



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

(21) BR 112020007413-6 A2



(22) Data do Depósito: 16/10/2018

(43) Data da Publicação Nacional: 08/12/2020

(54) Título: VETORES ADENOVIRAIS COM DOIS CASSETES DE EXPRESSÃO

(51) Int. Cl.: C12N 15/861; A61K 39/235.

(30) Prioridade Unionista: 16/10/2017 US 62/572,944.

(71) Depositante(es): GLAXOSMITHKLINE BIOLOGICALS S.A..

(72) Inventor(es): STEFANO COLLOCA.

(86) Pedido PCT: PCT EP2018078210 de 16/10/2018

(87) Publicação PCT: WO WO/2019/076880 de 25/04/2019

(85) Data da Fase Nacional: 14/04/2020

(57) Resumo: A presente invenção refere-se a um vetor adenoviral símio que compreende dois cassetes de expressão, em que cada cassete de expressão compreende um transgene e um promotor, e em que o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral símio e o segundo cassete de expressão é inserido em uma região do vetor adenoviral compatível com a replicação do vetor.

(i) RC1



(ii) RC3:



(iii) RC2:



Relatório Descritivo da Patente de Invenção para “**VETORES ADENOVIRAIS COM DOIS CASSETES DE EXPRESSÃO**”.

LISTAGEM DE SEQUÊNCIAS

[001] O presente pedido de patente contém uma listagem de sequências que foi submetida eletronicamente no formato ASCII e é incorporada por referência em sua totalidade. A dita cópia ASCII, criada em 11 de outubro de 2018, é denominada VU66441A_WO_SL.txt e tem 178.926 bytes de tamanho.

CAMPO DA INVENÇÃO

[002] Esta invenção está no campo de vetores adenovirais recombinantes. A invenção refere-se a um vetor adenoviral compreendendo dois cassetes de expressão. Em particular, a invenção refere-se a um adenovírus símio como um adenovírus de chimpanzé compreendendo dois cassetes de expressão.

ANTECEDENTES DA INVENÇÃO

[003] Os adenovírus recombinantes são úteis na terapia genética e como vacinas.

[004] Os adenovírus humanos têm sido amplamente utilizados para aplicações de transferência de genes devido à sua grande capacidade transgênica e capacidade de obter transferência de genes altamente eficiente em uma variedade de tecidos alvo.

[005] No entanto, a maioria dos seres humanos é exposta e desenvolve imunidade aos adenovírus humanos. Portanto, há uma demanda por vetores que efetivamente entregam moléculas a um alvo e minimizam o efeito da imunidade preexistente aos sorotipos de adenovírus humano. Os adenovírus símios são eficazes nesse sentido; estão suficientemente relacionados aos vírus humanos para serem eficazes na indução de imunidade a antígenos exógenos entregues aos quais os seres humanos têm pouca ou nenhuma imunidade preexistente. Portanto, vetores virais baseados em adenovírus símios podem fornecer

uma alternativa ao uso de vetores adenovirais derivados de seres humanos para o desenvolvimento de vacinas baseadas em ácidos nucleicos.

[006] Os adenovírus com defeito de replicação entregam seu genoma no interior de uma célula e, por não se replicarem, não amplificam a carga útil do transgene. Tipicamente, o gene E1 é substituído por um cassete de transgene compreendendo um promotor de escolha e uma sequência de ácidos nucleicos correspondente a um gene ou genes de interesse, resultando em um vírus recombinante com defeito de replicação.

[007] Há uma necessidade na técnica de adenovírus recombinantes melhorados.

SUMÁRIO DA INVENÇÃO

[008] A presente invenção refere-se a um vetor adenoviral símio compreendendo dois cassetes de expressão. Em particular, a invenção refere-se a um adenovírus símio como um adenovírus de chimpanzé compreendendo dois cassetes de expressão. Exemplos de adenovírus de chimpanzé adequados incluem ChAd155 e ChAd83.

[009] Os vetores adenovirais da invenção são úteis como componentes de composições imunogênicas para a indução de uma resposta imune em um indivíduo, métodos para sua utilização no tratamento e processos para a produção.

[0010] O termo “vetor” refere-se a um agente (como um plasmídeo ou vírus) que contém ou transporta material genético e pode ser usado para introduzir genes exógenos em um organismo. O vetor adenoviral da presente invenção é derivado de um adenovírus símio não humano, também conhecido como “adenovírus símio”. De preferência, o vetor adenoviral símio da presente invenção é um adenovírus símio.

[0011] Cada cassete de expressão no vetor adenoviral da invenção

compreende um transgene e um promotor. Um “transgene” é uma sequência de ácidos nucleicos, heteróloga às sequências de vetores que flanqueiam o transgene, que codifica um polipeptídeo de interesse. A sequência de codificação de ácidos nucleicos está operacionalmente ligada a componentes reguladores de uma maneira que permite a transcrição, tradução e/ou expressão de transgene em uma célula hospedeira. Um “promotor” é uma sequência nucleotídica que permite a ligação da RNA polimerase e direciona a transcrição de um gene. Tipicamente, um promotor está localizado em uma região não codificante de um gene, proximal ao sítio de início da transcrição.

[0012] Nos vetores adenovirais da invenção, o primeiro cassete de expressão é inserido na região E1 do vírus e o segundo cassete de expressão é inserido em uma segunda região do vetor adenoviral.

[0013] Em um vetor adenoviral símio compreendendo dois cassetes de expressão da invenção, o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral símio e o segundo cassete de expressão é inserido em uma região do vetor adenoviral que é compatível com a replicação do vetor. Uma região do genoma do vetor adenoviral é considerada “compatível com a replicação do vetor” se a interrupção dessa região não afetar a capacidade do vetor adenoviral de se replicar.

[0014] De preferência, nos vetores adenovirais da invenção, o primeiro cassete de expressão é inserido na região E1 do vírus, e o segundo cassete de expressão é inserido na região E3, HE1 ou HE2 do vetor adenoviral. Como é bem conhecido na técnica, os genes E3 são expressos na fase inicial da transdução para preparar a célula hospedeira para replicação viral. E3 está envolvido na modulação imunológica. O termo “HE1” é usado para descrever um sítio localizado entre os códons de parada de L5 e E4. O termo “HE2” foi utilizado para definir um sítio localizado entre a extremidade da ITR e o sítio de terminação do mRNA de E4.

[0015] Por exemplo, em um vetor de adenovírus ChAd155:

- HE1 ChAd155: sítio de inserção entre os pb 34611 e 34612 da SEQ ID NO: 1.
- HE2 ChAd155: sítio de inserção entre os pb 37662 e 37663 da SEQ ID NO: 1.

[0016] Em outro exemplo, em um vetor de adenovírus ChAd83:

- HE1 ChAd83: sítio de inserção entre os pb 33535 e 33536 da SEQ ID NO: 2.
- HE2 ChAd83: sítio de inserção entre os pb 36387 e 36388 da SEQ ID NO: 2.

