0088] Em algumas concretizações, o promotor do cassete de expressão não é um promotor GAPDH. Em uma concretização, se o cassete de expressão compreende um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômica não traduzida à jusante de um promotor de desidrogenase de gliceraldeído de 3-fosfato (GAPDH), onde o polipeptídeo codificado por uma sequência de polinucleotídio não é GAPDH, e onde a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótica inicia dentro de uma região abrangendo a partir da posição de nucleotídeo em torno de +1 para a posição de nucleotídeo em torno de +7000, onde a posição de nucleotídeo é relativa ao início de transcrição do mRNA GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico é de cerca de 100 a cerca de 15000 nucleotídeos e sendo que o cassete de expressão compreendendo adicionalmente de uma sequência DNA genômica não traduzida à montante de um promotor GAPDH eucariótico, onde a sequência de DNA genômico não traduzida à montante do promotor de GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição de nucleotídeo em torno de -3500, sendo que a posição de nucleotídeo é relativa ao início da transcrição do mRNA GAPDH, e sendo que o comprimento da sequência de DNA genômica não traduzida à montante do

promotor GAPDH eucariótico está em torno de 100 para em torno de 15000 nucleotídeos, o promotor do cassete de expressão pode ser um promotor GAPDH eucariótico, preferivelmente, um promotor GAPDH de mamífero, mais preferivelmente, um promotor GAPDH de roedor ou humano. Nesta concretização, a sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico iniciando dentro de uma região abrangendo da extremidade 5' do promotor GAPDH eucariótico para posição de nucleotídeo em torno de -3500 é preferivelmente localizado diretamente à montante do promotor GAPDH eucariótico, mais preferivelmente, nesta concretização, o cassete de expressão compreende uma sequência de DNA genômico de ocorrência natural compreendendo o promotor GAPDH eucariótico e estendendo-se para aposição e nucleotídeos em torno de -3500, sendo que a posição de nucleotídeo é relativa ao início da transcrição do mRNA GAPDH.

[0089] Em algumas concretizações, a sequência de DNA genômico não traduzida à jusante e/ou à montante do promotor GAPDH eucariótico é de origem de mamífero, por exemplo, o promotor de GAPDH eucariótico é um promotor GAPDH de mamífero e a sequência de DNA genômico não traduzida à jusante e/ou à montante do promotor GAPDH de mamífero são utilizadas como descrita aqui.

[0090] Em algumas concretizações, a sequência de DNA genômico não traduzida à jusante e/ou a montante do promotor GAPDH eucariótico é de origem de roedores ou humanos, por exemplo, o promotor GAPDH eucariótico é um promotor GAPDH de roedor ou humano e uma sequência de DNA genômica não traduzida à jusante e/ou a montante do promotor GAPDH humano ou de roedor é utilizado como aqui descrito.

[0091] Preferivelmente, a sequência de DNA genômico não traduzida à jusante e/ou a montante do promotor GAPDH eucariótico é selecionada a partir de origem humana, rato ou camundongo, mais preferivelmente, a partir da origem humana ou de camundongo, mais preferivelmente, a partir da origem humana.

[0092] Em algumas concretizações, a sequência de DNA genômica não traduzida à jusante e/ou a montante do promotor GAPDH eucariótico não é operativamente ligado à sequência de polinucleotídio codificante do polipeptídeo.

[0093] Em algumas concretizações, o cassete de expressão comprehende um sítio de poliadenilação. Preferivelmente, o sítio de poliadenilação é selecionado a partir do grupo consistindo de SV40 poli(A) e BGH (Hormônio de crescimento bovino) poli(A).

[0094] Em algumas concretizações, o promotor e a sequência de polinucleotídio codificante de um polipeptídeo do cassete de expressão são operativamente ligados.

[0095] Em algumas concretizações, o promotor do cassete de expressão é selecionado a partir do grupo consistindo do promotor SV40, promotor tk humano, promotor MPSV, CMV de camundongo, CMV humano, CMV de rato, alfa EF1 humano, EF1 alfa hamsters chinês, GAPDH humano, promotores híbridos, incluindo MYC, HYK e promotor CX.

[0096] Em algumas concretizações, o polipeptídeo codificado pelo cassete de expressão pode ser um polipeptídeo glicosilado e não-glicosilado. Os polipeptídios glicosilados se referem aos polipeptídios tendo pelo menos uma cadeia de oligossacarídeo.

[0097] Exemplos de proteínas não glicosilado são, por

exemplo, hormônios não glicosilados, enzimas não glicosilados, fatores de crescimento não glicosilados da família do fator de crescimento do nervo (NGF), do fator de crescimento epitelial (EGF) e da família do fator de crescimento do fibroblasto (FGF) e receptores não glicosilados para hormônios e fatores de crescimento.

[0098] Exemplos para proteínas glicosiladas são hormônios e fatores de liberação e hormônios, fatores de coagulação, fatores anti-coagulantes, receptores para hormônios ou fatores de crescimento, fatores neurotróficos citocinas e seus receptores, receptores de célula T, proteínas de membrana superficial, proteínas de transporte, receptores residentes, endereçamento, proteínas regulatórias, anticorpos, proteínas quiméricas, tais como imuno-adesinas, e fragmentos de qualquer proteínas glicosiladas. Preferivelmente, o polipeptídeo é selecionado a partir do grupo consistindo de anticorpos, fragmentos de anticorpos ou derivados de anticorpos (por exemplo, proteínas de fusão Fc e formatos de anticorpo particular do tipo anticorpos biespecíficos). O fragmento de anticorpo como utilizado aqui inclui, mas não estão limitados a, (i) um domínio, (ii) o fragmento Fab consistindo de VL, VH, CL ou CK e domínios CH1, incluindo Fab' e Fab'-SH, (iii) o fragmento Fd consistindo dos domínios VH e CH1, (vi) o fragmento dAB (Ward ES et al., 1989) "Nature", 341(6242):544-6) que consiste de um domínio variável único (v) fragmentos F(ab')2, um fragmento bivalente compreendendo dois fragmentos Fab ligados (vi) moléculas Fv de cadeia única (scFv), onde um domínio VH e um domínio VL estão ligados por um peptídeo ligante que permite os dois domínios associar para formar um sítio de ligação ao antígeno

(Bird RE et al., (1988) "Science", 242(4877):423-6; Huston JS et al., (1988) "Proc. Natl. Acad. Sci.", USA, 90(14):6444-8; Holliger P et al., (2000) "Methods Enzymol", 326: 461-79), (viii) scFv, diacorpo ou anticorpo do domínio fundido a uma região Fc e (ix) scFv fundido ao mesmo ou a um anticorpo diferente.

[0099] Em algumas concretizações, o cassete de expressão comprehende ainda um elemento genético selecionado a partir do grupo consistindo de um promotor adicional, um melhorador, elementos de controle de transcrição, e um marcador selecionável, preferivelmente um marcador selecionável que é expresso em células animais. Os elementos de controle de transcrição são, por exemplo, sequências Kozak ou elementos terminais de transcrição.

[0100] Em uma concretização, o elemento genético é um marcador selecionável onde o conteúdo dos sítios CpG contido na sequência de polinucleotídio codificante do marcador selecionável é 45 ou menos, preferivelmente, 40 ou menos, mais preferivelmente, 20 ou menos, em particular 10 ou menos, mais particularmente 5 ou menos, mais particularmente 0 (sítios CpG foi completamente removido).

[0101] Em um aspecto adicional, a presente descrição provê um vetor de expressão, preferivelmente, um vetor de expressão de mamíferos compreendendo um cassete de expressão como descrito *supra*. Em algumas concretizações, o vetor de expressão comprehende pelo menos duas unidades de transcrição separadas. Um vetor de expressão com duas unidades de transcrição separadas é também referido como um vetor de duplo gene. Um exemplo do mesmo é um vetor, no qual a primeira unidade de transcrição codifica da cadeia pesada de

um anticorpo ou um fragmento do mesmo e a segunda unidade de transcrição codifica a cadeia leve de um anticorpo. Um outro exemplo, é um vetor de duplo gene, no qual as duas unidades de transcrição codifica duas subunidades diferentes de uma proteína, tal como uma enzima. Entretanto, também é possível que o vetor de expressão da presente invenção compreende mais do que duas unidades de transcrição, por exemplo, três, quatro ou mesmo mais unidades de transcrição separadas cada uma das quais compreende uma sequência de nucleotídeo diferente codificante de uma cadeia de polipeptídeo diferente. Um exemplo, portanto, é um vetor com quatro unidades de transcrição separadas, cada uma das quais contém uma sequência de nucleotídeo diferente codificante de uma subunidade de uma enzima consistindo de quatro subunidades diferentes.

[0102] Em algumas concretizações, o vetor de expressão compreende adicionalmente um elemento genético selecionado a partir do grupo consistindo de um promotor adicional, um melhorador, elementos de controle de transcrição, uma origem de replicação e um marcador selecionável.

[0103] Em algumas concretizações, o vetor de expressão compreende adicionalmente uma origem de replicação e um marcador selecionável, onde o conteúdo dos sítios CpG conteúdos na sequência de polinucleotídio do vetor de expressão codificante da origem de replicação e os marcadores selecionáveis é 200 ou menos, preferivelmente, 150 ou menos, em particular 100 ou menos, mais particularmente 50 ou menos, mais particularmente 30 ou menos.

[0104] Qualquer seleção de marcador comumente empregada tal como quinase timidina (tk), reductase dihidrofolato (DHFR),

puromicina, neomicina ou sintetase glutamina (GS) pode ser utilizado para o cassette de expressão ou vetor de expressão da presente invenção. Preferivelmente, os vetores de expressão da invenção também compreende um número limitado dos sítios de restrição úteis para inserção do cassette de expressão para a secreção de uma proteína heteróloga da presente invenção. Quando utilizado em particular para expressão transitória/epissomal apenas, os vetores de expressão da invenção podem compreender adicionalmente uma origem de replicação tal como de origem oriP do vírus Epstein Barr (EBV) ou vírus SV40 para replicação autônoma/manutenção pessoal em células hospedeiras eucarióticas, mas pode ser desprovida de um marcador selecionável. A expressão transitória em célula perde os fatores relevantes para facilitar a replicação do vetor também é possível. O vetor de expressão acolhendo o cassette de expressão pode ainda compreender um cassette de expressão codificante para um marcador fluorescente, um cassette de expressão codificante para um ncRNA, um cassette de expressão codificante para uma proteína anti-apoptótica, ou um cassette de expressão codificante para uma proteína aumentando a capacidade da rota secretora.

[0105] Em um aspecto adicional, a presente descrição provê um vetor de expressão, que compreende na ordem:

- (a) uma sequência de DNA genômico não traduzida à montante e/ou a jusante de um promotor GAPDH eucariótico;
- (b) um promotor;
- (c) uma sequência de polinucleotídio codificante de um polipeptídeo;
- (d) um sítio de poliadenilação;

- (e) um melhorador;
- (f) uma sequência de DNA genômico não traduzido à jusante e/ou a montante de um promotor GAPDH eucariótico; ou
  - (a) uma sequência DNA genômica não traduzida à montante e/ou à jusante de um promotor GAPDH eucariótico;
  - (b) um melhorador;
  - (c) um promotor;
  - (d) uma sequência de polinucleotídio codificante de um polipeptídeo;
  - (e) um sítio de poliadenilação;
  - (f) uma sequência de DNA genômico não traduzido à jusante e/ou a montante de um promotor GAPDH eucariótico; ou
    - (a) um melhorador;
    - (b) uma sequência de DNA genômico não traduzido à montante e/ou à jusante de um promotor GAPDH eucariótico; ou
      - (c) um promotor;
      - (d) uma sequência de polinucleotídio codificante de um polipeptídeo;
      - (e) um sítio de poliadenilação;
      - (f) uma sequência de DNA genômico não traduzido à jusante e/ou a montante de um promotor GAPDH eucariótico; ou

jusante do promotor GAPDH eucariótico é de 100 a cerca de 15000 nucleotídeos e sendo que a sequência de DNA genômico não traduzido do promotor GAPDH eucariótico inicia dentro de uma região abrangendo a partir da extremidade 5' do promotor GAPDH eucariótico para a posição de nucleotídeo em torno de -3500, onde a posição do nucleotídeo é relativa ao início da transcrição do mRNA GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida à montante do promotor GAPDH eucariótico é de cerca de 100 a cerca de 15000 nucleotídeos, com a provisão de que se (a) ou (b) for uma sequência de DNA genômico não traduzida à montante de um GAPDH eucariótico, (f0 é uma sequência de DNA genômico não traduzido à jusante de um GAPDH eucariótico, e se (a) ou (b) for uma sequência de DNA genômico não traduzido à jusante de uma GAPDH eucariótico (f) é uma sequência de DNA genômico não traduzido à montante de um GAPDH eucariótico.

Em algumas concretizações, a presente descrição provê um vetor de expressão, que compreende na ordem: (a) uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico; (b) um promotor; (c) uma sequência de polinucleotídios codificante de um polipeptídio; (d) um sítio de poliadenilação; (e) um melhorador; (f) uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico, onde a inclusão do melhorador é opcional.

Em um aspecto adicional, a presente descrição provê um vetor de expressão, que compreende na ordem:

- (a) uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico;
- (b) um melhorador;

- (c) um promotor;
- (d) uma sequência de polinucleotídios codificantes de um polipeptídio;
- (e) um sítio de poliadenilação;
- (f) uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico, onde a inclusão do melhorador é opcional;

Em um aspecto adicional, a presente descrição provê um vetor de expressão, que comprehende na ordem:

- (a) um melhorador;
- (b) uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico;
- (c) um promotor;
- (d) uma sequência de polinucleotídios codificantes de um polipeptídio;
- (e) um sítio de poliadenilação;
- (f) uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótica; onde a inclusão do melhorador é opcional.

Em um aspecto adicional, a presente descrição provê um vetor de expressão, que comprehende na ordem:

- (a) uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico;
- (b) um promotor;
- (c) uma sequência de polinucleotídios codificante de um polipeptídio;
- (d) um sítio de poliadenilação;
- (e) um melhorador;
- (f) uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico, onde a inclusão do melhorador

é opcional;

[0107] Em um aspecto adicional, a presente descrição provê um vetor de expressão, que compreende na ordem:

- (a) uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico;
- (b) um melhorador;
- (c) um promotor;
- (d) uma sequência de polinucleotídios codificante de um polipeptídio;
- (e) um sítio de poliadenilação;
- (f) uma sequência de DNA genômico não traduzida à montante de um promotor GAPDH eucariótico, onde a inclusão do melhorador é opcional.

[0108] Em um aspecto adicional, a presente descrição provê um vetor de expressão, que compreende na ordem:

- (a) um melhorador;
- (b) uma sequência de DNA genômico não traduzido à jusante de um GAPDH eucariótico;
- (c) um promotor;
- (d) uma sequência de polinucleotídio codificante de um polipeptídeo;
- (e) um sítio de poliadenilação;
- (f) uma sequência de DNA genômico não traduzida à montante de um GAPDH eucariótico, onde a inclusão do melhorador é opcional.

[0109] A sequência de DNA genômico não traduzida à montante de um GAPDH eucariótico, melhorador, promotor, sequência de polinucleotídio codificante de um polipeptídeo, sítio de poliadenilação e sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico dos vetores de

expressão são, por exemplo, como descrito *supra*.

[0110] Em um aspecto adicional, a presente descrição provê uma célula hospedeira compreendendo um cassete de expressão ou um vetor de expressão como descrito *supra*. A célula hospedeira pode ser uma célula humana ou uma célula não humana. As células hospedeiras preferidas são células de mamíferos. Os exemplos preferidos de células hospedeiras de mamíferos incluem, sem estar restrita a, células de rim embriônico humano (Graham FL et al., "J. Gen. Virol.", 36: 59-74), fibroblastos humanos MRC5, células de melanoma humano 983M, células de rim canino MDCK, fibroblastos de pulmão de rato cultivados RF isolados a partir de ratos Sprague-Dawley, as células melanoma de murino B16BL6, células mastocitoma P815, células de adenocarcinoma mamário de murino MTI A2, células PER:C6 (Leiden, Holanda), e células ou linhas de célula de ovário de hamsters chineses (CHO (Puck et al., 1958; "J. Exp. Med.", 108: 945-955).

[0111] Em uma concretização preferida particular, a célula hospedeira é uma célula ou linha de célula de ovário de hamsters chinês (CHO). As linhas de células CHO apropriadas incluem, por exemplo, CHO-S (Invitrogen, Carlsbad, CA, USA), CHO K1 (ATCC CCL-610, CHO pro3-, CHO DG44, CHO P12 ou a linha de célula CHO dhfr- DUK-BII (Chasin et al., PNAS 77, 1980, 4216-4220), DUXBI 1 (Simonsen et al., PNAS 80, 1983, 2495-2499), ou CHO-K1SV (Lonza, Basel, Suíça).

[0112] Em um aspecto adicional, a presente invenção provê um método *in vitro* para a expressão de um polipeptídeo, compreendendo a transfecção de uma célula hospedeira com o cassete de expressão ou um vetor de expressão como descrito *supra* e recuperação do polipeptídeo. O polipeptídeo é

preferivelmente, um heterólogo, mais preferivelmente, um polipeptídeo humano.

[0113] Para transfecção, o cassete de expressão ou o vetor de expressão em uma célula hospedeira de acordo com a presente invenção qualquer técnica de transfecção tal como aquelas bem conhecidas do estado da técnica, por exemplo, eletroporação, co-precipitação de fosfato de cálcio, transfecção DEAE-dextran, lipofecção, pode ser empregado se apropriado para um determinado tipo de célula hospedeira. Deve ser observado que a célula hospedeira transfectada com o cassete de expressão ou o vetor de expressão da presente invenção deve ser construída como sendo uma linha de célula estavelmente transfectada ou transitória. Assim, de acordo com a presente invenção, o cassete de expressão ou vetor de expressão pode ser mantido epissomalmente, ou seja, transfectado transitoriamente ou pode ser estavelmente integrado no genoma da célula hospedeira, ou seja, estavelmente transfectado.

[0114] Uma transfecção transitória é caracterizada por não aparelho de qualquer seleção de pressão para um vetor do marcador de seleção de nascimento. Em experimentos de expressão transitória, que comumente levam de 2 até 10 dias após a transfecção, cassete de expressão ou vetor de expressão transfectados são mantidos como elementos epissomais e não são ainda integrados dentro do genoma. Ou seja, o DNA transfectado não integra, usualmente, dentro do genoma da célula hospedeira. As células hospedeiras tendem a perder o DNA transfectado e crescer além das células transfectadas na população durante a cultura da célula transitoriamente transfectada fundida. Portanto, a expressão

é mais forte no período imediatamente após a transfecção e diminui com o tempo. Preferivelmente, um transfectante transitório de acordo com a presente invenção é entendido como uma célula que é mantida na cultura celular na ausência de seleção de pressão até um tempo de 2 a 10 dias após a transfecção.

[0115] Em uma concretização preferida da invenção, a célula hospedeira, por exemplo, a célula hospedeira CHO é estavelmente transfectadas com o cassete de expressão ou com o vetor de expressão da presente invenção. Os meios de transfecção estáveis que introduziram recentemente o DNA exógeno tal como o vetor de DNA tornou-se incorporado dentro do DNA genômico, usualmente por eventos de recombinação aleatórios, não homólogos. O número de cópias do vetor de DNA e concomitantemente a quantidade do produto de gene pode ser aumentada por seleção das linhas de célula nas quais as sequências de vetores foram amplificadas após integração dentro do DNA da célula hospedeira. Portanto, é possível que as referidas integrações estáveis resultem em elevação, durante a exposição ao aumento adicional na pressão de seleção para a amplificação do gene, para o dobro de minutos nos cromossomos em células CHO. Além disso, uma transfecção estável pode resultar em perda de partes da sequência do vetor não diretamente relacionadas à expressão do produto de gene recombinante, tal como, por exemplo, as regiões de controle de número de cópias bacterianas resultantes do superfluxo durante a integração genômica. Além disso, uma célula hospedeira transfectada foi integrada pelo menos em parte ou partes diferentes do cassete de expressão ou vetor de expressão dentro do genoma.

[0116] Em um aspecto adicional, a presente invenção provê o uso do cassete de expressão ou um vetor de expressão como descrito *supra* para a expressão de um polipeptídeo heterólogo a partir da célula hospedeira de mamífero, em particular o uso do cassete de expressão ou um vetor de expressão como descrito *supra* para a expressão *in vitro* de um polipeptídeo heterólogo a partir de uma célula hospedeira de mamífero.

[0117] A expressão e a recuperação da proteína podem ser realizadas de acordo com os métodos conhecidos para um técnico no assunto.

[0118] Para a expressão de um polipeptídeo, a sequência de DNA genômico não traduzida à jusante e/ou a montante de um promotor GAPDH eucariótico do cassete de expressão ou do vetor de expressão como descrito *supra* e a célula hospedeira como descrito *supra* são utilizados e são usualmente da mesma origem. Surpreendentemente, foi observado que um aumento da expressão é obtido se a sequência do DNA genômico não traduzida á jusante e/ou a montante de um promotor GAPDH eucariótico do cassete de expressão ou do vetor de expressão e a célula hospedeira são de origem diferente, por exemplo, se a sequência de DNA humana á jusante e/ou a montante de um promotor GAPDH eucariótico são utilizados em células CHO.

[0119] Em um aspecto adicional, a presente descrição provê o uso do cassete de expressão ou do vetor de expressão como descrito *supra* para a preparação de um medicamento para o tratamento de um distúrbio.

[0120] Em um aspecto adicional, a presente descrição provê o cassete de expressão ou o vetor de expressão *supra* para o uso como um medicamento para o tratamento de um distúrbio.

[0121] Em um aspecto adicional, a presente descrição provê

o cassette de expressão ou o vetor de expressão como descrito supra para uso na terapia gênica.

#### EXEMPLOS

Exemplo 1 - clonagem de vetores de expressão:

##### I. Materiais e Métodos

###### I.1 Constructos de plasmídeos

###### I.1.1. Placas de cultura LB

[0122] 500 ml de água foram misturadas e fervidas com 6g de ágar LB (Invitrogen, Carlsbad, CA, USA) (1 litro de LB contém 10g de triptofano, 5g de extrato de levedura e 10g de NaCl). Após resfriar, o respectivo antibiótico foi adicionado à solução que foi então plaqueada (placas de ampicilina a 100 µg/ml e placas de canamicina a 50µg/ml).

###### I.1.2 - Reação em cadeia de polimerase (PCR) :

[0123] Todos os PCR foram realizados usando 1 µl de dNTPs (10mM para cada dNTP; Invitrogen, Carlsbad, CA, USA), 2 unidades de Phusion® DNA Polimerase (Finnzymes Oy, Espoo, Finlândia), 25 nmol de Primer A (Mycrosynth, Balgach, Suíça), 25 nmol de Primer B (Mycrosynth, Balgach, Suíça), 10µl de tampão 5X HF (7,5 mM de MgCl<sub>2</sub>, Finnzymes, Espoo, Finlândia), 1,5 µl de sulfóxido de dimetila (DMSO, Finnzymes, Espoo, Finlândia) e 1-3 µl do modelo (1-2 µg) em um volume final de 50 µl. Todos os primers utilizados estão listados na Tabela 1.

[0124] PCR foi iniciado por uma desnaturação inicial a 98°C durante 3 minutos, seguido por 35 ciclos de 30 segundos de desnaturação a 98°C, 30 segundos de anelamento em uma temperatura específica ao primer (de acordo com o conteúdo CG) e 2 minutos de elongação a 72°C. Uma elongação final a 72°C durante 10 minutos foi realizada antes do resfriamento e

mantida a 4°C.

Tabela 1

Resumo dos primers utilizados no PCR. GAPDH: sequência da desidrogenase gliceraldeído-3-fosfato, 5': sequência à montante; 3: sequência à jusante. O "T" (sublinhado) no primer GlnPr1172 foi introduzido de modo a evitar a formação de dímeros no primer.

Primer	Sequência do primer	Sequência amplificada	Seq ID
GlnPr 1171	ATTATTCGCGATGGCTCCTGGCA TCTCTGGGACCGAGGC	5' -GAPDH	SEQ ID NO:1
GlnPr 1172	ATCGTCGCGAAGCTTGAGATTGT CCAAGCAGGTAGCCAG		SEQ ID NO:2
GlnPr 1173	AGCAAGTACTTCTGAGCCTCA GTAATGGCTGCCTG	3' -GAPDH	SEQ ID NO:3
GlnPr 1174	TGGCAGTACTAACGCTGGCACCA CTACTTCAGAGAACAAAG		SEQ ID NO:4

### I.1.3 - Digestão por restrição

[0125] Para todas as digestões por restrição em torno de 1µg de DNA plasmídial (quantificado com NanoDrop, espectrofotômetro ND-1000) (Thermo Scientific, Wilmington, DE, USA)) foi misturado em 10-20 unidades de enzima cada, 4 ml de tampão 10X NE correspondente (NEB, Ipswich, MA, USA), e o volume foi completado a 40 µl com H<sub>2</sub>O estéril. Sem indicação adicional, as digestões foram incubadas por 1 hora a 37°C.

[0126] Após cada digestão preparativa da cadeia principal, 1 unidade de fosfatase alcalina intestinal de bezerro (CIP; NEB, Ipswich, MA, USA) foi adicionada e a mistura foi incubada 30 minutos a 37°C. Se a digestão foi feita em tampão NE3 (NEB, Ipswich, MA, USA), o tampão foi alterado para tampão NEB4 antes da adição de CIP porque esta enzima tem uma forte atividade neste tampão e pode também digerir alguns dos nucleotídeos na extremidade externa.

#### I.1.4 - Purificação por PCR e eletroforese em gel de agarose

##### I.1.4.1 - Limpeza por PCR

Para permitir a digestão, todos os fragmentos do CPR foram limpos antes da digestão por restrição usando um kit Macherey Nagel Extract II (Macherey Nagel, Oensingen, Suíça) após o manual do fabricante usando 40  $\mu$ l de tampão de eluição. Este protocolo foi também utilizado para mudar os tampões das amostras de DNA.

##### I.1.4.2 - Extração do DNA

[0127] Para a eletroforese em gel, 1% dos géis foram preparados usando UltraPure™ Agarose (Invitrogen, Carlsbad, CA, USA) e tampão de EDTA de ácido acético 50X Tris (TAE, pH 8,3; Bio Rad, Munique, Alemanha). Para a coloração do DNA, 1  $\mu$ g de Gel Red Gye (Biotum, Hayward, CA, USA) foi adicionado a 100 ml de gel de agarose. Como um marcador de tamanho, 2  $\mu$ g de um DNA Ladder de 1kd (NEB, Ipswich, MA, USA) foi utilizado. A eletroforese foi corrida em torno de 1 hora a 125 Volts. As bandas de interesse foram cortadas do gel de agarose usando o kit Extract II (Macherey-Nagel, Oensingen, Suíça), após o manual do fabricante usando 40  $\mu$ l de tampão de eluição.

#### I.1.5 - Ligação

[0128] Para cada ligação, 4  $\mu$ l de inserto foram misturados a 1  $\mu$ l do vetor, 400 unidades de ligase (T4 DNA ligase, NEG, Ipswich, MA< USA), 1  $\mu$ l de tampão 10X ligase (tampão T4 DNA Ligase, NEB, Ipswich, MA, USA), em um volume de 10  $\mu$ l. A mistura foi incubada durante 1-2 horas a RT.

#### I.1.6 - Transformação dos produtos de ligação dentro da bactéria competente:

[0129] Para a clonagem de pGLEX41-[REP] e para os

constructos feitos com o vetor pCR-Blunt que contém uma origem padrão de replicação, TOP 10 (One Shot® TOP 10 *E. coli* competente; Invitrogen, Carlsbad, CA, USA) foram utilizadas.

[0130] Para o início da replicação do plasmídeo contendo a origem de replicação R6K, a expressão da proteína  $\pi$ , codificada por uma sequência, *pir*, é requerida. A proteína  $\pi$  é expressa por uma *E. coli* competente Shot®PIR1 (Invitrogen, Carlsbad, Ca, USA). Estas bactérias foram utilizadas para todos os vetores contendo a sequência R6K.

[0131] Para transformar a bactéria competente com o produto de ligação, 25-50  $\mu$ l da bactéria foram descongelados em gelo por 5 minutos. Então, 3-5  $\mu$ l do produto de ligação foram adicionados às bactérias competentes e incubadas durante 20-30 minutos em belo antes do choque térmico druantel minuto a 42°C. Então, 500  $\mu$ l do meio S.O.C (Invitrogen, Carlsbad, CA, USA) foram adicionados por tubo e incubados durante 1 hora a 37°C sob agitação. Finalmente, a bactéria foi colocada em uma placa de LB com ampicilina (Sigma-Aldrich, St. Louis, MO, USA) e incubada durante a noite a 37°C. Para a clonagem em vetores pCR-Blunt, placas com canamicina (Sigma-Aldrich, St. Louis, MO, USA) foram utilizadas.

I.1.7 - Preparação do plasmídeo em pequena escala (mini) e média escala (midi):

I.1.7.1 - Minipreparação

[0132] Para a minipreparação, as colônias de bactérias transformadas foram crescidas durante 6-16 horas em 2,5 ml de LB e ampicilina ou canamicina a 37°C, 200 rpm. O DNA foi extraído com um kit de purificação do plasmídeo para *E. coli* (QuickPure, Macherey Nagel, Oensingen, Suíça), após o manual provido.

[0133] O DNA plasmídial a partir das minipreparações foi quantificada uma vez com o espectrofotômetro NanoDrop ND-1000 (Thermo Scientific, Wilmington, DE, USA), através a medida da absorbância a 260 nm e avaliação a proporção do OD 260 nm/OD280 nm que tinha que estar entre 1,8 e 2. Uma digestão controle foi realizada antes de enviar a amostra para Fasteris S.A (Genebra, Suíça) para confirmação da sequência.

[0134] Para a extração BAC, o kit QuickPure (Macherey Nagel, Oensingen, Suíça) foi utilizada com a modificação a seguir do protocolo: 10 ml de LB e cloranfenicol (12,5 µg/ml) (Sigma-Aldrich, St. Louis, MO, USA) foram semeadas com bactéria contendo o vetor pBACe3.6. Após incubação em uma plataforma de agitação a 37°C durante a noite, a cultura foi centrifugada durante 5 minutos a 13.300 rpm antes de ser ressuspensa em 500 µl de um tampão A1. 500 µl de tampão de lise A2 foram adicionados e a solução foi incubada durante 5 minutos a RT. Então, foi neutralizado com 600 ml de tampão A3 e centrifugado por 10 minutos a 13300 rpm. O sobrenadante foi carregado em uma coluna e a partir desta etapa em diante, o protocolo padrão do kit QuickPure miniprep foi utilizado.

#### I.1.7.2 Midipreparação

[0135] Para a midipreparação, a bactéria transformada foram crescidas a 37°C durante a noite em 200 a 400 ml de LB e ampicilina (ou canamicina). Então, a cultura foi centrifugada 20 minutos a 725 g e o plasmídeo foi purificado usando um kit comercial (NucleoBond Xtra Midi; Macherey nagel, Oensingen, Suíça) após o protocolo baixo do plasmídeo no manual do fabricante.

[0136] O DNA plasmídial a partir da midipreparação foi quantificado três vezes com o espectrofotômetro NanoDrop ND-

1000, confirmado pela digestão por restrição e finalmente envidado para o sequenciamento (Fasteris SA, Genebra, Suíça).

## II. Resultados e Discussão

II.1 -Clonagem das regiões de DNA à montante e a jusante do cassette de expressão GAPDH (5' e 3' de GAPDH).

[0137] O clone BAC RPCIB753F11841Q foi ordenado em Imagene (Berlin, Alemanha). Este clone contém a sequência GAPDH humana em uma cadeia principal do vetor pBACe3.6, contendo um gene de resistência ao cloranfenicol. Após a extração de DNA por minipreparação, a concentração do vetor foi determinada pelo Nanodrop para 27 ng/ $\mu$ l.

[0138] As sequências de DNA imediatamente em torno do cassette de expressão GAPDH à montante do promotor e à jusante do sítio de poliadenilação amplificada pelo PCR usando 27 ng do clone purificado RPCIB753F11841Q como padrão. O fragmento 3kb à montante do promotor foi amplificada com primers GlnPr1171 (SEQ ID NO:1) E gLNpR1172 (SEQ ID NO:2) conduzindo ao amplicon com SEQ ID NO:5. Como o primer GlnPr1172 (SEQ ID NO:2) carrega uma mudança de base (G para T) em relação a sequência padrão, todas as sequências derivadas a partir desta reação de PCR irá realizar a mudança de base, também. A mudança está localizada na posição -3721 em relação ao início da transcrição do gene GAPDH (bp 812 da SEQ ID NO:17, posição 23 em relação ao início da SEQ ID NO:5). O fragmento de 3kb à jusante do sítio de poliadenilação foi amplificada com primers GlnPr1173 (SEQ ID NO:3) e GlnPr1174 (SEQ ID NO:4) conduzindo ao amplicon com SEQ ID NO:6 (Tabela 1). A temperatura de anelamento utilizada para estes PCRs foi de 72°C.

[0139] Os fragmentos de GAPDH na 5' e na 3' (SEQ ID NOS: 5

e 6) foram clonados no pCR-Blunt, um vetor de clonagem produto de PCR disponível comercialmente (pCR-Blunt, kit de clonagem PCR Zero Blunt, Invitrogen). Os produtos de ligação foram transformados na bactéria competente TOP10 e plaqueados em placas de LB-agar canamicina. As colônias foram amplificadas e os plasmídeos foram isolados através de minipreps. As digestões de controle foram conduzidas para identificar clones positivos resultando em constructos pCR-Blunt-5'-GAPDH e pCR-Blunt-3'-GAPDH.

II.2- Preparação do fragmento de DNA codificante para as proteínas repórteres GFP e um anticorpo monoclonal IgG1 recombinante (LC-IRES-HC-IRES-GFP)

[0140] O constructo repórter (REP) utilizado no presente trabalho consistiu em um gene policistronico: anticorpo monoclonal IgG1 de cadeia leve (LC)-IRES-IgG1, anticorpo monoclonal de cadeia pesada (HC)-IRES-proteína verde fluorescente (GFP). A presença de sítios de entrada ribossomal interna (IRES) derivada do vírus encefalomielite (Gurtu et al., "Biochem Biophys Res Commun.; 229(1): 295-298, 1996)) permite a tradução dos peptídeos 3 do anticorpo monoclonal IgG1 de cadeia leve (LC), anticorpo monoclonal IgG1 de cadeia pesada (HC) e GFP (Figura 1). As células transfetadas irão, portanto secretar o anticorpo monoclonal IgG1 e acumula o GFP intracelular de uma maneira dependente. Entretanto, o mRNAs policistrônico não são comuns em células eucariótico e sua tradução não é muito eficiente, conduzindo aos títulos relativos baixos de expressão de IgG1 e GFP.