[0017] À medida que o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral, a região E1 nativa é deletada. A fim de aumentar a capacidade de clonagem do vetor, a região E3 nativa pode ser removida do vetor adenoviral. A região E3 nativa pode ser excluída do vetor adenoviral nas modalidades da invenção em que o segundo cassete de expressão é inserido na região E3, ou nas modalidades em que o segundo cassete de expressão não é inserido na região E3. A inserção no sítio HE1 ou HE2 não requer deleção de nenhuma sequência específica da cadeia principal do vetor.

[0018] De preferência, o segundo cassete de expressão é inserido na região HE1 ou HE2 do vetor adenoviral. Com a máxima preferência, o segundo cassete de expressão é inserido na região HE2 do vetor adenoviral. Em uma modalidade, a região E3 nativa é excluída do vetor adenoviral para aumentar a capacidade de clonagem do vetor e o segundo cassete de expressão é inserido na região HE1 ou HE2 do vetor adenoviral.

[0019] Nas modalidades da invenção, o primeiro cassete de expressão do vetor adenoviral pode compreender um CMV humano ou um promotor de CMV humano aprimorado e/ou o segundo cassete de expres-

são pode compreender um CMV humano ou um promotor de CMV humano aprimorado.

[0020] Em uma modalidade preferencial, o primeiro e o segundo cassete de expressão compreendem diferentes promotores. Por exemplo, em uma modalidade, o primeiro cassete de expressão pode compreender um promotor de CMV humano e o segundo cassete de expressão, um promotor de CMV humano aprimorado (ou vice-versa).

[0021] Em um aspecto da invenção, é fornecido um vetor adenoviral da invenção, em que o primeiro cassete de expressão é inserido na região E1 do vírus, e o segundo cassete de expressão é inserido em uma região do vetor adenoviral que é compatível com replicação de vetor, em que pelo menos um do primeiro e do segundo cassete de expressão compreende um promotor de CMV aprimorado. Em algumas modalidades, o promotor de hCMV aprimorado pode incluir uma sequência de ácidos nucleicos tendo pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98%, pelo menos cerca de 99% ou mais, de identidade de sequência com a SEQ ID NO: 6. Em algumas modalidades, o promotor compreende ou consiste em uma sequência de ácidos nucleicos da SEQ ID NO: 6.

[0022] Os vetores adenovirais da invenção são derivados de um vetor adenoviral símio, por exemplo, de chimpanzés (*Pan troglodytes*), bonobos (*Pan paniscus*), gorilas (*Gorilla gorilla*) e orangotangos (*Pongo abelii* e *Pongo pygmaeus*). Os adenovírus de chimpanzé incluem, mas não são limitados a AdY25, ChAd3, ChAd19, ChAd25.2, ChAd26, ChAd27, ChAd29, ChAd30, ChAd31, ChAd32, ChAd33, ChAd34, ChAd35, ChAd37, ChAd38, ChAd39, ChAd40, ChAd63, ChAd83, ChAd155, ChAd15, SadV41, sAd4310A, sAd4312, SAdV31, SAdV-A1337, ChAdOx1, ChAdOx2 e ChAd157. De preferência, o vetor adenoviral símio da invenção é um vetor de adenovírus ChAd83 ou

ChAd155, com mais preferência, um vetor de adenovírus ChAd155.

[0023] De preferência, o vetor adenoviral da invenção tem uma soroprevalência inferior a 30%, de preferência, menor que 10% em indivíduos humanos e, com mais preferência, nenhuma soroprevalência em indivíduos humanos.

[0024] Em uma modalidade preferencial, o vetor adenoviral símio da invenção é capaz de infectar uma célula de mamífero.

[0025] Em uma modalidade, o primeiro e o segundo cassete de expressão do vetor adenoviral da invenção compreendem transgenes do vírus sincicial respiratório (RSV). Por exemplo, em uma modalidade, um dos cassetes de expressão compreende um antígeno RSV F e o outro cassete de expressão compreende os antígenos RSV M e N. Em tais modalidades, o vetor codifica preferencialmente um antígeno RSV F0ΔTM (proteína de fusão (F) deletada das regiões transmembrana e citoplasmática) e antígenos RSV M2-1 (anti-terminação da transcrição) e N (nucleocapsídeo).

[0026] A presente invenção também fornece uma composição compreendendo um vetor adenoviral símio e um excipiente farmacêuticamente aceitável.

[0027] Além disso, a presente invenção fornece um vetor adenoviral símio ou composição compreendendo esse vetor adenoviral para uso como medicamento, vacina e/ou para terapia ou profilaxia de uma doença.

[0028] A invenção também fornece um método para induzir uma resposta imune em um indivíduo, compreendendo a administração do vetor ou composição adenoviral símia ao indivíduo.

DESCRIÇÃO DAS FIGURAS

[0029] Figura 1: Construtos adenovirais símios com um cassete de expressão único. Repetições terminais invertidas (ITR) flanqueiam as extremidades 3' e 5'; E1 é o primeiro gene 1; CMV é o promotor do

citomegalovírus; CASI é o promotor CASI, RG é um antígeno modelo, WPRE é o Elemento Regulador Pós-Transcricional de Hepatite da Marmota, $\Delta E3$ indica que o gene 3 inicial é excluído; fibra representa o gene adenoviral que codifica a proteína da fibra e E4 é o gene inicial 4.

[0030] Três vetores adenovirais símios diferentes são mostrados na Figura 1. O vetor da Figura 1(i) foi construído inserindo um cassete de expressão de transgene no local da região E3 do genoma adenoviral (“RC1”) (painel superior), o vetor da Figura 1(ii) foi formado pela inserção de um cassete de expressão de transgene na região HE1, isto é, entre os códons de parada do gene da fibra e da região E4 (“RC3”) (painel intermediário) e o vetor da Figura 1(iii) foi feito pela inserção de um cassete de expressão de transgene na região HE2, isto é, entre a extremidade da ITR e o sítio de terminação do mRNA de E4 (“RC2”) (painel inferior).

[0031] Figura 2A: Produção de ChAd155 e ChAd83 com cassete de transgene inserido nos sítios E3 e HE2 (vetores RC1 e RC2 da Figura 1) em uma linhagem celular humana primária.

[0032] Figura 2B: Produção de ChAd83 com cassete de transgene inserido em E3, HE1 e HE2 (vetores RC1, RC2 e RC3 da Figura 1) em uma linhagem celular MRC5 humana dois e sete dias após a infecção. As células foram infectadas nas multiplicidades de infecção de 250 e 1250.

[0033] Figura 3A: Número total de cópias do genoma viral do vetor RC1 e RC2 (ChAd155 e ChAd83) da Figura 1 em uma linhagem celular humana primária.

[0034] Figura 3B: Número total de cópias do genoma viral das versões RC1, RC2 e RC3 do vetor ChAd83 da Figura 1 em uma linhagem celular MRC5 humana dois e sete dias após a infecção. As células foram infectadas nas multiplicidades de infecção de 250 e 1250.

[0035] Figura 4: Número total de cópias do genoma viral dos vetores

ChAd155 RC1 e RC2 e ChAd83 RC1 e RC2 da Figura 1 em uma linhagem de células murinas (Figura 4(a), painel superior) e em uma linhagem de células de primatas não humanos (Figura 4(b), painel inferior). As células foram infectadas em multiplicidades de infecção de 50 e 250.

[0036] Figura 5: Comparação dos níveis de expressão dos vetores ChAd155 RC1 e RC2 que expressam um modelo de transgene de glicoproteína da raiva (RG) em uma linhagem de células murinas, demonstrada por western blot dois e cinco dias após a infecção (Figura 5(a), painel superior). Comparação dos níveis de expressão dos vetores ChAd155 RC1 e RC2 com os vetores ChAd83 RC1 e RC2 que expressam um modelo de transgene de glicoproteína da raiva (RG) em uma linhagem de células murinas, demonstrada por western blot dois e cinco dias após a infecção (Figura 5(b), painel inferior). As células foram infectadas em multiplicidades de infecção de 50, 250 e 1250.

[0037] Figura 5(c): Comparação dos níveis de expressão dos vetores ChAd83 RC1, RC2 e RC3 que expressam um modelo de transgene de glicoproteína da raiva (RG) em uma linhagem celular humana MRC5, demonstrada por western blot dois e sete dias após a infecção. As células foram infectadas em multiplicidades de infecção de 250 e 1250.

[0038] Figura 6: Outro construto adenoviral símio com um cassete de expressão único. Repetições terminais invertidas (ITR) flanqueiam as extremidades 3' e 5'; o CMV humano (hCMV) é o promotor do citomegalovírus; FΔTM (F0DTM) e N.M2-1 são antígenos de RSV; 2A é uma sequência de ligação de autoclivagem; 4E4 indica que o gene inicial 4 é deletado; fibra representa o gene adenoviral que codifica a proteína da fibra. No vetor da Figura 6, o cassete de expressão do transgene é inserido no local da região E1 do genoma adenoviral.

[0039] Figura 7: Um construto adenoviral símio de acordo com a invenção com um cassete de expressão duplo. Repetições terminais invertidas (ITR) flanqueiam as extremidades 3' e 5'; o CMV humano

(hCMV) é o promotor do citomegalovírus; o hCMV aprimorado é o promotor de citomegalovírus aprimorado; N-M2-1 e F Δ TM (F0DTM) são os antígenos de RSV; WPRE é o elemento regulador pós-transcricional da hepatite da marmota; Δ E3 indica que o gene 3 inicial é deletado; fibra representa o gene adenoviral que codifica a proteína da fibra; e Ad5E4orf6 em um substituto na região inicial do gene 4 (E4).

[0040] O vetor da Figura 7 foi construído inserindo um primeiro cassete de expressão de transgene no local da região E1 do genoma adenoviral e um segundo cassete de expressão de transgene na região HE2, isto é, a jusante da ITR direita.

[0041] Figura 8: Comparação dos níveis de expressão de vetores que expressam o transgene F0TM em uma linhagem celular MRC5, demonstrada por western blot 48 horas e 96 horas após a infecção em condições não redutoras. As células foram infectadas em multiplicidades de infecção de 500 e 1250.

[0042] Figura 9: Comparação dos níveis de expressão de vetores que expressam o transgene NM2-1 em uma linhagem celular MRC5, demonstrada por western blot 48 horas após a infecção em condições redutoras. As células foram infectadas em multiplicidades de infecção de 250 e 1250.

[0043] Figura 10: Comparação da imunogenicidade dos vetores ChAd155 que expressam o antígeno RSV F0TM (F Δ Tm). Os dados foram coletados 4 semanas e 8 semanas após a vacinação com uma dose de 5×10^8 partículas de vírus.

[0044] Figura 11: Comparação da imunogenicidade dos vetores ChAd155 que expressam o antígeno RSV M2. Os dados foram coletados 3 semanas após a vacinação com uma dose de 10^7 ou 10^6 partículas de vírus.

[0045] Figura 12A e 12B: Ilustram os resultados do experimento do Exemplo 9 para investigar as respostas de células T do pulmão a partir

de vetores ChAd155. A Figura 12A mostra a resposta CD4+ e a Figura 12B mostra a resposta CD8+.

[0046] Figuras 13A e 13B: Mostram os resultados do experimento do Exemplo 9 para investigar as respostas de células T periféricas dos vetores ChAd155. A Figura 13A mostra a resposta PBMC CD4+ e a Figura 13B mostra a resposta PBMC CD8+.

[0047] Figuras 14A e 14B: Mostram também os resultados do Exemplo 9. A Figura 14A mostra os títulos de Ab neutralizantes do RSV e a Figura 14B ilustra a razão do nAb do dia D90 para D0.

[0048] A Figura 15A, 15B e 15C: Mostram os resultados do experimento de imunogenicidade do Exemplo 10.

[0049] Figura 16A e 16B: Western blots obtidos utilizando a expressão em células HeLa dos vetores no Exemplo 11.

[0050] Figura 17: Ilustra os resultados do experimento de CRPV do Exemplo 12.

[0051] Figura 18: Mostra os resultados da caracterização do vetor de cassete duplo de HPV do Exemplo 13.

ANOTAÇÃO DAS SEQUÊNCIAS

SEQ ID NO: 1 - Sequência polinucleotídica que codifica ChAd155 selvagem

SEQ ID NO: 2 - Sequência polinucleotídica que codifica ChAd83 selvagem

SEQ ID NO: 3 - Sequência polinucleotídica que codifica o promotor CASI

SEQ ID NO: 4 - Sequência de polinucleotídeos que codifica ChAd155/RSV

SEQ ID NO: 5 - Sequência de aminoácidos RSV F0 Δ TM-N-M2-1

SEQ ID NO: 6 - Sequência polinucleotídica que codifica o promotor de hCMV aprimorado

SEQ ID NO: 7 - Sequência polinucleotídica que codifica o cassete hCMV

NM2 bghpolyA

SEQ ID NO: 8 - Sequência de aminoácidos (proteína) NM2

SEQ ID NO: 9 - Sequência polinucleotídica que codifica o cassete hCMV
F0 WPRE bghpolyA

SEQ ID NO: 10 - Sequência de aminoácidos (proteína) F0

SEQ ID NO: 11 - Sequência de aminoácidos de um ligante flexível

SEQ ID NO: 12 - Sequência de aminoácidos de um ligante flexível

DESCRIÇÃO DETALHADA DA INVENÇÃO

Adenovírus

[0052] Os adenovírus são vírus icosaédricos não envelopados com um genoma linear de DNA de fita dupla de aproximadamente 36 kb. Os adenovírus podem transduzir numerosos tipos de células de várias espécies de mamíferos, incluindo células que se dividem e que não se dividem, sem se integrar ao genoma da célula hospedeira. Eles têm sido amplamente utilizados para aplicações de transferência de genes devido à sua segurança comprovada, capacidade de obter transferência de genes altamente eficiente em uma variedade de tecidos alvo e grande capacidade de transgene. Atualmente, os vetores adenovirais humanos são usados em terapia gênica e vacinas, mas têm a desvantagem de uma alta prevalência mundial de imunidade preexistente, após exposição prévia a adenovírus humanos comuns.