[0141] Um vetor contendo o constructo REP foi digerido usando as enzimas de restrição NheI e BstBI (BstBI é

utilizado a 65°C). O fragmento REP contendo o constructo de expressão foi cortado, purificados utilizados para etapas de clonagem adicionais.

#### II.3- Clonagem de vetores de expressão:

[0142] O vetor pGLEX41, um vetor de expressão derivado a partir de pcDNA 3.1 (+) (Invitrogen, Carlsbad, CA) foi utilizado para a produção de linha de célula estável. Foi utilizada como uma cadeia principal inicial que foi modificada para gerar a segunda geração de vetores A e B com e sem as sequências GAPDH. Para todos os vetores, a mesma combinação do promotor-ítron (mCMV e um fragmento doador-receptor codificante para o primeiro ítron (IgDA)) foi utilizado(Gorman et al., (1990) "Proc. Natl. Acad. Sci. USA, 87: 5459-5463).

#### Clonagem do vetor intermediário pGLEX41-HM-MCS-ampiA:

[0143] O desenvolvimento da geração do novo vetor foi iniciada a partir de pGLEX41. Este vetor foi cortado usando as enzimas de restrição NruI e BspHI de modo a liberar o cassete de resistência à ampicilina. O fragmento de cadeia principal foi CIPeD e purificado pela eletroforese em gel. O fragmento de DNA codificante para uma versão de códon otimizada (para expressão em *E. coli*) do gene de resistência à ampicilina (incluindo o promotor bla) foi ordenado da GeneArt. O inserto foi cortado do vetor de clonagem GeneArt # 1013237 usando as enzimas de restrição NruI e BspHI (as mesmas enzimas utilizadas para a cadeia principal), purificada e clonada dentro da cadeia principal. Os minipreps foram analisados através da digestão por restrição. O clone pGLEX41-HM-MCS-ampiA#2 teve o perfil de restrição esperado e a integração do fragmento correto foi confirmada pelo

sequenciamento.

**Clonagem do vetor intermediário pGLX41-MCS-56K-ampiA:**

[0144] A ordem para trocar o pUC de origem de replicação do vetor pGLEX41-HM-MCS-ampiA#2 foi digerido usando um PvuI e BspHI. Os fragmentos da cadeia principal foi CIPed e purificado. O novo fragmento inserido contém a origem R6K de replicação e uma sequência SV40 modificada poli(A) como parte do cassete de expressão. As sequências da cadeia principal viral ou bacteriana não necessária em torno de SV40 poli(A) foi eliminada (ver Tabela 2 abaixo). O fragmento inserido foi ordenado da GeneArt; foi cortado do vetor de clonagem da GeneART#1013238 usando as enzimas PvuI e BspHI (a mesma como utilizada para a cadeia principal), purificada e clonado dentro do fragmento de cadeia principal. Os minipreps foram preparados e foram confirmados pela análise de sequência. O clone pGLEX41-MCS-R6K-ampiA#2 teve a sequência correta.

Tabela 2 – Conteúdo de CpG nos vetores diferentes

	Conteúdo CpG nos vetores de expressão		
	pGLEX41	Vetores de códon otimizado: "A"	Vetores de CpG reduzidos "B"
Resistência à ampicilina	49	43	19
Resistência à puromicina	93	36	1
Resistência à geneticina	74	51	0
Origem da replicação	45	9	9
Soma:	261	139	29

**Clonagem de vetor intermediário pGLEX41-MCS-R6K-ampiB**

[0145] O vetor pGLEX41-MCS-R6K-ampiA#1 foi aberto usando a enzima de restrição BspHI e CIPed de modo a liberar a resistência à ampicilina. O novo fragmento inserido contém o códon de resistência a ampicilina otimizado para expressão em

*E. coli*, mas em todas as sequências CpG que poderiam ser eliminadas pelo códon alternativa usados foram substituídos (ver tabela 2 acima). Este fragmento foi ordenado na GeneArt. De modo a liberar o fragmento inserido, o vetor GeneArt de clonagem #1016138 foi digerido usando BspHI. Após purificação de ambos os fragmentos de cadeia principal e inserido pela eletroforese em gel, eles foram ligados e transformados em uma bactéria PIR1. Os minipreps foram diretamente enviados para sequenciamento, pPLEX41-R6K-MCS-ampiB#1 foi a sequência correta e foi utilizada para etapas de clonagem adicionais. Clonagem do constructo repórter em vetores de expressão derivados pPLEX41.

[0146] De modo a clonar o constructo repórter REP nos vetores de expressão pPLEX41-MCS-R6K-ampiA e pPLEX41-MCS-R6K-ampiB, os vetores foram cortados usando as enzimas de restrição NheI e ClaI. O vetor de expressão pPLEX41-HM-MCS foi aberto usando as enzimas de restrição NheI e BstBI (a 65°C). Todas as cadeias principais do vetor foram tratadas com CIP após a digestão e as cadeias principais purificadas pela eletroforese em gel. As cadeias principais foram ligadas com o fragmento NheI/BstBI (BstBI é compatível com ClaI) codificador para o constructo repórter REP. Os produtos de ligação foram transformados em bactéria competente PIR1 ou TOP10 e plaqueados em placas de LB-agar ampicilina. As colônias foram amplificadas e os plasmídeos foram isolados através de minipreps. Os clones positivos poderiam ser identificados por digestão de restrição dos minipreps e subsequentemente confirmação da sequência por Fasteris SA. Adição de sequências GAPDH nos flancos em vetores de expressão derivados de pPLEX41.

[0147] Todas as digestões por restrição deste parágrafo foram realizadas em um volume final de 80  $\mu$ l e incubadas durante a noite a 37°C.

[0148] A sequência 5'GAPDH (SEQ ID NO:7) foi cortada de pCR-blunt-5'GAPDH usando a enzima de restrição NruI e ligado nos vetores de expressão pLEX41-R6K-ampiA-[REP] e pLEX41-R6K-ampiB-[REP] que foram linearizada usando NruI e tratado com CIP de modo a evitar a recircularição. Após a amplificação de colônias PIR1 (obtida pela transformação de produtos de ligação) minipreps foram analisados pela digestão por restrição. Os clones pLEX41-R6K-ampiA-5'GAPDH-[REP]#2 e pLEX41-R6K-ampiA-5'GAPDH-[REP]#1 mostrou bandas do tamanho esperado na análise de restrição, foram subsequentemente confirmado pelo sequenciamento e utilizado para etapas de clonagem adicional. Estes novos vetores foram então abertos com ScaI e tratados com CIP. O fragmento 3'GAPDH (SEQ ID NO:8) foi cortado do pCR-Blunt-3'GAPDH usando a mesma enzima e ligado dentro de duas cadeias principais de modo a gerar os vetores de expressão pLEX41-R6K-ampiA-GAPDH-[REP] e pLEX41-R6K-ampiB-GAPDH-[REP].

[0149] O controle da digestão de clones pLEX41-R6K-ampiA-GAPDH-[REP]#2 e pLEX41-R6K-ampiB-GAPDH-[REP]#8 mostraram bandas de tamanho esperado na análise de restrição. A inserção do fragmento 3'-GAPDH na orientação correta foi subsequentemente confirmada pelo sequenciamento (Fasteris).

#### II.4- Clonagem de vetores de resistência:

[0150] O ponto de partida para a clonagem dos vetores de resistência foi o vetor pLEX-MCS-R6K-ampiA#1. Como para a expressão dos genes de resistência um promotor fraco é suficiente, o promotor mCMV foi substituído pelo promotor

SV40. Os genes codificantes para os genes de resistência foram ordenados da GeneART SA (Regensburg, Alemanha) e também otimizado para expressão em hamsters chineses (puromicina: puroA e neomicina: neoA) ou reduzida no conteúdo CpG através do uso do códon seletivo (puromicina: puroB e neomicina: neoB).

#### Clonagem de pGLEX-R6K-ampiA-PuroA/PuroB:

[0151] De modo a clonar a resistência à puromicina no cassete de expressão, o vetor pGLEX41-MCS-R6K-ampiA#1 foi aberto usando as enzimas de restrições NruI e XbaI seguido pelo tratamento com CIP. O fragmento inserido foi ordenado da GeneArt e foi provido como inserto no vetor de clonagem GeneArt #1013239. Ele contém o promotor SV40 e o códon do gene otimizado para a resistência de puromicina (para o uso do códon de células CHO). O inserto foi cortado do vetor de clonagem GeneART usando as enzimas NruI e XbaI (a mesma usada para a cadeia principal), purificada e clonada dentro do fragmento da cadeia principal. Os minipreps foram preparados e analisados pela digestão por restrição. O clone pGLEX41-MCS-R6K-ampiA-puroA-#1 mostrou o perfil correto e poderia ser confirmada pelo sequenciamento.

[0152] Este vetor foi utilizado para a clonagem do vetor pGLEX41-MCS-R6K-ampiA-puroB através da troca da região codificante para o gene de resistência à puromicina, enquanto conduz o promotor SV40. O novo fragmento inserido contém uma versão otimizada do códon do gene puromicina, onde todas as sequências Cpg que poderiam ser eliminadas devido a alternativa de uso do códon ter sido substituído. O fragmento adquirido da GeneART e foi liberado no vetor de clonagem #1016139. De modo a liberar o fragmento do inserto, o vetor

GeneArt foi digerido usando as enzimas de restrição XbaI e NotI. O fragmento inserido foi purificado por eletroforese em gel e clonado dentro da cadeia principal do pGLEX-MCS-R6K-ampiA-puroA, após liberação da estrutura de leitura aberta da puromicina através da digestão por restrição usando XbaI e NotI, seguido pelo tratamento CIP. O vetor resultante pGLEX-MCS-R6K-ampiA-puroB#1 foi confirmada diretamente pela análise de sequência.

Clonagem dos vetores pGLEX-R6K-ampiA-NeoA e pGLEX-R6K-ampiA-NeoB:

[0153] De modo a clonar a resistência à neomicina no cassete de expressão, o vetor pGLEX-R6K-puroA#1 foi aberto usando as enzimas de restrições XbaI e NotI, seguido pelo tratamento com CIP. O fragmento inserido foi ordenado da GeneArt e foi provido como inserto no vetor de clonagem GeneArt #1013242 (neoA) e #1026894 (neoB). Eles contêm o gene do códon otimizado para a resistência à neomicina para uso do códon de células CHO e a versão reduzida de CpG para a resistência à neomicina, respectivamente. O inserto foi cortado do vetor de clonagem GeneART usando as enzimas XbaI e NotI (a mesma usada para a cadeia principal), purificada e clonada dentro do fragmento da cadeia principal. Os minipreps foram preparados e os clones foram confirmados por sequenciamento.

Clonagem de vetores pGLEX-R6K-ampiB-NeoB e pGLEX41-R6K-ampiB-puroB:

[0154] O vetor pGLEX41-R6K-puroB#1 foi aberto usando as enzimas de restrições BspHI e subsequentemente CIPed. O fragmento inserido contém o gene de resistência à ampicilina que usa o códon otimizado para expressão em *E.coli*, enquanto

todas as sequências CpG que poderiam ser eliminadas devido ao uso de códon alternativo ter sido substituído. Este fragmento foi adquirido na GeneArt e arranjado no vetor de clonagem #1016138. De modo a liberar o fragmento inserido, o vetor de clonagem GeneART foi digerido usando BspHI. Após a purificação de ambos, o fragmento inserido e o fragmento da cadeia principal por eletroforese em gel, eles foram ligados e transformados na bactéria PIR1. Os minipreps foram diretamente enviados, por sequenciamento e poderiam ser confirmados (pGLEX41-ampiB-R6K-puroB#1).

[0155] A clonagem conduziu ao vetor pGLEX-R6K-neoB-ampiB foi feita por abertura do pGLEX-R6K-neoB-ampiA usando as enzimas de restrição BspHI de modo a criar o fragmento da cadeia principal. A digestão de pGLEX-R6K-ampiB-higroB usando a mesma combinação de enzima de restrição resultou no fragmento inserido codificante para ampiB. O inserto ampiB foi克lonado dentro da cadeia principal de pGLEX-R6K-neoB-ampiA.

II.5- Adição das sequências à montante e à jusante do gene GAPDH humano dentro dos vetores de resistência:

[0156] O vetor pCR-blunt-5'GAPDH foi digerido com NruI de modo a obter o inserto 5'-GAPDH (3164 bps). Os vetores codificantes para os genes de resistência digeridos com NruI, subsequentemente tratados com CIP (fosfatase intestinal de bezerro, NEB, Ipswich, MA) de modo a preparar os fragmentos da cadeia principal. Os 4 diferentes fragmentos da cadeia principal (pGLEX-R6K-neoA-ampiA, pGLEX-R6K-neoB-ampiB, pGLEX-R6K-puroA-ampiA e pGLEX-puroB-ampiB) foram ligados com o inserto 5'GAPDH de 3164 bps e transformado em uma bactéria PIR1 competente. A digestão por restrição de minipreps usando

Apali permitiu a identificação de clones pGLEX-R6K-neoB-ampiB-5'GAPDH#5, pGLEX-R6K-neoA-ampiA-5'GAPDH#6, pGLEX-R6K-puroA-ampiA-5'GAPDH#16, e pGLEX-puroB-ampiB-5'GAPDH#5.

[0157] Estes vetores intermediários foram então cortados com a enzima de restrição ScaI e tratados com CIP de modo a preparar a cadeia principal para ligação. O vetor carregando o segundo fragmento inserido, pCR-Blunt-3'GAPDH, foi cortado usando ScaI de modo a liberar o fragmento inserido (3224 bps) a região flanqueando o GAPDH à jusante. As quatro moléculas de cadeias principais diferentes foram ligadas com o fragmento inserido purificado de 3224 bps e transformadas em células pIR1 competentes. Os minipreps foram analisados por digestão com restrição. Os clones mostrando fragmentos de restrição do tamanho esperado foram pGLEX-R6K-neoB-ampiB-GAPDH#8, pGLEX-R6K-neoA-ampiA-GAPDH#1, pGLEX-R6K-puroA-ampiA-GAPDH#1 e pGLEX-puroB-ampiB-GAPDH#4. Os clones foram subsequentemente confirmados pela análise de sequenciamento (Fasteris, Genebra, Suíça).

#### II.1.5 - Midipreparações de plasmídeos clonados para transfecção:

[0158] De modo a ter quantidades suficientes de plasmídeos, os midipreps forma preparados usando o kit Macherey Nagel (NucleoBond Xtra Midi; Macherey Nagel, Oensingen, Suíça). Após confirmação através da digestão por restrição e sequenciamento, os plasmídeos foram linearizados e utilizados para transfecção em células CHO-S. A tabela 3 resume as concentrações de plasmídeos de DNA em bateladas obtidos nas midipreparações, os preps de DNA linearizados que foram preparados para transfecção, às enzimas utilizadas para linearização e as sequências depositadas na Fasteris SA

confirmando a identidade e a informação da sequência do respectivo plasmídeo. Todos os midipreps foram confirmados através do sequenciamento antes de serem utilizados para transfecção.

Tabela 3

Resumo dos plasmídeos clonados. Concentração de midipreparação de DNA e midipreparação linearizada (com a enzima correspondente). O número de GSC codifica para o respectivo plasmídeo e permite identificar os arquivos de sequenciamento relevantes.

Plasmídeos	Conc. da midipreparação ( $\mu\text{g}/\text{ml}$ )	Enzima para linearização	Conc. dos plasmídeos linearizados ( $\mu\text{g}/\text{ml}$ )	Código do plasmídeo da Glenmark
pGLEX41-R6K-ampiA-[REP]-GAPDH	1538	EcoRV	1019	GSC2774
pGLEX41-R6K-ampiB-[REP]-GAPDH	1243	EcoRV	1233	GSC 2773
pGLEX-R6K-ampiA-neoA-GAPDH	890	AseI	766	GSC 2776
pGLEX-R6K-ampiB-neoB-GAPDH	594	AseI	979	GSC 2777
pGLEX-R6K-ampiA-puroA-GAPDH	917	AseI	858	GSC2778
pGLEX-R6K-ampiB-puroB-GAPDH	868	AseI	1049	GSC 2779
pGLEX41-[REP]	2119	BspHI	868	GSC 2239
pGLEX41-R6K-ampiA-[REP]	865	BspHI	779	GSC 2240
pGLEX41-R6K-ampiB-[REP]	1751	BspHI	806	GSC 2249
pGLEX-R6K-ampiA-neoA	890	BspHI	764	GSC 2214
pGLEX-R6K-ampiB-neoB	767	BspHI	654	GSC 2244
pGLEX-R6K-ampiA-puroA	708	BspHI	659	GSC 2220
pGLEX-R6K-ampiB-puroB	574	BspHI	746	GSC 2213

Exemplo 2: Transfecção de células com vetores de expressão:

1. Materiais e Métodos:

[0159] Células CHO-S e células HEK293 -

[0160] As células são os hospedeiros preferidos para expressar proteínas porque elas são capazes de corrigir as dobras, arranjo e modificação de pós-transcrição das proteínas recombinantes. A linha de célula CHO foi utilizada devido elas serem bem caracterizadas e não servir como um hospedeiro para muitas viroses humanas patogênicas, tornando elas um hospedeiro relativamente seguros para produção de proteína terapêutica estável. As células de ovário de hamsters chineses (CHO-S, Invitrogen, Carlsbad, CA, USA) foram cultivadas em suspensão em meio PowerCHO-2 CD (Lonza, Verviers, Bélgica), suplementado com 4 mM de L-glutamina (AppliChem, Alemanha), e incubada em um incubador com agitação (200 rpm com um movimento circular de 2,5 cm) a 37°C, 5% de CO<sub>2</sub> e 80% de umidade. As células HEK293 são utilizadas devido ao fato de elas serem fácil para transfectar e permitir a produção rápida de proteínas recombinantes até a quantidade de gramas inferior. As células utilizadas são células HEK293-EBNA (ATCC, Manassas, VA) e são rotineiramente cultivadas em suspensão e meio Ex-cell 293 (Sigma-Aldrich, St. Louis, MI).

[0161] As subculturas de células CHO-S e HEK293-EBNA foram rotineiramente realizadas a cada 3-4 dias usando uma densidade para semeação de  $0,5 \times 10^6$  células viáveis/ml em meio fresco. As células forma cultivadas usando 10 ml de meio em tubos de bioreator de 50 ml (Tubespin Bioreator 50; TPP, Transadingen, Suíça) contendo um filtro permeável permitindo a troca de gás. A viabilidade celular e concentração foram determinadas com o contador de célula automatizado Countess (Invitrogen, Carlsbad, CA, USA) usando o método de exclusão

celular Trypan Blue. A concentração celular foi confirmada através da determinação do método do volume celular embalado (PCV) usando tubos PVC (TPP< Trasadingen, Suíça) para células CHO-S.

#### Volume de célula embalada (PCV)

[0162] O método PCV é baseado na centrifugação de um volume específico do líquido de cultura em um tubo de mini-PCV (PCV Packed Cell Volume Tube; TPP, Transadingen, Suíça) por 1 minuto a 5000 rpm. Durante a centrifugação, as células são pelotizadas no capilar graduado na base do tubo. A porcentagem do volume celular embalado é então determinada pela avaliação do volume de pelotas em relação à quantidade do fluido de cultura celular centrifugado. Por exemplo, 1% de PCV indicou que 10 $\mu$ l de pelotas celulares estava presente em 1 ml de fluido de cultura.

[0163] Para contagem de células de rotina, 200  $\mu$ l de cada amostra foi pipetada em um tubo PCV e o volume da pelota correspondente (em  $\mu$ l) foi lido com uma régua (dispositivo de medida "de fácil leitura"; TPP, Transadingen, Suíça). Este volume foi multiplicado por 5 ter o valor para 1 ml e então foi multiplicado usando um fator de correlação específica celular para obter uma estimativa de concentração de células viáveis (em milhões de células/ml).

#### Contagem de célula "automatizada":

[0164] A concentração celular e viabilidade foram determinadas com o contador celular automatizado Countess® (Invitrogen, Carlsbad, CA, USA) em mistura de amostra com a mesma quantidade de Trypan Blue. A solução é então pipetada dentro da lâmina da câmara Countess® antes de ser lida pelo instrumento. Este instrumento permite uma leitura automática

da câmara de Neubauer que, após calibração, determina a viabilidade celular e a concentração de células mortas e vivas.

#### Análise da citometria de fluxo:

[0165] A citometria de fluxo é uma técnica para análise de parâmetros múltiplos de células individuais. Esta técnica permite que a análise quantitativa e qualitativa das células sejam fenotipicamente diferentes umas das outras, por exemplo, mortas de células viáveis (de acordo com o tamanho e a granularidade de células). Também permite que a quantificação de células que expresse uma proteína de interesse, tal como GFP. As células foram coletadas a partir da cultura por pipetas estéreis de 300  $\mu$ l de amostras e foram analisadas com um citômetro de fluxo de calibração da classificação celular associada à fluorescência (FACS) (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) equipado com um laser de argônico resfriado a ar emitindo em 488 nm. As análises foram feitas com o software CellQuest. A emissão GFP foi detectada com o FL-1, usando um filtro de passagem de banda de 530/30 nm.

[0166] No primeiro portal, os resíduos celulares bem como as células mortas foram excluídos a partir da análise em pontos traçados SSC/FSC em escala linear. Então, a fluorescência GFP de células vivas foi mostrada em um histograma em escala logarítmica. O valor médio da distribuição de fluorescência foi utilizado para avaliar o nível de expressão GFP das populações de células analisadas.

#### Método de quantificação de IgG: OCTET QK-

[0167] O sistema Octet QK (FortéBio, Menlo Park, CA, USA) desempenha quantificação livre do marcador de anticorpos,

proteínas, peptídeos, DNA e outras biomoléculas e provê caracterização cinética de interações de ligação biomolecular. Uma correlação entre a taxa de ligação (nm) e a concentração IgG1 acumulada ( $\mu\text{g}/\text{ml}$ ) da amostra permite a quantificação da titulação IgG com uma curva de calibração.

[0168] As amostras celulares foram centrifugadas por 5 minutos a 300g. O sobrenadante foi então diluído (1/5 para o anticorpo IgG1) com o tampão Octet em uma placa de 96 poços antes de ser analisado com o Octet usando os biosensores de Proteína A (Proteína A DIP e READ<sup>TM</sup> Biosensor, Forté Bio, USA) para obter a concentração do anticorpo por poço.

Transfecção transitória usando JetPEI:

[0169] A transfecção transitória e estável de células CHO-S e HEK293-EBNA foram realizadas usando polietilenimina (PEI; JetPEI, transfecção-PolyPlus, Illkirch, França). PEI é um polímero catiônico que pode complexar com as moléculas carregadas negativamente tal como DNA. O complexo DNA-PEI com carga positiva se liga à superfície celular carregada negativamente é internalizada pela endocitose. Ela alcança o compartimento de lisossomo a partir de onde ele é liberado pela lise para o núcleo. A eficiência de alta transfecção com os complexos DNA-PEI é devida à capacidade de PEI e proteger o DNA a partir da degradação lisossomal. As células foram transfectadas de acordo com o manual provido pelo fabricante.

[0170] Todos os plasmídeos foram linearizados antes da transfecção estável (100  $\mu\text{g}$  de DNA ressuspenso em 100  $\mu\text{l}$  de Tris-EDTA, pH 7,5). Para transfecção transitória, plasmídeos circulares foram diretamente utilizados a partir da midipreparação de DNA. Neste estudo, as transfecções transitórias foram mantidas em tubos de 50 ml do biorreator e

nenhum antibiótico foi adicionado.

[0171] Os clones CHO-S estáveis expressando IgG1 e GFP foram obtidos através da co-transfecção de um vetor de expressão e dois vetores de resistência (codificantes para resistência de puromicina ou neomicina, respectivamente).

Seleção de grupamentos estáveis e mini-agrupamentos:

[0172] A eficiência da transfecção foi determinada 24 horas após a transfecção pela citometria de fluxo (citômetro BD FACS Calibur, # 1293) através da análise da expressão GFP intracelular. Se a porcentagem de células positivas GFP foi maior que 20%, as células transfectadas foram diluídas com meio seletivo e distribuídas em placas de 96 poços (para limitar a diluição para gerar mini-agrupamentos estáveis) ou frascos T (para gerar os agrupamentos estáveis). O meio seletivo utilizado foi o PowerCHO-2, 4mM de glutamina, suplementado com concentrações diferentes de geneticina e puromicina.

[0173] Sete dias após a transfecção após a transfecção, ao rigor da seleção foi renovado pela adição de meio de seleção para as células. Assim que as colônias nos 96 poços foram confluentes, as placas foram lidas usando um leitor de fluorescência.

[0174] O agrupamento em frascos T foi expandido para escala de tubos spin usando 4 mM de L-glutamina PowerCHO-2, livre de antibiótico. Sua viabilidade e concentração foram avaliados com o contador de célula automatizado Countess (Invitrogen, Carlsbad, CA, USA). Assim que a densidade celular permitiu, o início da formação do grão para cada agrupamento através das células semeadas em uma densidade de  $0,5 \times 10^6$  células/ml em 10 ml do meio em tubos de biorreator de 50 ml (incubados em um

agitador (200 rpm) a 5% de CO<sub>2</sub>, 37°C e 80% de umidade). Cada formação de semente foi passada duas vezes por semana pelo semeio das células em 0,5x10<sup>6</sup> células/ml no meio de crescimento (a concentração celular foi determinada através da análise PCV). A formação semeada foi utilizada para inocular todas as corridas de produção (batelada).

[0175]

[0176] Para as produções das próximas 4-5 semanas, as corridas foram semeadas uma vez por semana em duplicada. A estabilidade do agrupamento foi avaliada por FACS e expressão de IgG como descrito acima para populações de clonagem.

Corridas de produção (fermentação em batelada).

[0177] As corridas de batelada dos grupamentos celulares semeados em uma concentração de 0,5x10<sup>6</sup> células/ml usando a formação semeada para inoculação e as células foram então cultivadas por 7 dias no meio Feed. No dia 4 e 8, 200 ml de células foram centrifugadas durante 5 minutos a 300g e o sobrenadante foi analisado quanto a IgG acumulada usando o Octed. Em adição, a expressão GFP de cada batelada foi analisada através de FACS.

## 2. Resultados:

### 2.1- Expressão transitória em células CHO:

[0178] Os vetores comparados neste estudo diferem principalmente em sua cadeia principal. O cassete de expressão total (promotor, primeiro ítron, constructo de expressão, poli(A)) é exatamente a mesma para todos os vetores. Os vetores são derivados do vetor pGLEX41 como descrito no Exemplo 1. Em um vetor, o gene de resistência a ampicilina foi códon otimizado para expressão em *E. coli* e a cadeia principal da bactéria foi reduzida a um mínimo:

pGLEX41-R6K-AmpiA-[REP] (em A curto). Em um segundo vetor, o gene de resistência a ampicilina foi códon otimizado para expressão em *E. coli*, mas todas as sequências CpG foram evitadas, pelo uso de códons alternados (quando possível): este vetor é chamado pGLEX41-R6K-AmpiB-[REP] (em B curto). A terceira modificação incluiu o uso das sequências de flanco GAPDH que foram clonadas à montante e à jusante do cassete de expressão dos vetores A e B resultando em vetores pGLEX41-R6K-AmpiA-[REP]-GAPDH (em GAPDH\_A curto) e pGLEX41-R6K-AmpiB-[REP]-GAPDH (em GAPDH\_B curto).

[0179] As transfecções transitórias de células CHO-S (Invitrogen) foram feitas de modo a comparar o nível de expressão das proteínas repórter expressas no contexto da cadeia principal do plasmídeo diferente. As transfecções (em duplicita) foram realizadas em tubos do bioreator de 50 ml (TPP, Transadingen, Suíça) usando 10 ml de volume de meio final e analisado no dia 5 após transfecção por Octet (Figura 2).

[0180] Todos os vetores (A e B) com a cadeia principal corrigida mostram um nível de expressão levemente maior do que os vetores controle pGLEX41. Existe apenas uma diferença mínima entre os vetores A e B. Isto é esperado, devido apenas a diferença na cadeia principal que é resistente à ampicilina, o que poderia não ter um impacto sobre a expressão transitória.

[0181] A observação mais impressionante é o efeito positivo das sequências GAPDH sobre a expressão. Um nível de expressão 2 vezes maior é obtido com o plasmídeo colhido nas sequências de flanco de GAPDH comparada a uma sem as sequências GAPDH. Isto é verdade para ambos os constructos A e B. comparado ao

vetor pGLEX41, uma expressão 3 vezes maior pode ser observada. Isto é ainda mais surpreendente se o tamanho do plasmídeo for levado em consideração. O vetor A (7048 bps) é quase metade do tamanho comparado ao vetor GAPDH-A (13436 bps). Portanto, assume-se que a quantidade do DNA liberado durante o processo de transfecção transitória é o mesmo para todos os plasmídeos, apenas metade da quantidade molar de GAPDH-A é liberada ao núcleo.

## 2.2- Expressão transitória nas células HEK293:

[0182] As transfecções transitórias de células HEK293 EBNA foram feitas de modo a comparar o nível de expressão das proteínas repórter expressas no contexto de diferentes cadeias principais de plasmídeos. As transfecções (em duplicata) foram realizadas em tubos de biorreator de 50 ml (TPP, Transadingen, Suíça) usando 10 ml de volume de meio final e forma analisadas no dia 10 após transfecção por Octet (Figura 3).

[0183] Os resultados mostrados na figura 3 mostram um aumento significante na expressão que pode ser obtido usando as regiões dos flancos GAPDH nas células HEK293 EBNA. O vetor GAPDH está mostrando um aumento triplo na expressão, enquanto o vetor GAPDH-A mostra um aumento ainda maior na expressão de 5 vezes. Estes vetores não contém o elemento oriP e podem, portanto, ter um potencial para titulações ainda mais altas.

## 2.3 - Expressão em linhas de células CHO estáveis: estabelecimento das células transfectadas estáveis.

[0184] As populações estáveis foram geradas através da co-transfecção de um vetor de expressão e vetores codificantes dos genes de resistência, seguido pela seção da pressão mediada por antibióticos. A seleção da pressão foi removida

14 dias após a transfecção. Estas etapas permitiu a geração de mini-agrupamentos estáveis e agrupamentos estáveis que foram cultivados em intervalos regulares nas corridas de produção de modo a comparar os níveis de expressão das proteínas repórter (anticorpo IgG1 e GFP) de constructos diferentes e a estabilidade da expressão.

Estudo da expressão da proteína repórter sobre as corridas de produção realizadas com grupamentos celulares:

[0185] Os agrupamentos foram gerados pela transfecção estável. Durante o procedimento de seleção (os primeiros 14 dias após a transfecção) os agrupamentos foram analisados por meio de FACS. Um aumento na fração celular GFP positiva junto com a viabilidade da cultura poderia ser observado em relação ao tempo. A seleção da pressão mediada pelos antibióticos foi removida a partir do grupamento após 14 dias. Usando esta abordagem nenhum grupamento celular transfectado com o plasmídeo "B" poderia ser obtido. O nível de expressão dos grupamentos gerados foi avaliado assim que as células pudesse ser cultivadas em tubos de biorreator de 50 ml. As bateladas foram feitas em duplicatas. As células foram analisadas pelo FACS para expressão GFP e o acúmulo de IgG no sobrenadante foi avaliado por Octet após 8 dias de expressão.

[0186] Uma relação proporcional poderia ser observada entre a titulação IgG e a expressão de GFP dos agrupamentos. Portanto, apenas os dados de IgG estão mostrados na figura 4. Todos os grupamentos transfectados com os vetores contendo a sequência GAPDH mostraram expressão maior comparada ao vetor pGLEX41 ou com o mesmo vetor sem a sequência GAPDH (fator de 2,8 entre a e A-GAPDH). Nenhuma conclusão poderia ser conseguida entre B e B-GAPDH quando nenhum grupamento B

sobrevive.

[0187] As transfecções realizadas com A-GAPDH e B-GAPDH induziu uma expressão maior de IgG (2,7 e 3,5 vezes mais respectivamente) do que a transfecção de pLEX41 (para a batelada 2). Portanto, nos grupamentos, as sequências de flancos GAPDH mostram ser favoráveis para a produção de proteínas. Finalmente, as transfecções realizadas com os vetores B-GADPH induziu uma expressão maior de IgG do que a transfecção realizada com A-GAPDH (fator de 1,25). Portanto, a redução CpG nos genes de resistência mostra ser favorável para a produção estável de proteínas, também.

Estudo do nível de expressão nas populações de clonagem:

[0188] As células foram transfectadas e distribuídas em placas de 96 poços no meio seletivo de modo a obter as populações clonais ou oligoclonais. Após 7 dias, a seleção da pressão foi refrescada pela adição do meio seletivo para as células. A expressão de GFP foi avaliada 14 dias após a transfecção pelo uso de um leitor de placa ELISA. Os resultados estão mostrados nas figuras 5.

[0189] A confirmação dos resultados obtidos no grupamento celular, as células transfectadas com os vetores contendo as sequências de flancos GAPDH expressou significativamente mais GFP do que as mesmas cadeias principais sem as sequências à jusante e à montante GAPDH (fatores de 1,7 a 2 vezes) ou outros vetores utilizados como controle (pLEX41: 2,5 vezes) (figura 5). Em adição, a população com vetores contendo sequências de resistência que foi reduzida CpG (B) induziu uma expressão maior que os vetores correspondentes que foram apenas códon otimizado (A) (1,5 vezes entre A e B; 1,2 vezes entre B e B-GAPDH).

[0190] A partir do estudo de expressão, várias conclusões poderiam ser extraídas. Primeira, a sequência GAPDH à montante e à jusante GAPDH permite maior expressão do que o vetor padrão que foi utilizado como uma referência (pGLEX41). Ademais, um nível de expressão menor foi obtido quando as células foram transfectadas com a mesma cadeia principal do vetor sem as sequências GAPDH confirmado que o efeito benéfico da expressão é relacionado às sequências dos flancos GAPDH inserido. Em adição, a redução do número de CpG na expressão e seleção dos plasmídeos parece ser levemente favorável para expressão também.

Exemplo 3: Nível de expressão transitória de células CHO-S GMP transfectadas com novos vetores desenhados.

[0191] Tem sido descrito na literatura que a região 5' do promotor GAPDH abriga um potencial de insulina bem como um elemento de resposta de éster forbol (Alexander-Bridges et al., (1992) "Advan. Enzyme Regul.", 32: 149-159). O elemento de resposta de éster forbol (-1040-1010 bps) está situado à montante do qual é usualmente referido como o promotor GAPDH (-488 - + 20). Em um estudo de deleção realizado em linhas de células H35 Hepatoma, os autores não foram capazes de demonstrar um efeito significante da deleção de pares de base -1200 para -488 (relativa ao ponto inicial de transcrição). Portanto, o elemento de resposta éster forbol pode não ser funcionalmente ligado à direção de expressão do promotor de GAPDH. No entanto, um experimento de transfecção transitória foi realizado de modo a avaliar a combinação de insulina e PMA (forbol-12-meristato-13-acetato, o éster forbol mais comum) no aumento, em transição e expressão estável que foi observado usando os plasmídeos contendo os elementos de

flancos GAPDH.