[0053] Os adenovírus têm uma morfologia característica com um capsídeo icosaédrico compreendendo três proteínas principais, hexônio (II), base pentônio (III) e uma fibra com uma protuberância (IV), juntamente com várias outras proteínas menores, VI, VIII, IX, IIIa e IVa2. O hexônio é responsável pela maioria dos componentes estruturais do capsídeo, que consiste em 240 capsômeros de hexônio triméricos e 12 bases pentônio. O hexônio possui três tambores duplos conservados e o topo tem três torres, cada torre contendo uma alça de cada subuni-

dade que forma a maior parte do capsídeo. A base do hexônio é altamente conservada entre os sorotipos adenovirais, enquanto as alças de superfície são variáveis. O pentônio é outra proteína do capsídeo adenoviral; forma uma base pentamérica à qual a fibra se liga. A proteína da fibra trimérica se projeta da base de pentônio em cada um dos 12 vértices do capsídeo e é uma estrutura em forma de bastão. O papel principal da proteína da fibra é fixar o capsídeo viral à superfície celular através da interação da região da protuberância com um receptor celular. Variações no eixo flexível, bem como nas regiões das protuberâncias das fibras, são características dos diferentes sorotipos adenovirais.

[0054] O genoma adenoviral foi bem caracterizado. O DNA linear de dupla fita está associado à proteína VII altamente básica e a um pequeno peptídeo pX (também denominado mu). Outra proteína, V, é empacotada com este complexo de DNA-proteína e fornece uma ligação estrutural ao capsídeo via proteína VI. Existe uma conservação geral na organização geral do genoma adenoviral em relação a quadros de leitura aberta específicos posicionados de maneira similar, por exemplo, a localização dos genes E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 e L5 de cada vírus. Cada extremidade do genoma adenoviral compreende uma sequência conhecida como repetição terminal invertida (ITR), necessária para a replicação viral. A extremidade 5' do genoma adenoviral contém os elementos cis 5' necessários para o empacotamento e replicação; isto é, as sequências 5' ITR (que podem funcionar como origens da replicação) e os domínios aprimoradores nativos de empacotamento 5', que contêm sequências necessárias para empacotar genomas adenovirais lineares e elementos potencializadores para o promotor E1. A extremidade 3' do genoma adenoviral inclui elementos cis 3', incluindo as ITRs, necessárias para o empacotamento e encapsidação. O vírus também compreende uma protease codificada por vírus, necessária

para processar algumas das proteínas estruturais necessárias para produzir vírions infecciosos.

[0055] A estrutura do genoma adenoviral é descrita com base na ordem em que os genes virais são expressos após a transdução das células hospedeiras. Mais especificamente, os genes virais são referidos como genes precoces (E) ou tardios (L), de acordo com se a transcrição ocorre antes ou após o início da replicação do DNA. Na fase inicial da transdução, os genes E1A, E1B, E2A, E2B, E3 e E4 do adenovírus são expressos para preparar a célula hospedeira para replicação viral. O gene E1 é considerado uma chave principal, atua como um ativador da transcrição e está envolvido na transcrição gênica precoce e tardia. E2 está envolvido na replicação do DNA; E3 está envolvido na modulação imune e E4 regula o metabolismo do RNAm viral. Durante a fase tardia da infecção, a expressão dos genes tardios L1 a L5, que codificam os componentes estruturais das partículas virais, é ativada. Os genes tardios são transcritos do Promotor Tardio Principal (MLP) com splicing alternativo.

[0056] Os sítios HE1 e HE2 foram identificados como sítios de inserção em potencial para um transgene, pois a inserção nesses pontos específicos não interrompe as sequências de codificação ou sequências de regulação importantes de um adenovírus de chimpanzé, como um adenovírus de chimpanzé Tipo C ou E, por exemplo, ChAd155 e ChAd83. Os sítios HE1 e HE2 podem ser identificados por alinhamento de sequência em qualquer adenovírus de chimpanzé. Portanto, a clonagem de cassetes de expressão nos sítios HE1 e HE2 dos genomas ChAd não afeta o ciclo de replicação do vírus.

Replicação adenoviral

[0057] Historicamente, o desenvolvimento da vacina contra adenovírus se concentrou em vetores com defeito e não replicantes. Eles são tornados defeituosos na replicação por deleção dos genes da região E1,

que são essenciais para a replicação. Normalmente, os genes da região E3 não essenciais também são deletados para abrir espaço para transgenes exógenos. Um cassete de expressão compreendendo o transgene sob o controle de um promotor exógeno é então inserido. Esses vírus com defeito de replicação são produzidos em células que complementam a E1.

[0058] O termo “adenovírus com defeito de replicação” ou “incompetente de replicação” refere-se a um adenovírus que é incapaz de replicação porque foi projetado para compreender pelo menos uma deleção funcional (ou mutação de “perda de função”), ou seja, uma deleção ou mutação que prejudique a função de um gene sem removê-lo completamente, por exemplo, introdução de códons artificiais de parada, deleção ou mutação de sítios ativos ou domínios de interação, mutação ou deleção de uma sequência reguladora de um gene etc., ou uma remoção completa de um gene codificando um produto genético essencial à replicação viral, como um ou mais genes adenovirais selecionados de E1A, E1B, E2A, E2B, E3 e E4 (como E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, E3 ORF9, E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 e/ou E4 ORF1). Adequadamente, E1 e opcionalmente E3 e/ou E4 são deletadas. Se deletada, a região do gene deletada mencionada acima não será adequadamente considerada no alinhamento ao determinar a porcentagem de identidade em relação a outra sequência.

Vetores da Invenção

[0059] Os vetores virais baseados em adenovírus símios não humanos representam uma alternativa ao uso de vetores derivados humanos para terapia gênica e vacinas genéticas. Certos adenovírus isolados de símios não humanos estão intimamente relacionados aos adenovírus isolados de seres humanos, como demonstrado por sua propagação eficiente em células de origem humana. Como os seres humanos desenvolvem

pouca ou nenhuma imunidade aos adenovírus símios, eles prometem fornecer uma alternativa melhorada aos usos adenovirais humanos.

[0060] “Baixa soroprevalência” pode significar ter um nível de anticorpo neutralizante preexistente reduzido em comparação com o adenovírus humano 5 (Ad5). Da mesma forma ou alternativamente, “baixa soroprevalência” pode significar menos de cerca de 30% de soroprevalência, menos de cerca de 20% de soroprevalência, menos de cerca de 15% de soroprevalência, menos de cerca de 10% de soroprevalência, menos de cerca de 10% de soroprevalência, menos de cerca de 5% de soroprevalência, menos de cerca de 4% soroprevalência, menos de cerca de 3% de soroprevalência, menos de cerca de 2% de soroprevalência, menos de cerca de 1% de soroprevalência ou nenhuma soroprevalência detectável. A soroprevalência pode ser medida como a porcentagem de indivíduos com um título de neutralização clinicamente relevante (definido como um título de neutralização de 50% >200) usando métodos como descrito em Hum. Gene Ther. (2004) 15:293.