[0192] De modo a obter o meio de crescimento livre de insulina, PowerCHO<sup>2</sup> foi preparado a partir do meio de energia e sem insulina foi adicionado. PMA foi comprado da Sigma (St. Louis, MO), e foi dosado em uma concentração final de 1,6  $\mu$ M (correspondente à concentração utilizada por Alexander-Bridges em linhas de célula H35 Hepatoma) em PowerCHO2 (+/- insulina).

[0193] As transfecções foram realizadas em tubos de biorator de 50 ml (Tubespins, TPP, Trasadingen, Suíça) como descrito previamente. De modo a evitar a presença de insulina provida pelo OptiMEM (Life Technologies, Carlsbad, CA), o meio de transfecção foi mudado para RPMI 1640 (PAA, Pasching, Áustria) suplementado com 4 mM de Gln e 25 mM HEPES. Após a transfecção, as células foram distribuídas em placas de 12 poços e 1 ml de quatro meios diferentes foi adicionado (PowerCHO2, 4 mM Gln, +/- insulina; PowerCHO2, 4 mM de Gln, 1,6  $\mu$ M de PMA, +/- insulina). Novamente, o constructo repórter expressando IgG1 e GFP usando dois IRES foi utilizado (descrito no exemplo 2). Este vetor permitiu a verificação da eficiência de transfecção. A porcentagem e a viabilidade de células transfectadas foram observadas similares em todas as quatro preparações de meio diferente.

[0194] Como mostrado na figura 6, nenhum efeito significante da depleção de insulina e/ou adição de PMA poderia ser observada durante este experimento. Títulos similares foram obtidos em todos os meios utilizados para expressão. Isto sugere que o éster forbol potencial e os elementos de resposta de insulina presente na sequência de flancos à montante do gene GAPDH não afetam a expressão do

transgene transitório.

**Exemplo 4:** Análise da fragmentação dos flancos do DNA no cassette de expressão GAPDH à montante do promotor, e à montante do sítio poli-A, de modo a estudar o efeito sobre a expressão do gene repórter.

[0195] O *locus* GAPDH humano é localizado no cromossomo 12 do genoma humano. GAPDH é descrito por ser constitutivamente ativo em todas as células de origem de mamífero, quando a enzima é a peça chave no metabolismo de glicose. À montante do promotor, o gene GAPDH é flanqueado por NCAPD2, um gene que abrange mais do que 30000 bps. À jusante do sítio de poliadenilação, o gene GAPDH é flanqueado pelo IFFO1 (ver figura 7 para detalhes).

[0196] Não apenas GAPDH e o promotor, mas também as regiões dos flancos são bem conservadas entre as espécies diferentes (ver tabela 4).

Tabela 4: Extensão da alta homologia entre as regiões dos flancos GAPDH humanos, de ratos e de camundongos. A análise foi feita usando gerenciamento do clone 9 (ScieED, Cary, NC, USA). O número é relativo à primeira base do elemento do flanco à montante ou à jusante, respectivamente (Sequência ID NO: 7 e Sequência ID NO: 8, respectivamente). As sequências utilizadas para alinhamento foram para as bases de camundongos 532-3731 (à montante) e 8164-11364 (à jusante) da Sequência ID No. 18 e para as bases de ratos 719-3918 (à montante) e 8495-11058 (à jusante) da Sequência ID NO. 19.

Região à montante				Região à jusante			
Sequências de homologia (rato)		Sequências de homologia (camundongo)		Sequências de homologia (rato)		Sequências de homologia (camundongo)	
>80%	>90%	>80%	>90%	>80%	>90%	>80%	>80%
161-249	279-331	15-69	278-329	1608-1764	1706-1764	1614-1671	1904-2061
256-338	554-623	159-249	546-626	1894-2067	1912-20611	1888-2072	2927-3071
515-659		273-342				2918-3082	
2296-2349		515-647					
2381-2513		1143-1223					
2736-2818		1957-2009					
		2029-2080					
		2375-2485					
		2730-2821					

[0197] Uma comparação da homologia do DNA entre roedores e humanos mostra um mínimo de conservação do DNA de 38%. A presença de uma extensão conservada do DNA fora de uma região do promotor ou uma região codificante para um gene indica que pode existir uma pressão de seleção sobre a célula para manter a sequência de DNA ou para permitir apenas determinadas/mínimas mudanças. Em um caso específico, as regiões dos flancos GAPDH podem ser importantes para as células devido ao fato de elas manterem um nível de expressão alto dos genes de GAPDH. As mudanças na sequência de DNA conduzindo a diminuição da expressão seriam contra seleção.

[0198] De modo a avaliar a contribuição do elemento GAPDH à montante e à jusante para o aumento observado na expressão, constructos foram feitos contendo apenas a região do flanco

GAPDH à montante (SEQ ID NO:7), fragmentos à montante da região do flancos GAPDH ou a região do flancos GAPDH à jusante (SEQ ID NO:8). O anticorpo do tipo IgG1 repórter foi expresso por um constructo IRES (cadeia leve-IRES-cadeia pesada), portanto, evitando a co-transfecção de plasmídeos múltiplos. Os detalhes sobre a fragmentação do fragmento GAPDH à montante são mostrados na figura 8. Os fragmentos seguintes da região dos flancos GAPDH foram utilizados. O fragmento 1 (SEQ ID NO:9), fragmento 2 (SEQ ID NO:10), o fragmento 3 (SEQ ID NO:11), fragmento 4 (SEQ ID NO:12), fragmento 8 (SEQ ID NO:13), fragmento 9 (SEQ ID NO:14), fragmento 11 (SEQ ID NO:15), fragmento 17 (SEQ ID NO:16).

[0199] A região dos flancos de GAPDH à montante (SEQ ID NO:7) utilizada contém 2 - 3 vezes (no total 6) nucleotídeos do sítio de restrição NruI do qual três são ligados ao DNA genômico em sua extremidade 5' e três estão ligadas ao DNA genômico em sua extremidade 3'. A região dos flancos GAPDH à jusante (SEQ ID NO:8) utilizadas contém duas vezes 3 (no total 6) nucleotídeos do sítio de restrição ScaI dos quais três estão ligados ao DNA genômico em sua extremidade 5' e três estão ligados ao DNA genômico em sua extremidade 3'. A região do flanco GAPDH à montante e a região do flanco de GAPDH à jusante sem os nucleotídeos do respectivo sítio de restrição são mostrados na SEQ ID NO:20 (região do flanco GAPDH à montante sem sítios de restrição) e SEQ ID NO:21 (região do flanco GAPDH à montante sem os sítios de restrição). Os fragmentos da região do flanco GAPDH à montante utilizado contém cada um 3 nucleotídeos do respectivo sítio de restrição em sua extremidade 5' e/ou 3' ligada ao DNA genômico (Fragmento 1 contém 3 nucleotídeos do

sítio de restrição NruI em sua extremidade 5'; Fragmento 2 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 3'; fragmento 3 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 5'; fragmento 4 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 3'; fragmento 8 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 3'; fragmento 9 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 5' e 3 nucleotídeos do sítio de restrição NruI em sua extremidade 3'; fragmento 11 contém 3 nucleotídeos do sítio de restrição NruI em sua extremidade 3'). O fragmento 17 não contém nucleotídeos de um sítio de restrição. Os fragmentos à montante da região dos flancos GAPDH sem os nucleotídeos do respectivo sítio de restrição são mostrados na SEQ ID NO:22 (fragmento 1 sem sítio de restrição), SEQ ID NO:23 (fragmento 2 sem o sítio de restrição) SEQ ID NO:24 (fragmento 3 sem o sítio de restrição), SEQ ID NO:25 (fragmento 4 sem o sítio de restrição), SEQ ID NO:26 (fragmento 8 sem o sítio de restrição), SEQ ID NO:27 (fragmento 9 sem o sítio de restrição), SEQ ID NO:28 (fragmento 11 sem o sítio de restrição).

[0200] O efeito dos elementos GAPDH à montante e à jusante sobre a expressão foi avaliado no dia 10 após a transfecção usando o Octet (ForteBio, Menlo, CA, USA) de modo a quantificar a quantidade da IgG1 secretada no sobrenadante (ver figura 9). O pGLEX41, o vetor original apresenta resultados de expressão inferiores (80%) comparado ao desenho do novo vetor melhorado utilizado na cadeia principal do pGLEX41-ampA. Comparado à cadeia principal do pGLEX41 original, o novo desenho inclui códon de otimização do gene

ampiA necessário para resistência à ampicilina em *E. coli*, uma origem diferente de replicação (R6K ao invés da origem pUC de replicação) e eliminação das sequências ligantes desnecessárias (ou espaçadoras) de origem bacteriana. Ambos os vetores tem aproximadamente o mesmo tamanho.

[0201] Surpreendentemente, o pGLEX41-ampiA incluindo o elemento à montante (SEQ ID NO:7) e à jusante (SEQ ID NO:8), (denominado pGLEX41 à montante/à jusante, na figura 9 mostrando os resultados de expressão) apresenta maior expressão (fator 1,5) quando comparado ao mesmo vetor sem as sequências à montante e à jusante. Se um considera a diferença no tamanho (fragmentos à montante/jusante aumenta o tamanho do plasmídeo em aproximadamente 6000 bps) e, portanto, as diferenças nas cópias de plasmídeos liberadas durante a transfeção, o efeito pode ser mais importante em uma base por plasmídeo.

[0202] O vetor contendo apenas o fragmento à montante (superior) está mostrando um nível de expressão similar ao do constructo de expressão original pGLEX41-ampiA. O vetor contendo apenas o fragmento à jusante (inferior) está mostrando um aumento significante (fator 1,2) na expressão comparada ao do constructo de expressão original pGLEX41-ampiA. Um aumento adicional na expressão pode ser observado se ambos, os fragmentos à montante e à jusante estiverem presentes. Isto é confirmado pela fragmentação dos fragmentos à montante. O fragmento 9 e o promotor proximal ao fragmento 8 não mostram qualquer diferença na expressão quando comparado ao pGLEX41-ampiA. Os fragmentos 1, 11 e 17 mostram um aumento na expressão. O maior aumento foi observado para o fragmento 4. Deve ser destacado que o promotor proximal ao

fragmento 8 não está mostrando qualquer efeito. Portanto, o aumento na expressão não pode ser explicado pelas sequências previamente publicadas (Alexander-Bridges et al., (1992) "Advan. Enzyme Regul.", 32:149-159), Graven et al., (1999) "Biochimica et Biophysic Acta", 147: 203-218).

[0203] Interessantemente, os fragmentos 2 e 3 conduzem a uma diminuição significante na expressão. Isto é inesperado, especialmente em vista do fato de que estes fragmentos clonados na direção oposta (anti-sentido (AS) na figura 9) não causam este efeito. Para os fragmentos 1, 8, 9, 11 e 17 nenhuma diferença na expressão foi observada para os fragmentos que foram integrados no sentido ou na orientação anti-sentido (dados não mostrados). O fragmento 11, apesar de uma parte do fragmento 2, não mostrou este efeito. Portanto, o elemento da sequência que parece ser prejudicial à expressão poderia estar pelo menos parcialmente sobre o fragmento BstBI-BstBI que foi deletado no fragmento 2 de modo a obter o fragmento 11.

[0204] Em adição, a hipótese de que um elemento negativo está localizado (pelo menos parcialmente) sobre o fragmento BstBI-BstBI é suportada pelo aumento na expressão observado entre o fragmento 3 (que inclui o fragmento BstBI-BstBI) e o fragmento 1.

[0205] Embora pareça fácil localizar o fragmento tendo um efeito negativo (BstBI-BstBI), a partir deste estudo é menos óbvio como este efeito negativo observado para os fragmentos 2 e 3 é compensado pelos elementos da sequência presentes no fragmento à montante completo. Pode ser que este efeito negativo seja balanceado pelo pequeno efeito positivo que foi observado pelo fragmento 1 e o fragmento 4 (as o aumento na

expressão para o fragmento 1 foi menor que para o fragmento 4). Todavia, o efeito positivo para o fragmento 4 (fator 1,25) observado parece menos importante comparado ao efeito negativo (fator 0,4). Além disso, o fragmento 9, que é a região completa à montante sem o fragmento BstBI-BstBI não mostra a expressão aumentada comparada a região de flanco DE GAPDH à montante total (no entanto, o fragmento 9 inclui o fragmento EcoRV-BstBI que é parte dos fragmentos 2 e 3 e pode ter um efeito negativo sobre a expressão).

[0206] Pode ser especulado apenas em relação ao mecanismo por trás dos efeitos observados. A dependência da orientação do efeito negativo sobre a expressão observada com os fragmentos 2 e 3 exclui a expressão das estruturas de abertura de leitura não identificadas (por exemplo, expressão de um ncRNA), devido a não existência de promotores ao redor, o que poderia despertar a expressão de apenas uma orientação. O fato de que a expressão é reduzida abaixo do nível basal mostra não apenas a ausência de um efeito positivo (por exemplo, uma atividade melhorada), mas também a presença de um efeito negativo dependente da orientação.

[0207] Em resumo, um aumento surpreendente de expressão na transição nas células CHO é observado se ambas as regiões dos flancos, a região à montante e a região à jusante, estão presentes no plasmídeo de expressão. Apesar de o fragmento 4 parecer ter um efeito positivo significante sobre a expressão, nenhum fragmento sozinho poderia ter sido identificado como responsável pelo aumento total da expressão que foi observado. O aumento de expressão do vetor de expressão pGLEX41-ampiA (à jusante/à montante) parece ser o efeito resumido de ambas, as regiões de flancos à montante e

à jusante.

Exemplo 5: Clonagem da sequência de DNA genômica não traduzida à montante do gene GAPDH de hamsters chinês e do promotor do hamsters chinês -

[0208] 1.1- Clonagem da sequência de DNA genômica não traduzida à montante do gene GAPDH de hamsters chinês dentro de um vetor de expressão.

[0209] A sequência de DNA genômico não traduzida à montante do gene GAPDH foi ampliada a partir do DNA genômica de células CHO-S (Life Technologies) através do PCR. O DNA genômico foi extraído como descrito no exemplo 1. Os constructos foram preparados usando o promotor CMV de camundongo ou o promotor GAPDH de hamsters chineses para a expressão do constructo de gene repórter {REP} descrito no Exemplo 1.

[0210] Para clonagem da sequência de DNA genômico à montante do gene GAPDH de hamsters chinês em combinação com o promotor CMV de camundongo, os primers GlnPr1896 e GlnPr1897 foram utilizados para amplificação dos fragmentos de 3 kbs (bps 672 a 3671 da SEQ ID NO:29) usando o protocolo de PCR descrito para o Exemplo 1 e conduzindo ao amplicon com a SEQ ID NO:30. O amplicon contém a sequência de DNA genômico à montante do gene GAPDH de hamsters chinês e os sítios de restrição 5' e 3' que foram introduzidos pelos primers.

[0211] Para clonagem da sequência de DNA genômico à montante do gene GAPDH de hamsters chinês em combinação com o promotor de GAPDH de hamsters chineses, os primers GlnPr1902 e GlnPr1905 foram utilizados de modo a amplificar o fragmento de 3508 bps contendo a sequência de DNA genômico incluindo a sequência de DNA genômico à montante do gene GAPDH de

hamsters chinês e o promotor GAPDH (bps 672 a 4179 da SEQ ID NO:29) conduzindo ao amplicon com a SEQ ID NO:31. Em um segundo PCR, GlnPr1901 e GlnPr1902 foram utilizados para amplificação do fragmento de 508 bps contendo apenas a região do promotor (bps 3672 a 4179 da SEQ ID NO:29), conduzindo à SEQ ID NO:32). O ítron utilizado no vetor "A" (descrito no Exemplo 1) foi amplificado usando primers GlnPr1903 e GlnPr1904.

[0212] Uma primeira fusão por PCR foi realizada com primers GlnPr1904 e GlnPr1901 usando o amplicom com a SEQ ID NO:32 e o amplicon com a sequência ítron como molde. O amplicon contém o promotor GAPDH de hamsters chinês, um ítron e os sítios de restrição 5' e 3' que foram introduzidos pelos primers. Todos os primers estão mostrados na tabela 5.

[0213] Uma segunda fusão por PCR foi realizada com primers GlnPr1905 e GlnPr1904 usando o amplicon com a SEQ ID NO:31 e o amplicon com a sequência ítron como molde. O amplicon contém o promotor GAPDH de hamsters chinês, um ítron e os sítios de restrição 5' e 3' que foram introduzidos pelos primers.

[0214] Após a purificação em um gel de agarose a 1%, as bandas de interesse foram cortadas e purificadas usando o kit "NucleoSpin Gel e PCR Clean-up" (Macherey Nagel, Oensingen, Suíça). Os fragmentos purificados foram clonados dentro do plasmídeo pCR\_Blunt usando o kit de clonagem do PCR Zero-Blunt (Invitrogen, Carlsbad, CA, USA). Os produtos de ligação foram transformados dentro de *E. coli* competentes TOP10 (One Shot® TOP10 Competent *E.coli*; Invitrogen, Carlsbad, CA, USA) e analisadas por meio da análise de restrição e minipreps. Isto conduz aos plasmídeos pCR\_b blunt[CHO-à montante-GAPDH],

contendo a sequência de DNA genômico à montante do gene GAPDH de hamsters chinês, pCR-Blunt[CHO-A montante-GAPDH\_GAPDH-promotor] contendo a sequência de DNA genômico à montante do gene de GAPDH de hamsters chinês e o promotor GAPDH e o ítron a partir do vetor "A" e o pCR-Blunt[CHO-GAPDH-promotor] contendo o promotor GAPDH e o ítron do vetor "A".