[0061] O vetor adenoviral da presente invenção é derivado de um adenovírus símio não humano, também conhecido como “adenovírus símio”. Numerosos adenovírus foram isolados de símios não humanos, como chimpanzés, bonobos, macacos rhesus, orangotangos e gorilas. Os vetores derivados desses adenovírus podem induzir fortes respostas imunes aos transgenes codificados por esses vetores. Certas vantagens de vetores baseados em adenovírus símios não humanos incluem uma relativa falta de anticorpos neutralizantes cruzados para esses adenovírus na população alvo humana, portanto, seu uso supera a imunidade preexistente aos adenovírus humanos. Por exemplo, alguns adenovírus símios não têm reatividade cruzada com anticorpos neutralizantes humanos preexistentes e a reação cruzada de certos adenovírus de chimpanzés com anticorpos neutralizantes humanos preexistentes está presente apenas em 2% da população alvo, em comparação a 35% no

caso de certos vetores de adenovírus humanos candidatos (Sci. Transl. Med. (2012) 4:1).

[0062] Os vetores adenovirais da invenção são derivados de um adenovírus símio, por exemplo, de chimpanzés (*Pan troglodytes*), bonobos (*Pan paniscus*), gorilas (*Gorilla gorilla*) e orangotangos (*Pongo abelii* e *Pongo pygmaeus*). Eles incluem adenovírus do Grupo B, Grupo C, Grupo D, Grupo E e Grupo G. Os adenovírus dos chimpanzés incluem, mas não estão limitados a AdY25, ChAd3, ChAd19, ChAd25.2, ChAd26, ChAd27, ChAd29, ChAd30, ChAd31, ChAd32, ChAd33, ChAd34, ChAd35, ChAd37, ChAd38, ChAd39, ChAd40, ChAd63, ChAd83, ChAd155, ChAd15, SadV41 e ChAd157 ChAd3, ChAd19, ChAd25.2, ChAd26, ChAd27, ChAd29, ChAd30, ChAd31, ChAd32, ChAd33, ChAd34, ChAd35, ChAd37, ChAd38, ChAd39, ChAd40, ChAd63, ChAd83, ChAd155, ChAd15, SadV41, sAd4310A, sAd4312, SAdV31, SAdV-A1337, ChAdOx1, ChAdOx2 e ChAd157. Alternativamente, os vetores adenovirais podem ser derivados de adenovírus símios não humanos isolados de bonobos, como PanAd1, PanAd2, PanAd3, PanAd3, Pan 5, Pan 6, Pan 7 (também conhecido como C7) e Pan 9. Os vetores podem incluir, no todo ou em parte, um nucleotídeo que codifica a fibra, pentônio ou hexônio de um adenovírus não humano.

[0063] Em uma modalidade dos vetores adenovirais da invenção, o vetor adenoviral tem uma soroprevalência inferior a 30%, inferior a 20%, inferior a 10% ou inferior a 5% em indivíduos humanos, de preferência nenhuma soroprevalência em indivíduos humanos e com mais preferência, não há soroprevalência em indivíduos humanos que não haviam estado em contato com um vetor adenoviral de chimpanzé.

[0064] Em modalidades dos vetores adenovirais da invenção, o DNA adenoviral é capaz de entrar em uma célula alvo de mamífero, isto é, é infeccioso. Um vetor adenoviral recombinante infeccioso da invenção pode ser utilizado como uma vacina profilática ou terapêutica e para

terapia genética. Dessa forma, em uma modalidade, o vetor adenoviral recombinante compreende uma molécula endógena para entrega em uma célula alvo. A célula alvo é uma célula de mamífero, por exemplo, uma célula bovina, uma célula canina, uma célula caprina, uma célula de cervo, uma célula de chimpanzé, uma célula de quiróptero, uma célula equina, uma célula felina, uma célula humana, uma célula de lobo, uma célula ovina, célula suína, célula de roedor, célula de urso ou célula vulpina. A molécula endógena para entrega em uma célula alvo é um cassete de expressão.

[0065] Em uma modalidade da invenção, o vetor compreende uma região ITR esquerda, uma região E1 deletada, depois uma região E3 deletada e, opcionalmente, elementos intensificadores adicionais; estes são seguidos por uma região de fibra, uma região E4 e uma ITR direita. A tradução ocorre nas direções direita e esquerda. Nesta modalidade, o primeiro cassete de expressão é inserido na região E1 deletada e o segundo cassete de expressão é inserido na região E3 deletada. Em uma outra modalidade, os promotores dos dois cassetes de expressão são promotores de CMV. Ainda em uma modalidade adicional, o elemento intensificador é o Elemento Regulador Pós-Traducional da Hepatite B (HPRE) ou o Elemento Regulador Pós-Traducional de Hepatite B da Marmota (WPRE).

[0066] Em uma modalidade da invenção, o vetor compreende regiões ITR esquerda e direita; uma região E1 deletada; pelo menos uma região E3 parcialmente deletada; uma região de fibra; uma região E4; dois cassetes de expressão, cada um compreendendo: um promotor e pelo menos um antígeno de interesse e, opcionalmente, um ou mais elementos intensificadores. O primeiro cassete de expressão é inserido na região E1 deletada e o segundo cassete de expressão é inserido no sítio HE1, ou seja, entre os códons de parada do gene da fibra e uma região E4 (“o sítio HE1”). O sítio de inserção ChAd155 HE1 está entre

os pb 34611 e 34612 da sequência ChAd155 selvagem. O sítio de inserção ChAd83 HE1 está entre os pb 33535 e 33536 da sequência ChAd83 selvagem. A tradução ocorre nas direções direita e esquerda. Em uma outra modalidade, os promotores são promotores de CMV. Em uma modalidade preferencial, um promotor é um promotor de CMV e o outro é um promotor de eCMV. Ainda em uma modalidade adicional, o elemento intensificador é HPRE ou WPRE.

[0067] Em uma outra modalidade, o vetor compreende regiões ITR esquerda e direita; uma região E1 deletada; pelo menos uma região E3 parcialmente deletada; uma região de fibra; uma região E4; dois cassetes de expressão, cada um compreendendo: um promotor, pelo menos um antígeno de interesse e, opcionalmente, um ou mais elementos intensificadores. O primeiro cassete de expressão é inserido na região E1 deletada e o segundo cassete de expressão é inserido no sítio HE2, ou seja, entre a extremidade da ITR esquerda e o sítio de terminação do mRNA de E4 (“o sítio HE2”). O sítio de inserção ChAd155 HE2 está entre os pb 37662 e 37663 da sequência ChAd155 selvagem. O sítio de inserção ChAd83 HE2 está entre os pb 36387 e 36388 da sequência ChAd83 selvagem. A tradução ocorre nas direções direita e esquerda. Em uma outra modalidade, os promotores são promotores de CMV. Em uma modalidade preferencial, um promotor é um promotor de CMV e o outro é um promotor de eCMV. Ainda em uma modalidade adicional, o elemento intensificador é HPRE ou WPRE (o elemento intensificador aumenta a expressão do transgene).