[0215] Para avaliação dos amplicon sobre seus efeitos na expressão de um gene secretado, o vetor "A" (descrito no Exemplo 1) foi utilizado. Como utilizado previamente, o cassete de expressão utilizado neste vetor contém um gene de policistrônico codificante para uma IgG1 secretada e GFP (ver Exemplo 1). As células transfetadas irão secretar, portanto, o anticorpo monoclonal IgG1 e acumular intracelularmente o GFP de uma maneira dependente.

[0216] De modo a liberar o fragmento inserido de 3 kb, contendo a sequência de DNA genômico à montante do gene GAPDH de hamsters chinês, o plasmídeo pCR\_Blunt[CHO-A montante-GAPDH] foi digerido usando a enzima de restrição NaeI. Este inserto foi clonado na cadeia principal de "A", digerido usando a enzima de restrição NruI e CIPed (CIP; NEB, Ipswich, MA, USA). A cadeia principal e o inserto foram ligados juntos usando a T4 DNA ligase (T4 DNA Ligase, NEB, Ipswich, MA, USA) e, subsequentemente, transformada em *E. coli* competente PIR1. Os clones foram selecionados para preparação miniprep e, subsequentemente análise de restrição. O plasmídeo resultante foi chamado de "A\_GAPDH-UP", confirmado pela análise de sequenciamento e produzido em escala minipreps usando o kit NucleoBond Xtra Midi (Macherey Nagel, Oensingen, Suíça).

[0217] Para a clonagem dos constructos de expressão usando o promotor GAPDH de hamsters chinês, os fragmentos inseridos

foram o plasmídeo pCR\_Blunt[CHO - à montante-GAPDH\_GAPDH-promotor] e pCR\_Blunt[CHO-GAPDH-promotor] através da digestão usando as enzimas de restrição NheI e NruI. Os fragmentos resultantes foram clonados na cadeia principal do vetor "A", aberta usando a mesma enzima e CPIed. Após ligação com a T4 DNA Ligase e transformação em *E. coli* PIR1 competente, os clones foram selecionados para análise de restrição miniprept. Os plasmídeos resultantes foram chamados de "A\_GAPDH\_UP\_Prom" (plasmídeo com sequência de DNA genômico não traduzido à montante do gene GAPDH de hamsters chinês e do promotor) e "A\_PR" (plasmídeo apenas com o promotor) confirmado pela análise de sequenciamento e produzido em escala minipreps usando o kit NucleoBond-Xtra Midi (Macherey Nagel, Oensingen, Switzerland).

2. Avaliação do efeito da sequência de DNA genômico não traduzido à montante do gene GAPDH de hamsters chineses sobre a expressão do constructo do gene repórter.

[0218] As células CHO-S transfetadas em tubos de biorreator usando 10 ml de volume do meio (como descrito no Exemplo 2). As células transfetadas foram incubadas em um incubador com agitação com 200 rpm de agitação a 37°C, 5% de CO<sub>2</sub> e 80% de umidade. Os sobrenadantes das células foram analisados para expressão da IgG1 usando o sistema Octet QK com os bio-sensores de Proteína A, (FortéBio, Menlo Park, CA, USA). Os resultados estão mostrados na figura 10.

[0219] O nível de expressão do plasmídeo contendo o promotor GAPDH ("A\_PR") comparado ao promotor CMV de camundongo (A) foi reduzido em 50%, indicando que o promotor GAPDH de hamsters chineses não é tão forte quanto o promotor viral. O plasmídeo contendo a sequência de DNA genômico não

traduzida do gene GAPDH de hamsters chinês em combinação com o promotor GAPDH de hamsters chinês ("A\_GAPDH\_UP\_Prom") mostra um aumento de duas vezes na expressão comparado ao constructo tendo apenas o promotor GAPDH ("A\_PR"). O plasmídeo contendo a sequência de DNA genômico não traduzida à montante do gene de hamsters chinês GAPDH e o promotor CMV de camundongo ("A\_GAPDH\_UP") mostra a maior expressão e um aumento maior que 40% sobre o plasmídeo contendo apenas o promotor CMV de camundongo ("A"). Isto confirma que a sequência de DNA genômico não traduzido à montante do gene GAPDH de hamsters chinês tem um efeito melhorado sobre a expressão da proteína repórter.

Tabela 5

Primers utilizados para clonagem no Exemplo 5

Primer	SEQ ID NO:	Sequência	Oriente- tação	Sítio de restrição
GlnPr1896	SEQ ID NO:33	TACGGCCGGCTTCACTGTACAGTGGC ACAT	à frente	NaeI
GlnPr1897	SEQ ID NO:34	TCAGGCCGGCCGTGGTTCTTCGGTAG TGAC	reverso	NaeI
GlnPr1901	SEQ ID NO:35	TACTCGCGAAGAAGATCCTCAACTTT TCCACAGCC	à frente	NruI
GlnPr1902	SEQ ID NO:36	GTTCACTAACGAGCTCTGCTATTAA TAGGAACTGGGGTG	Reverso	/
GlnPr1903	SEQ ID NO:37	CACCCCAGTTCCCTATAAATAGCAGAG CTCGTTAGTGAAC	à frente	/
GlnPr1904	SEQ ID NO:38	CGCTAGCACCGGTCGATCGA	Reverso	NheI
GlnPr1905	SEQ ID NO:39	TACTCGCGATTCACTGTACAGTGGCA CATAC	à frente	NruI

REIVINDICAÇÕES

1. Cassete de expressão, caracterizado pelo fato de compreender um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e elemento de melhora da expressão, sendo que o elemento de melhora de expressão comprehende uma sequência de DNA genômico não traduzida à jusante de um promotor da desidrogenase 3-fosfato-gliceraldeído (GAPDH) eucariótico, onde o polipeptídeo codificado pela sequência de polinucleotídio não é GAPDH, e onde a sequência de DNA genômico não traduzido à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo a posição de nucleotídeo em torno de +1 até a posição de nucleotídeo em torno de +7000, sendo que a posição de nucleotídeo é relativa ao inicio da transcrição do mRNA GPDH, e sendo que o comprimento da sequência de DNA genômico não traduzida à jusante ao promotor GAPDH eucariótico é em torno de 100 para em torno de 15000 nucleotídeos.
2. Cassete de expressão, de acordo com a reivindicação 1, caracterizado pelo fato de compreender ainda uma sequência de DNA genômica não traduzidas a montante de um promotor GAPDH eucariótico, sendo que a sequência de DNA genômica não traduzida a montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição de nucleotídeo em torno de -3500, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, e sendo que o comprimento da sequência de DNA genômica não traduzida a montante do promotor GAPDH eucariótico é em torno de 100 para em torno de 15000 nucleotídeos.
3. Cassete de expressão, caracterizado pelo fato de

compreender um promotor, uma sequência de polinucleotídio codificante de um polipeptídeo, e uma sequência de DNA genômico não traduzida a montante de um promotor GAPDH eucariótico, sendo que o polipeptídeo codificado por uma sequência de polinucleotídio não é GAPDH, e sendo que a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo em torno da extremidade 5' do promotor GAPDH eucariótico para a posição de nucleotídeo em torno de -3500, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, sendo que o comprimento da sequência de DNA genômica não traduzida a montante do promotor GAPDH eucariótico está em torno de 100 para em torno de 15000 nucleotídeos, com a previsão de que o cassete de expressão não compreende um promotor GAPDH eucariótico ou fragmentos dos mesmos.

4. Cassete de expressão, de acordo com a reivindicação 3, caracterizado pelo fato de compreender ainda uma sequência de DNA genômico não traduzida à jusante de um promotor GAPDH eucariótico, sendo que a sequência de DNA genômico não traduzida a montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo a partir da posição de nucleotídeo em torno de +1 a posição de nucleotídeo em torno de +7000, sendo que a posição de nucleotídeo é relativa ao início da transcrição do mRNA de GAPDH, e sendo que o comprimento da sequência de DNA genômica não traduzida a montante do promotor GAPDH eucariótico está em torno de 100 para em torno de 15000 nucleotídeos.

5. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de a

sequência de DNA genômico não traduzida a montante e/ou à jusante do promotor GAPDH eucariótico não ser operativamente ligada à sequência de polinucleotídio codificante de polipeptídeo.

6. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de o cassete de expressão compreender ainda um sítio de poliadenilação.

7. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de o comprimento da sequência de DNA genômico não traduzido à jusante do promotor GAPDH eucariótico ser em torno de 10 nucleotídeos e estender-se em seu máximo para o segundo do último ítron do gene IFF01 ou para uma parte do mesmo.

8. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de a sequência de DNA genômico não traduzida a montante do promotor GAPDH eucariótico iniciar à jusante do sítio de poliadenilação GAPDH eucariótica a montante de o promotor GAPDH eucariótico ser pelo menos em torno de 100 nucleotídeos e estender-se ao seu máximo para o segundo dos últimos ítron do gene IFF01.

9. Cassete de expressão, de acordo com a reivindicação 2, caracterizado pelo fato de o comprimento da sequência de DNA genômico não traduzido a montante do promotor de GAPDH eucariótico ser pelo menos em torno de 100 nucleotídeos e estender-se em seu máximo para o códon inicial do gene NCAPD2.

10. Cassete de expressão, de acordo com a reivindicação 2, caracterizado pelo fato de o comprimento da sequência de DNA genômico não traduzida à jusante do promotor de GAPDH

eucariótico ser pelo menos em torno de 100 nucleotídeos e estender-se em seu máximo no terceiro do último ítron do gene NCAPD2.

11. Cassete de expressão, de acordo com a reivindicação 3, caracterizado pelo fato de o comprimento da sequência de DNA genômico não traduzido à jusante do promotor de GAPDH eucariótico ser pelo menos de 100 nucleotídeos e estender-se em seu máximo para o códon inicial do gene NCAPD2.

12. Cassete de expressão, de acordo com a reivindicação 3, caracterizado pelo fato de o comprimento da sequência de DNA genômico não traduzida à jusante do promotor de GAPDH eucariótico ser pelo menos 100 nucleotídeos e estender-se em seu máximo para o terceiro último ítron do gene NCAPD2.

13. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de a sequência de DNA genômico não traduzido a montante e/ou à jusante do promotor de GAPDH eucariótico ser de origem mamária.

14. Cassete de expressão, de acordo com a reivindicação 13, caracterizado pelo fato de a sequência de DNA genômico não traduzido a montante e/ou à jusante do promotor de GAPDH eucariótico ser de origem de roedores ou de humanos.

15. Cassete de expressão, de acordo com qualquer uma das reivindicações 1 ou 4, caracterizado pelo fato de a sequência de DNA genômico não traduzido a montante do promotor de GAPDH eucariótico compreender a sequência de nucleotídeo selecionada a partir do grupo consistindo de SEQ ID NOs: 8 e 21 ou fragmentos dos mesmos.

16. Cassete de expressão, de acordo com qualquer uma das reivindicações 1 ou 4, caracterizado pelo fato de a sequência

de DNA genômico não traduzida à montante do promotor de GAPDH eucariótico compreender a sequência de nucleotídeos complementar à sequência de nucleotídeos selecionada a partir do grupo consistindo de SEQ ID NOS: 8 e 21 ou fragmentos das mesmas.

17. Cassete de expressão, de acordo com qualquer uma das reivindicações 1 ou 4, caracterizado pelo fato de a sequência de DNA genômico não traduzido à montante do promotor de GAPDH eucariótico compreender a sequência de nucleotídeos pelo menos 80% idêntica à sequência de nucleotídeos selecionada a partir do grupo consistindo de SEQ ID NOS: 8 e 21 ou fragmentos das mesmas.

18. Cassete de expressão, de acordo com qualquer uma das reivindicações 2 ou 3, caracterizado pelo fato de a sequência de DNA genômico não traduzida à jusante do promotor de GAPDH eucariótico compreender uma sequência de nucleotídeos selecionada a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 e 28 ou fragmentos dos mesmos.

19. Cassete de expressão, de acordo com a reivindicação 18, caracterizado pelo fato de as sequências de nucleotídeos selecionados a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27, e 28 ou fragmentos dos mesmos compreende cinco ou menos modificações de ácido nucléico.

20. Cassete de expressão, de acordo com qualquer uma das reivindicações 2 ou 3, caracterizado pelo fato de a sequência de DNA genômico não traduzido à jusante do promotor de GAPDH eucariótico compreender uma sequência de nucleotídeos complementares para a sequência do nucleotídeo selecionado a

partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 e 28 ou fragmentos dos mesmos.

21. Cassete de expressão, de acordo com qualquer uma das reivindicações 2 ou 3, caracterizado pelo fato de a sequência de DNA genômico não traduzido à jusante do promotor de GAPDH eucariótico compreender uma sequência de nucleotídeos pelo menos 80% idêntica à sequência do nucleotídeos selecionada a partir do grupo consistindo das SEQ ID NOS: 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 24, 25, 26, 27 e 28 ou fragmentos dos mesmos.

22. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de o promotor e a sequência de polinucleotídio codificante de um polipeptídeo serem operativamente ligadas.

23. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de a sequência de DNA genômico não traduzida a montante do promotor de GAPDH eucariótico ser orientada na mesma direção que a sequência de polinucleotídio codificante de um polipeptídeo.

24. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de a sequência de DNA genômico não traduzida a montante do promotor de GAPDH eucariótico ser orientada em uma direção oposta em relação a da sequência de polinucleotídio codificante de um polipeptídeo.

25. Cassete de expressão, de acordo com qualquer uma das reivindicações 2 ou 3, caracterizado pelo fato de a sequência de DNA genômico não traduzida à jusante do promotor de GAPDH

eucariótico ser orientada na mesma direção que a sequência de polinucleotídio codificante de um polipeptídeo.

26. Cassete de expressão, de acordo com qualquer uma das reivindicações 2 ou 3, caracterizado pelo fato de a sequência de DNA genômico não traduzida à jusante do promotor de GAPDH eucariótico ser orientada na direção oposta em relação a sequência de polinucleotídio codificante de um polipeptídeo.

27. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de o promotor ser selecionado a partir do grupo consistindo do promotor SV40, promotor MPSV, CMV de camundongo, tk humano, CMV humano, CMV de rato, EF1 alfa humano, EF1 alfa de hamsters Chinês, GAPDH humano, promotores híbridos, incluindo os promotores MYC, HYK e CX.

28. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de o polipeptídeo ser selecionado a partir do grupo consistindo de anticorpos, fragmentos de anticorpos ou derivados dos anticorpos.

29. Cassete de expressão, de acordo com a reivindicação 6, caracterizado pelo fato de o sítio de poliadenilação ser selecionado a partir do grupo consistindo de BGH poli(A) e SV40 poli(A).

30. Cassete de expressão, de acordo com qualquer uma das reivindicações de 1 a 4, caracterizado pelo fato de compreender adicionalmente um elemento genético selecionado a partir do grupo consistindo de um promotor adicional, um melhorador, elementos de controle de transcrição, e um marcador selecionável.

31. Cassete de expressão, de acordo com a reivindicação 30,

caracterizado pelo fato de o elemento genético é um marcador selecionável onde o conteúdo dos sítios CpG contido na sequência de polinucleotídio codificante dos marcadores selecionáveis ser 45 ou menos.

32. Vetor de expressão, caracterizado pelo fato de compreender um cassete de expressão, definido em qualquer uma das reivindicações de 1 a 29.

33. Vetor de expressão, caracterizado pelo fato de compreender na ordem: (a) uma sequência de DNA genômico não traduzida a montante e/ou à jusante de um promotor GAPDH eucariótico; (b) um promotor; (c) uma sequência de polinucleotídios codificantes de um polipeptídio; (d) um sítio de poliadenilação; (e) um melhorador; (f) uma sequência de DNA genômico não traduzida à jusante e/ou a montante de um promotor GAPDH eucariótico; ou

(a) uma sequência de DNA genômico não traduzida a montante e/ou à jusante de um promotor GAPDH eucariótico

(b) um melhorador;

(c) um promotor;

(d) uma sequência de polinucleotídios codificantes de um polipeptídio;

(e) um sítio de poliadenilação;

(f) uma sequência de DNA genômico não traduzida à jusante e/ou a montante de um promotor GAPDH eucariótico; ou

(a) um melhorador;

(b) uma sequência de DNA genômico não traduzida a montante e/ou à jusante de um promotor GAPDH eucariótico;

(c) um promotor;

(d) uma sequência de polinucleotídios codificantes de um polipeptídio;

(e) um sítio de poliadenilação;

(f) uma sequência de DNA genômico não traduzida à jusante e/ou a montante de um promotor GAPDH eucariótico;

onde a inclusão do melhorador é opcional, e sendo que o polipeptídeo codificado pela sequência de polinucleotídio não é GAPDH, e sendo que a sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo a partir da posição de nucleotídeo em torno de +1 a posição de nucleotídeo em torno de +7000, onde a posição do nucleotídeo é relativa ao inicio da transcrição do mRNA de GAPDH, e onde o comprimento da sequência de DNA genômico não traduzida à jusante do promotor GAPDH eucariótico está em torno de 100 a 15000 nucleotídeos e sendo que a sequência de DNA genômico não traduzida a montante do promotor GAPDH eucariótico inicia dentro de uma região abrangendo a partir da extremidade 5' do promotor GAPDH eucariótico para a posição de nucleotídeo em torno de -3500, onde a posição de nucleotídeo é relativa ao início da transcrição do mRNA do GAPDH, e sendo que o comprimento da sequência de DNA genômico não traduzido a montante de um promotor GAPDH eucariótico é em torno de 100 para 15000 nucleotídeos, com a provisão e que se (a) ou (b) for uma sequência de DNA genômico não traduzido a montante de um GAPDH eucariótico, (f) é uma sequência de DNA genômico não traduzido à jusante de um GAPDH eucariótico, e se (a) ou (b) for uma sequência de DNA genômico não traduzido à jusante de um GAPDH eucariótico, (f) é uma sequência de DNA genômico não traduzido à jusante de um GAPDH eucariótico.

34. Vetor de expressão, de acordo com qualquer uma das reivindicações 32 ou 33, caracterizado pelo fato de o vetor

compreender ainda um elemento genético selecionado a partir do grupo consistindo de um promotor adicional, um melhorador, elementos de controle de transcrição, uma origem de replicação e um marcador selecionável.

35. Vetor de expressão, de acordo com qualquer uma das reivindicações 32 ou 33, caracterizado pelo fato de o vetor compreender ainda uma origem de replicação e um marcador selecionável onde o conteúdo dos sítios CpG contido na sequência de polinucleotídio do vetor de expressão codificante da origem de replicação e o marcador selecionável ser 200 ou menos.

36. Célula hospedeira, caracterizada pelo fato de compreender um cassete de expressão conforme definido em qualquer uma das reivindicações de 1 a 31 ou um vetor de expressão definido em qualquer uma das reivindicações de 32 a 35.

37. Cassete de expressão, conforme definido em qualquer uma das reivindicações de 1 a 31, ou o vetor de expressão, definido em qualquer uma das reivindicações 32 a 35, caracterizado pelo fato de ser utilizado como um medicamento para o tratamento de um distúrbio.

38. Cassete de expressão, conforme definido em qualquer uma das reivindicações de 1 a 31, ou o vetor de expressão, definido em qualquer uma das reivindicações 32 a 35, caracterizado pelo fato de ser utilizado na terapia gênica.

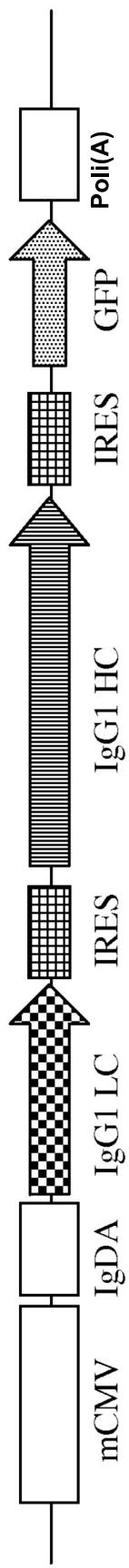
39. Método *in vitro* para expressão de um polipeptídeo, caracterizado pelo fato de compreender a transfecção de uma célula hospedeira com o cassete de expressão, conforme definido em qualquer uma das reivindicações de 1 a 31 ou o vetor de expressão, definido em qualquer uma das reivindicações de 32 a 35 e recuperação do polipeptídeo.

40. Método, de acordo com a reivindicação 39, caracterizado pelo fato de o cassete de expressão ou o vetor de expressão ser estavelmente transfetado.

41. Método, de acordo com a reivindicação 39, caracterizado pelo fato de o cassete de expressão ou vetor de expressão ser transitoriamente transfetado.

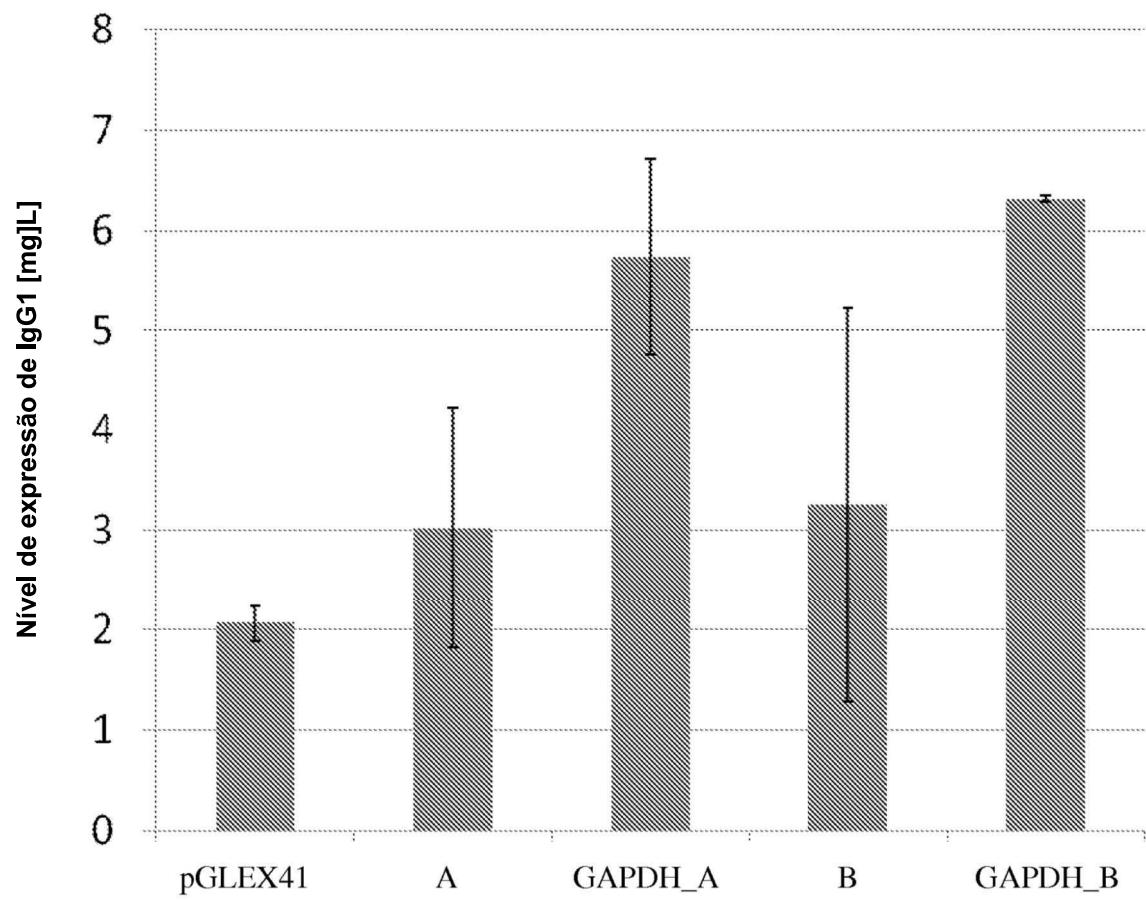
42. Uso de um cassete de expressão ou vetor de expressão, conforme definido em qualquer uma das reivindicações de 1 a 31 ou nas reivindicações de 32 a 35, caracterizado pelo fato de ser para expressão de um polipeptídeo heterólogos a partir de uma célula hospedeira de mamífero.

**1/8**

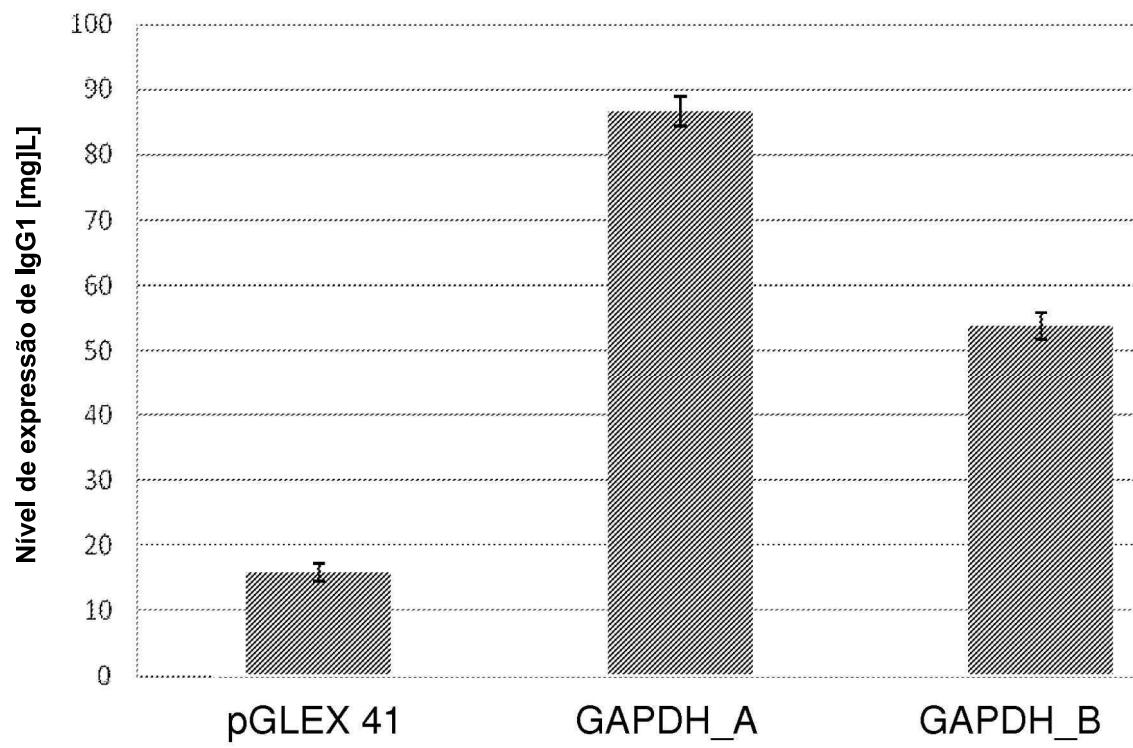


**FIG.1**

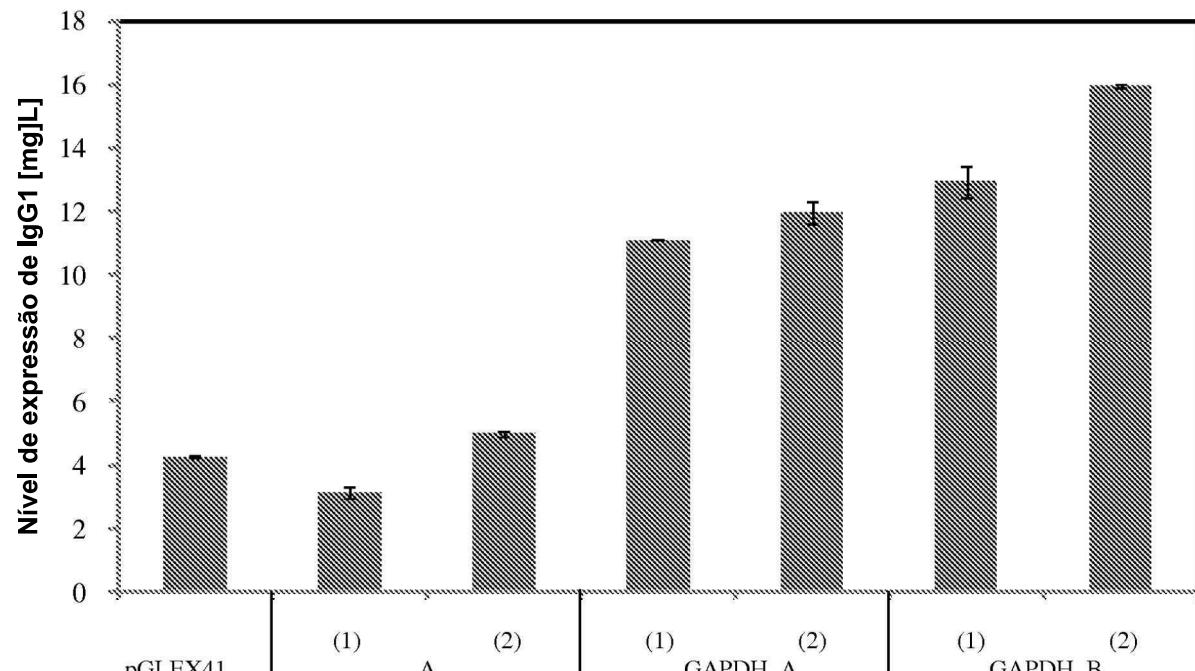
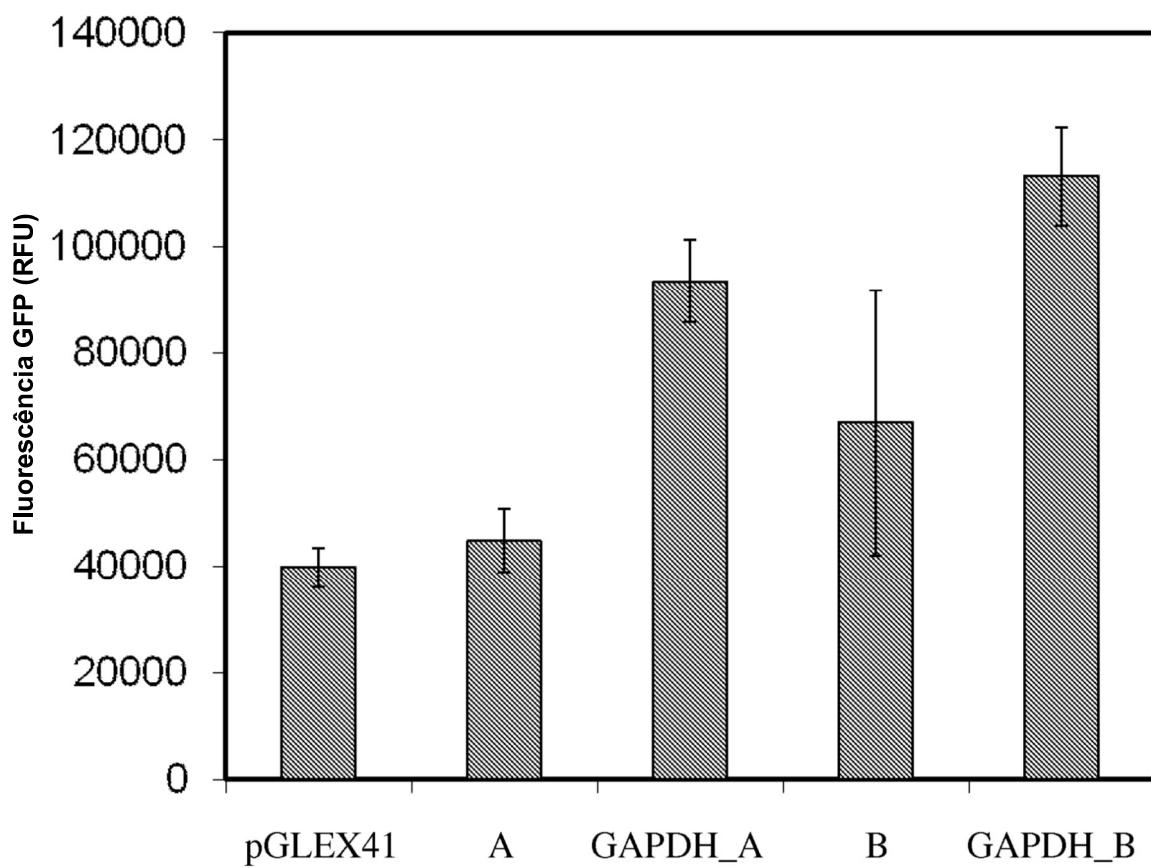
**2/8**