[0068] Os sítios HE1 e HE2 foram identificados como sítios de inserção para um transgene, uma vez que a inserção nesses pontos específicos não interrompe as sequências de codificação ou de regulação de ChAd155 e ChAd83. Portanto, a inserção de cassetes de expressão nos sítios HE1 ou HE2 do genoma ChAd não afeta o ciclo de replicação viral.

[0069] Em uma modalidade da invenção, o vetor é um derivado funcional ou imunogênico de um vetor adenoviral. Por “derivado de um vetor adenoviral” entende-se uma versão modificada do vetor, por exemplo, um ou mais nucleotídeos do vetor são deletados, inseridos, modificados ou substituídos.

Elementos de Regulação

[0070] Elementos reguladores, isto é, sequências de controle de expressão, incluem sequências apropriadas de iniciação, terminação, promotoras e intensificadoras de transcrição; sinais eficientes de processamento de RNA, como sinais de junção e poliadenilação (poli A), incluindo beta-globina poliA de coelho; sistemas reguláveis por tetraciclina, microRNAs, elementos reguladores pós-transcricionais (por exemplo, WPRE, elemento regulador pós-transcricional do vírus de hepatite da marmota); sequências que estabilizam o mRNA citoplasmático; sequências que aumentam a eficiência da tradução (por exemplo, sequência consenso de Kozak); sequências que melhoram a estabilidade das proteínas; e quando desejado, sequências que aumentam a secreção de um produto codificado.

[0071] Um “promotor” é uma sequência nucleotídica que permite a ligação da RNA polimerase e direciona a transcrição de um gene. Tipicamente, um promotor está localizado em uma região não codificante de um gene, proximal ao sítio de início da transcrição. Os elementos de sequência dentro dos promotores que funcionam no início da transcrição são frequentemente caracterizados por sequências de nucleotídeos consenso. Exemplos de promotores incluem, mas não estão limitados a promotores de bactérias, leveduras, plantas, vírus e mamíferos, incluindo símios e humanos. Um grande número de sequências de controle de expressão, incluindo promotores internos, nativos, constitutivos, indutíveis e/ou específicos para tecidos, são conhecidos na técnica e podem ser utilizados.

[0072] Os promotores da invenção serão tipicamente promotores heterólogos. Os promotores da invenção podem ser constitutivos.

[0073] Exemplos de promotores incluem, mas não estão limitados a promotores de bactérias, leveduras, plantas, vírus e mamíferos (incluindo seres humanos).

[0074] Exemplos de promotores incluem, sem limitação, o promotor TBG, o promotor LTR do vírus retroviral do sarcoma de Rous (opcionalmente com o intensificador), o promotor do citomegalovírus (CMV) (opcionalmente com o intensificador do CMV, ver, por exemplo, Boshart *et al*, Cell, 41:521-530 (1985)), o promotor CASI, o promotor SV40, o promotor da di-hidrofolato redutase, o promotor da β -actina, o promotor da fosfoglicerol quinase (PGK) e o promotor EF1a (Invitrogen).

[0075] Promotores adequados incluem o promotor de citomegalovírus (CMV) e o promotor CASI. O promotor CMV é forte e onipresente ativo. Tem a capacidade de conduzir altos níveis de expressão de transgene em muitos tipos de tecidos e é bem conhecido na técnica. O promotor de CMV pode ser utilizado em vetores da invenção, com ou sem um intensificador de CMV.

[0076] O promotor CASI é um promotor sintético descrito como uma combinação do intensificador de CMV, o promotor de beta-actina de frango e um doador de junção e um aceitor de junção que flanqueiam o intensificador de ubiquitina (UBC) (US 8865881).

[0077] Em algumas modalidades, o promotor CASI pode incluir uma sequência de ácidos nucleicos com pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98%, pelo menos cerca de 99% ou mais, de identidade de sequência com a SEQ ID NO: 3. Em algumas modalidades, o promotor compreende ou consiste em uma sequência de ácidos nucleicos da SEQ ID NO: 3.

[0078] Em algumas modalidades, o promotor de hCMV aprimorado

pode incluir uma sequência de ácidos nucleicos tendo pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98%, pelo menos cerca de 99% ou mais, de identidade de sequência com a SEQ ID NO: 6. Em algumas modalidades, o promotor compreende ou consiste em uma sequência de ácidos nucleicos da SEQ ID NO: 6.

[0079] Opcionalmente, vetores transportando transgenes que codificam produtos terapeuticamente úteis ou imunogênicos também podem incluir marcadores selecionáveis ou genes repórter. O gene repórter pode ser escolhido dentre aqueles conhecidos na técnica. Genes repórter adequados incluem, mas não estão limitados a proteína fluorescente verde aprimorada, proteína fluorescente vermelha, luciferase e fosfatase alcalina embrionária segregada (seAP), que podem incluir sequências que codificam resistência à geneticina, higromicina ou purimicina, entre outras. Esses genes repórter ou marcadores selecionáveis (que podem ou não estar localizados fora do genoma viral para serem empacotados em uma partícula viral) podem ser usados para sinalizar a presença dos plasmídeos nas células bacterianas, como resistência à ampicilina. Outros componentes do vetor podem incluir uma origem de replicação.

[0080] Um “elemento regulador pós-transcricional”, como utilizado neste pedido, é uma sequência de DNA que, quando transcrita, melhora a expressão do(s) transgene(s) ou fragmentos dos mesmos que são entregues pelos vetores virais da invenção. Os elementos reguladores pós-transcricionais incluem, entre outros, o Elemento Regulador Pós-Transcricional do Vírus da Hepatite B (HPRE) e o Elemento Regulador Pós-transcricional da Hepatite de Marmota (WPRE). O WPRE é um elemento tripartido cis-atuante que demonstrou melhorar a expressão do transgene acionado por certos promotores, mas não todos.

[0081] Nas modalidades da invenção, um vetor ChAd155 pode

compreender um ou mais dentre um promotor, um intensificador e um gene repórter. Por exemplo, os vetores da invenção podem compreender hCMV-SeAP, ChAd155-CASI-seAP e ChAd155-hCMV-seAP, opcionalmente com um controle transcricional liga/desliga de tetraciclina e ChAd155-CMV-hFerL-chEF1-seAP com um controle transcricional liga/desliga de tetraciclina.

[0082] Nas modalidades da invenção, um vetor ChAd83 pode compreender um ou mais dentre um promotor, um intensificador e um gene repórter. Por exemplo, os vetores da invenção podem compreender hCMV-SeAP aprimorado por ChAd155, hCMV-SeAP aprimorado por ChAd83, ChAd155-CASI-seAP e ChAd83-hCMV-seAP, opcionalmente com um controle transcricional liga/desliga de tetraciclina e ChAd83-CMV-hFerL-chEF1-seAP com um controle transcricional liga/desliga de tetraciclina.

[0083] Os vetores da invenção são gerados usando técnicas fornecidas neste pedido, em conjunto com técnicas conhecidas pelos versados na técnica. Tais técnicas incluem técnicas convencionais de clonagem de cDNA, como as descritas nos textos, uso de sequências oligonucleotídicas sobrepostas dos genomas de adenovírus, reação de polimerase em cadeia e qualquer método adequado que forneça a sequência nucleotídica desejada.