**FIG.2**



**FIG.3**

**FIG.4****FIG.5**

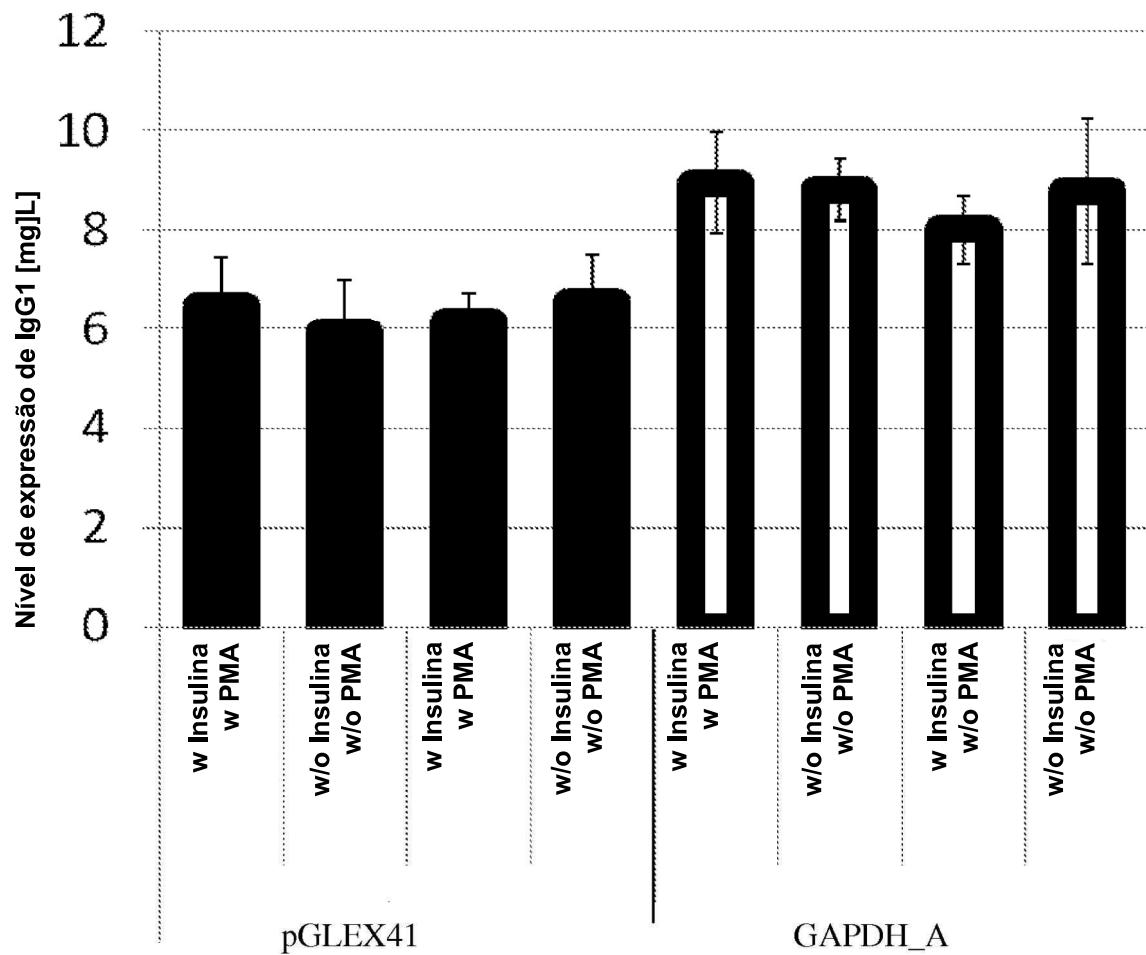


FIG.6

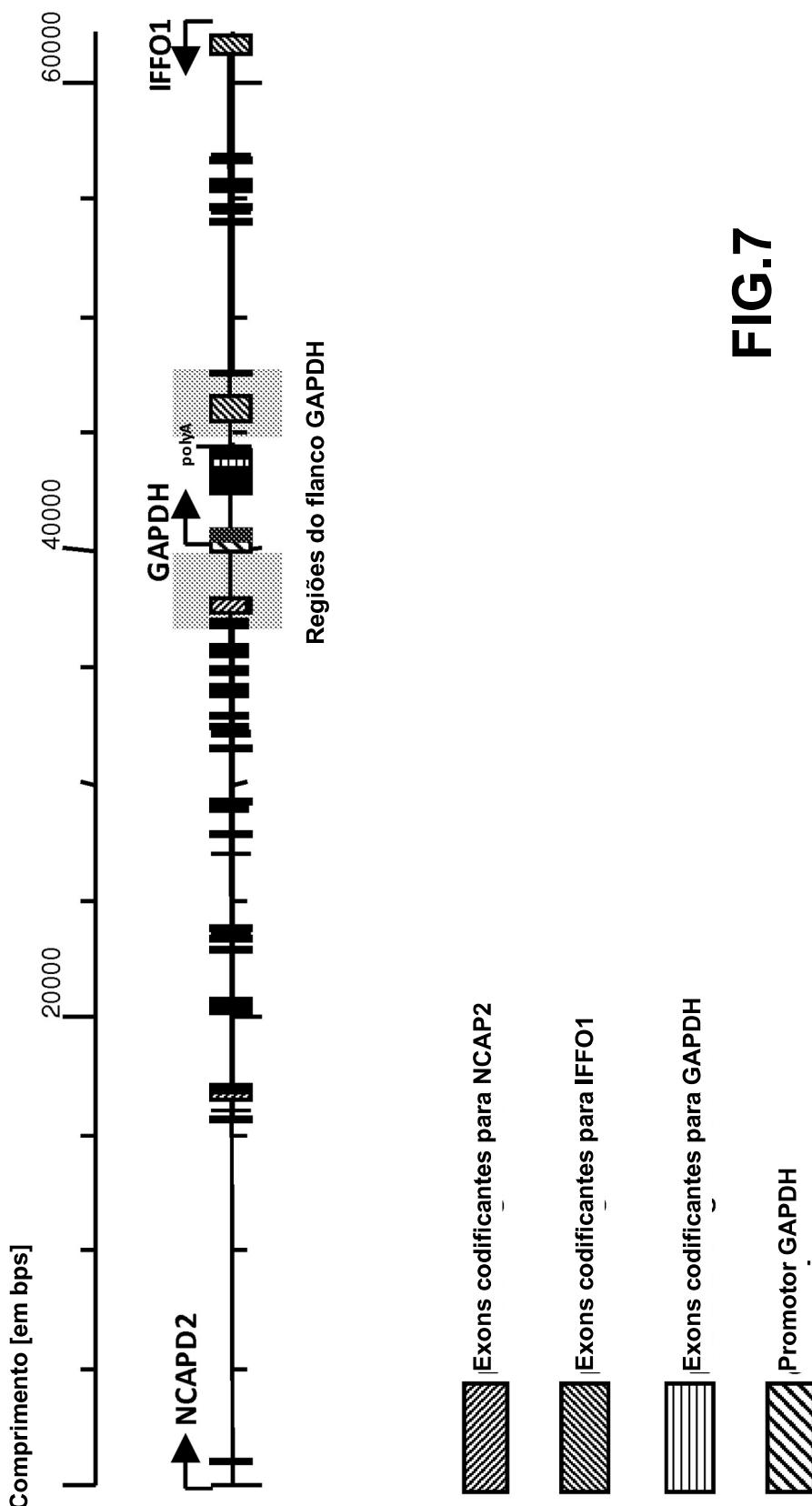
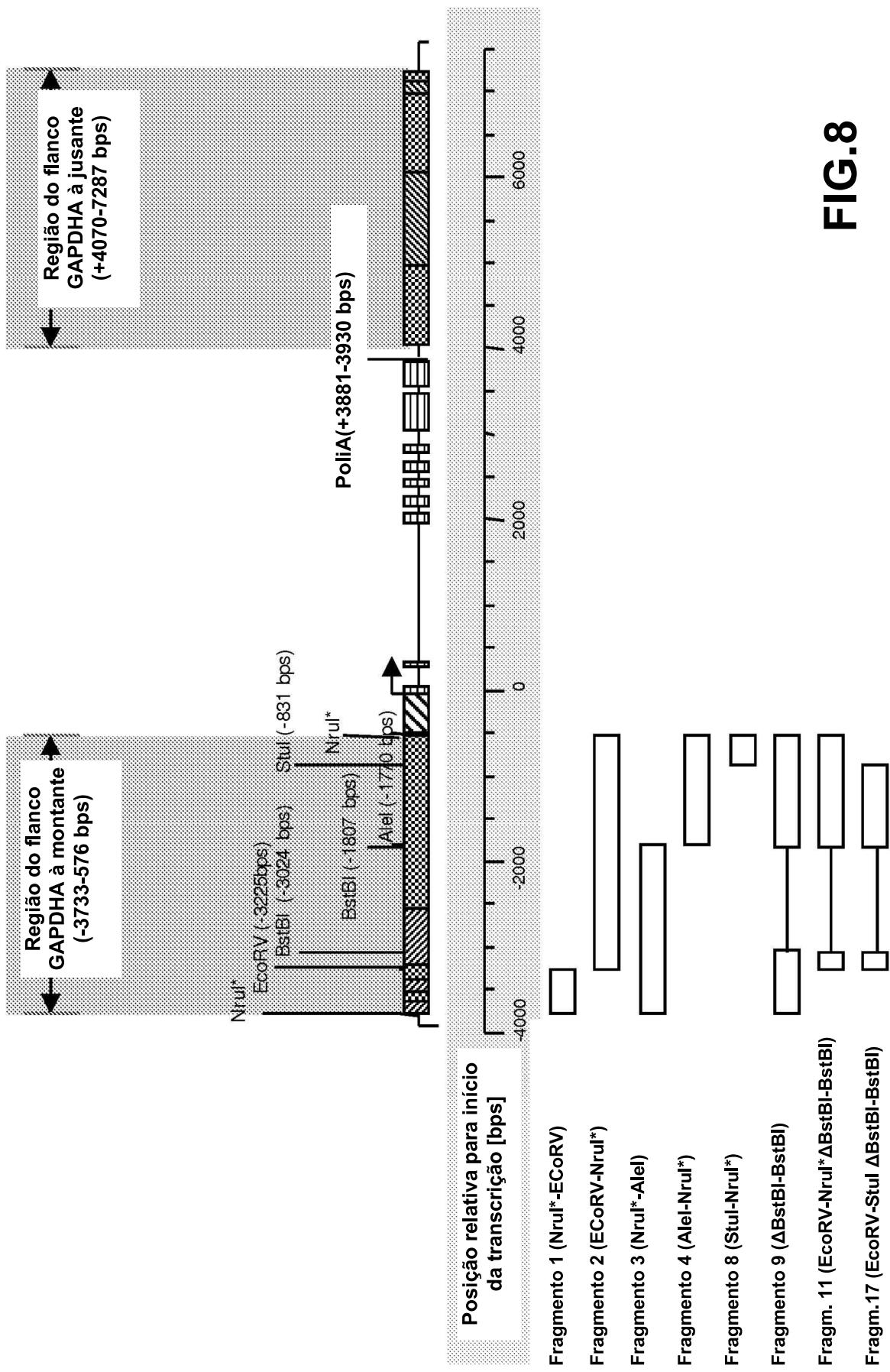
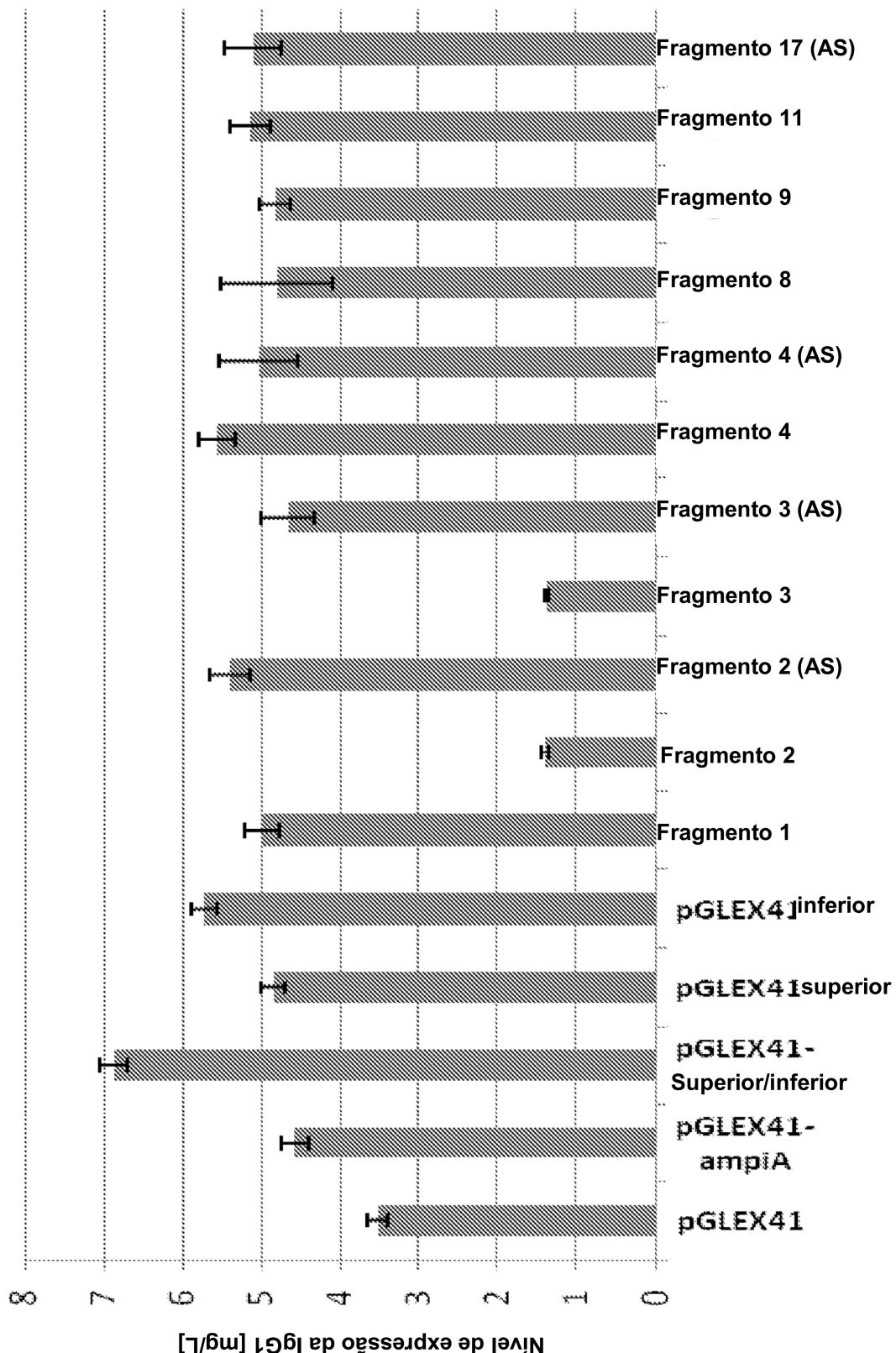


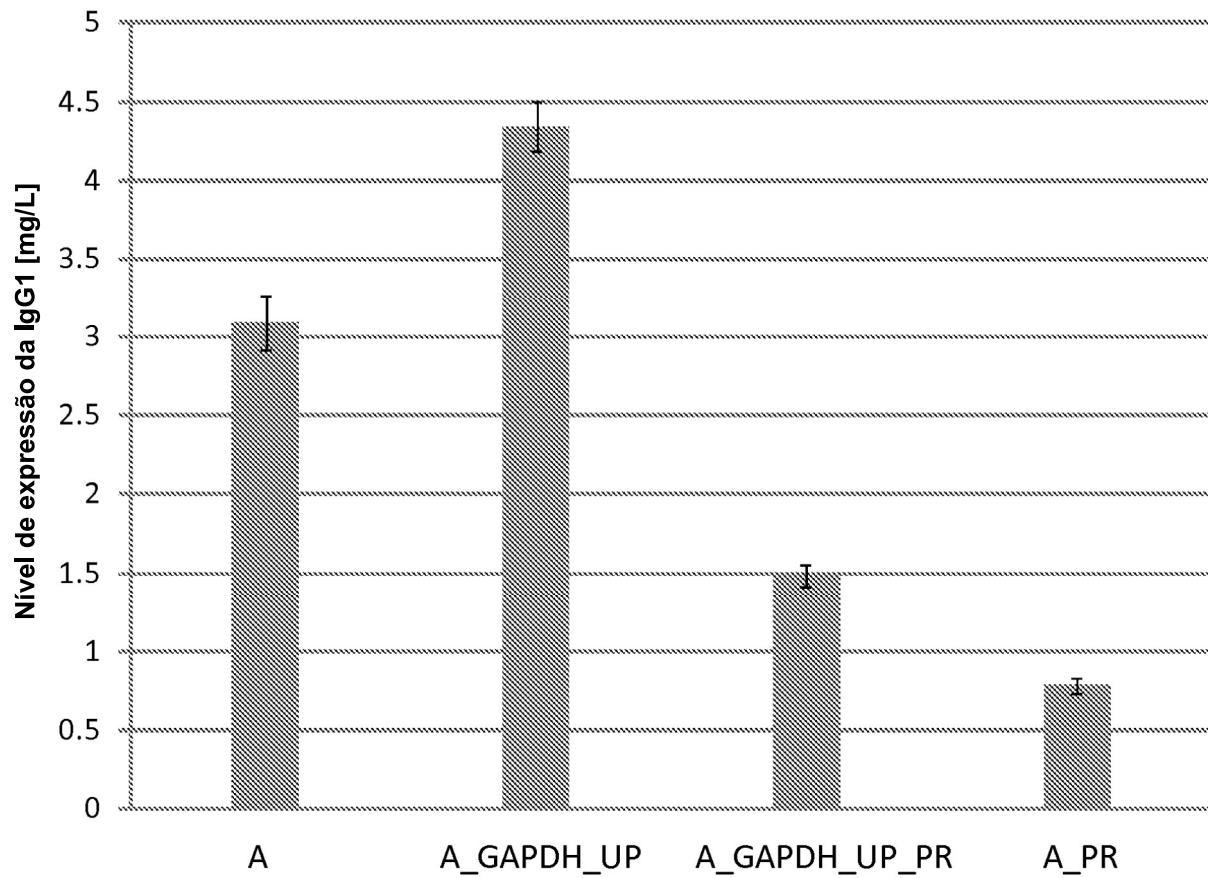
FIG.7

**FIG.8**

**7/8**



**FIG.9**

**FIG.10**

RESUMO

“CASSETE DE EXPRESSÃO, VETOR DE EXPRESSÃO, CÉLULA HOSPEDEIRA, MÉTODO *IN VITRO* PARA EXPRESSÃO DE UM POLIPEPTÍDEO E USO DE UM CASSETE DE EXPRESSÃO OU VETOR DE EXPRESSÃO”

A presente invenção refere-se a um cassete de expressão útil para a expressão de uma sequência de polinucleotídios codificantes de um polipeptídeo.

Este anexo apresenta o código de controle da listagem de sequências biológicas de que trata a Resolução INPI 228 de 11/11/2009:

**Código de Controle**

Campo 1



435F81C0EC67B3C2

Campo 2



B5BA8AB01D4CA666

**Outras Informações:**

- Nome do Arquivo: GLEM015.txt
- Data de Geração do Código: 03-06-2014
- Hora de Geração do Código: 16:22:29
- Código de Controle:
  - Campo 1: 435F81C0EC67B3C2
  - Campo 2: B5BA8AB01D4CA666