Transgenes

[0084] Um “transgene” é uma sequência de ácidos nucleicos, heteróloga às sequências de vetores que flanqueiam o transgene, que codifica um polipeptídeo de interesse. A sequência de codificação de ácidos nucleicos está operacionalmente ligada a componentes reguladores de uma maneira que permite a transcrição, tradução e/ou expressão de transgene em uma célula hospedeira. Nas modalidades da invenção, os vetores expressam transgenes em nível terapêutico ou profilático. Um

“derivado funcional” de um polipeptídeo transgênico é uma versão modificada de um polipeptídeo, por exemplo, em que um ou mais aminoácidos são deletados, inseridos, modificados ou substituídos.

[0085] O transgene pode ser utilizado para profilaxia ou tratamento, por exemplo, como uma vacina para induzir uma resposta imune, para corrigir deficiências genéticas, corrigindo ou substituindo um gene defeituoso ou ausente, ou como terapia para o câncer. Como utilizado neste pedido, a indução de uma resposta imune refere-se à capacidade de uma proteína para induzir uma célula T e/ou uma resposta imune de anticorpo humoral à proteína.

[0086] A resposta imune provocada pelo transgene pode ser uma resposta de célula B específica ao antígeno, que produz anticorpos neutralizantes. A resposta imune provocada pode ser uma resposta de célula T específica para o antígeno, que pode ser uma resposta sistêmica e/ou local. A resposta das células T específicas para o antígeno pode compreender uma resposta das células T CD4+, como uma resposta envolvendo células T CD4+ que expressam citocinas, por exemplo, interferon gama (IFN gama), fator de necrose tumoral alfa (TNF alfa) e/ou interleucina 2 (IL2). Alternativamente ou adicionalmente, a resposta da célula T específica para o antígeno compreende uma resposta da célula T CD8+, como uma resposta envolvendo células T CD8+ que expressam citocinas, por exemplo, IFN gama, TNF alfa, TNF alfa e/ou IL2.

[0087] A composição da sequência de transgene dependerá do uso ao qual o vetor resultante será aplicado. Em uma modalidade, o transgene é uma sequência que codifica um produto que é útil em biologia e medicina, como um transgene profilático, um transgene terapêutico ou um transgene imunogênico, por exemplo, proteína ou RNA. Os transgenes de proteínas incluem antígenos. Os transgenes antigênicos da invenção induzem uma resposta imunogênica para um organismo causador de doença.

[0088] Os transgenes da invenção incluem, mas não estão limitados a antígenos do vírus da raiva, por exemplo, glicoproteína da raiva (RG), antígenos do vírus sincicial respiratório (RSV), antígenos do vírus da imunodeficiência humana (HIV) ou fragmentos dos mesmos.

[0089] Como resultado da redundância no código genético, um polipeptídeo pode ser codificado por uma variedade de diferentes sequências de ácidos nucleicos. A codificação é enviesada para usar alguns códons sinônimos, ou seja, códons que codificam o mesmo aminoácido, mais do que outros. Por “códon otimizado”, entende-se que modificações na composição do códon de um ácido nucleico recombinante são feitas sem alterar a sequência de aminoácidos. A otimização de códons tem sido usada para melhorar a expressão de mRNA em diferentes organismos usando frequências de uso de códons específicas para cada organismo.

[0090] Além de e independentemente do viés do códon, alguns pares de códons sinônimos são usados com mais frequência do que outros. Esse viés do par de códons significa que alguns pares de códons estão super-representados e outros estão sub-representados. A desotimização do par de códons tem sido utilizada para reduzir a virulência viral. Por exemplo, foi relatado que os poliovírus modificados para conter pares de códons sub-representados demonstraram menor eficiência de tradução e foram atenuados em comparação com o poliovírus selvagem (Science (2008) 320:1784). A engenharia de um vírus sintético atenuado por desotimização de pares de códons pode produzir vírus que codificam as mesmas sequências de aminoácidos selvagens, mas usam diferentes arranjos de pares de códons sinônimos. Os vírus atenuados pela desotimização do par de códons geraram até 1000 vezes menos placas em comparação com o tipo selvagem, produziram menos partículas virais e exigiram cerca de 100 vezes mais partículas virais para formar uma placa.

[0091] Em contraste, os poliovírus modificados para conter pares de códons que são super-representados no genoma humano agiram de maneira semelhante ao RNA selvagem e geraram placas idênticas em tamanho ao RNA selvagem (Coleman *et al.* (2008) Science 320:1784). Isso ocorreu apesar do fato de que o vírus com pares de códons super-representados continha um número semelhante de mutações que o vírus com pares de códons sub-representados e demonstrou tradução aprimorada em comparação com o tipo selvagem. Esta observação sugere que se espera que os construtos otimizados do par de códons atuem de maneira semelhante às suas contrapartes não otimizadas do par de códons e que não forneça uma vantagem funcional. Sem desejar ser limitado pela teoria, isso pode ocorrer porque a evolução natural otimizou o emparelhamento de códons.

[0092] Um construto da invenção pode compreender uma sequência de ácidos nucleicos otimizada por códon. Alternativamente ou adicionalmente, um vetor da invenção compreende uma sequência otimizada por códon de um transgene ou um derivado imunogênico ou fragmento do mesmo. Um construto da invenção pode compreender uma sequência de ácidos nucleicos otimizada para o par de códons. Alternativamente ou adicionalmente, um vetor da invenção compreende ou consiste em uma sequência otimizada do par de códons de um transgene ou um derivado imunogênico ou fragmento do mesmo.

Transgenes do Vírus Sincicial Respiratório (RSV)

[0093] Em uma modalidade, a presente invenção fornece o uso de um vetor adenoviral derivado de símio recombinante compreendendo dois cassetes de expressão, em que cada cassete de expressão compreende um transgene imunogênico derivado do vírus sincicial respiratório humano (RSV), no tratamento ou profilaxia da infecção por RSV. Em uma modalidade, o vetor adenoviral derivado de símio recombinante

da presente invenção compreende um antígeno RSV F em um dos cassetes de expressão e outro antígeno viral de RSV no outro cassete de expressão. Antígenos adequados são discutidos mais adiante. Em uma modalidade, o vetor adenoviral derivado de símio recombinante compreende antígenos RSV M e N no segundo cassete de expressão. Em tais modalidades, o vetor codifica preferencialmente um antígeno RSV F0ΔTM (proteína de fusão (F) deletada das regiões transmembrana e citoplasmática) e antígenos RSV M2-1 (anti-terminação da transcrição) e N (nucleocapsídeo).

[0094] A infecção por RSV não confere imunidade protetora total. A infecção na infância é seguida por reinfecções sintomáticas do RSV, que continuam por toda a vida adulta. Essas reinfecções geralmente não são diagnosticadas porque geralmente se apresentam como infecções agudas comuns do trato respiratório superior. Em pessoas mais vulneráveis (por exemplo, adultos ou idosos imunocomprometidos), as reinfecções também podem levar a doenças graves. Ambos os braços do sistema imunológico (imunidade humoral e celular) estão envolvidos na proteção contra doenças graves [Guvenel AK, Chiu C and Openshaw PJ. Current concepts and progress in RSV vaccine development. *Expert Rev Vaccines*. 2014; 13(3): 333-44.].

[0095] A resposta imune humoral é capaz de neutralizar o vírus e inibir a replicação viral, desempenhando, assim, um papel importante na proteção contra infecção respiratória inferior por RSV e doenças graves [Piedra PA, Jewell AM, Cron SG, *et al.*, Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies. *Vaccine*. 2003; 21(24): 3479-82.]. Foi demonstrado que a imunização passiva, na forma de anticorpos monoclonais neutralizadores de RSV (Synagis) de imunoglobulina G (IgG) administrados profilática-

mente, tem demonstrado prevenir a doença por RSV em alguma extensão em bebês e recém-nascidos prematuros com displasia broncopulmonar ou doença cardiopulmonar subjacente [Cardenas S, Auais A and Piedimonte G. Palivizumab in the prophylaxis of respiratory syncytial virus infection. *Expert Rev Anti Infect Ther.* 2005; 3(5): 719-26].

[0096] As células T também estão envolvidas no controle da doença por RSV. Infecções letais por RSV foram descritas em pacientes com baixa contagem de células T CD8, como no caso de imunodeficiência combinada grave, receptores de medula óssea e transplante de pulmão [Hertz, 1989]. A histopatologia de casos fatais de infecção por RSV de recém-nascidos mostra que existe uma escassez relativa de células T CD8 no infiltrado pulmonar [Welliver TP, Garofalo RP, Hosakote Y, *et al.*, Severe human lower respiratory tract illness caused by respiratory syncytial virus and influenza virus is characterized by the absence of pulmonary cytotoxic lymphocyte responses. *J Infect Dis.* 2007. 195(8): 1126-36.]. Além disso, a presença de células T CD8 produtoras de interferon gama (IFN- γ) tem sido associada a respostas Th2 diminuídas e eosinofilia reduzida em modelos animais de RSV [Castilow EM and Varga SM. Overcoming T cell-mediated immunopathology to achieve safe RSV vaccination. *Future Virol.* 2008; 3(5): 445-454; Stevens WW, Sun J, Castillo JP, *et al.*, Pulmonary eosinophilia is attenuated by early responding CD8(+) memory T cells in a murine model of RSV vaccine-enhanced disease. *Viral Immunol.* 2009; 22(4): 243-51].

[0097] Antígenos adequados de RSV que são úteis como imunógenos para imunizar um ser humano ou um animal não humano podem ser selecionados dentre: a proteína de fusão (F), a proteína de ligação (G), a proteína de matriz (M2) e a nucleoproteína (N). O termo “proteína F” ou “proteína de fusão” ou “polipeptídeo de proteína F” ou “polipeptídeo de proteína de fusão” refere-se a um polipeptídeo ou proteína que possui a totalidade ou parte de uma sequência de aminoácidos de um

polipeptídeo da proteína de fusão de RSV. Da mesma forma, o termo “proteína G” ou “polipeptídeo de proteína G” refere-se a um polipeptídeo ou proteína que possui a totalidade ou parte de uma sequência de aminoácidos de um polipeptídeo da proteína de ligação de RSV. O termo “proteína M” ou “proteína da matriz” ou “polipeptídeo da proteína M” refere-se a um polipeptídeo ou proteína que possui a totalidade ou parte de uma sequência de aminoácidos de uma proteína da Matriz RSV e pode incluir um ou ambos os M2-1 (que pode ser escrito neste pedido como produtos dos genes M2.1) e M2-2. Da mesma forma, o termo “proteína N” ou “proteína de nucleocapsídeo” ou “polipeptídeo da proteína N” refere-se a um polipeptídeo ou proteína que possui a totalidade ou parte de uma sequência de aminoácidos de uma nucleoproteína de RSV.

[0098] Dois grupos de cepas humanas de RSV foram descritos, os grupos A e B, baseados principalmente em diferenças na antigenicidade da glicoproteína G. Numerosas cepas de RSV foram isoladas até o momento, qualquer uma das quais é adequada no contexto dos antígenos das combinações imunogênicas divulgadas neste pedido. Cepas exemplares indicadas pelo número de acesso GenBank e/ou EMBL podem ser encontradas no pedido de publicação publicado US 2010/0203071 (WO2008114149), que é incorporado neste pedido por referência com o objetivo de divulgar as sequências de ácidos nucleicos e polipeptídeos de proteínas RSV F e G adequadas para uso na presente invenção. Em uma modalidade, a proteína RSV F pode ser um ectodomínio de uma proteína RSV F (F0ΔTM).

[0099] Exemplos de ácidos nucleicos e sequências de proteínas das proteínas M e N podem ser encontrados, por exemplo, no pedido publicado US 2014/0141042 (WO2012/089833), que é incorporado neste pedido com o objetivo de divulgar as sequências de ácidos nucleicos e polipeptídeos das proteínas RSV M e N adequadas para uso na

presente invenção.

[00100] Adequadamente, para utilização na presente invenção, os ácidos nucleicos do transgene codificam um antígeno RSV F e antígenos RSV M e N. Mais especificamente, os ácidos nucleicos codificam uma proteína do antígeno RSV F0ΔTM (proteína de fusão (F) deletada das regiões transmembrana e citoplasmática) e antígenos RSV M2-1 (anti-terminação da transcrição) e N (nucleocapsídeo).

Proteína de fusão (F) deletada das regiões transmembrana e citoplasmática (F0ΔTM)

[00101] A proteína RSV F é um antígeno de superfície principal e medeia a fusão viral para as células alvo. A proteína F é um antígeno que é altamente conservado entre subgrupos e cepas de RSV. A proteína F é um alvo para anticorpos neutralizantes, incluindo o anticorpo monoclonal profilático neutralizante para RSV, Synagis. A deleção da região transmembrana e da cauda citoplasmática permite a secreção da proteína F0ΔTM. Anticorpos neutralizantes, incluindo Synagis, que reconhecem essa forma solúvel da proteína F, inibem a infectividade por RSV *in vitro* [Magro M, Andreu D, Gómez-Puertas P, *et al.*, Neutralization of human respiratory syncytial virus infectivity by antibodies and low-molecular-weight compounds targeted against the fusion glycoprotein. *J Virol.* 2010; 84(16): 7970-82].

Proteína nucleocapsídica (N)

[00102] A proteína N é um antígeno interno (não exposto), altamente conservado entre as cepas de RSV e conhecida por ser uma fonte de muitos epítomos de células T. A proteína N é essencial para a replicação e transcrição do genoma do RSV. A função principal da proteína N é encapsular o genoma do vírus para fins de transcrição, replicação e empacotamento de RNA, protegendo-o das ribonucleases.

Proteína de transcrição anti-terminação (M2-1)

[00103] A proteína M2-1 é um fator de anti-terminação de transcrição

importante para a síntese eficiente de RNAs mensageiros completos (mRNAs), bem como para a síntese de mRNAs de leitura policistrônica, que são característicos de vírus de RNA de cadeia negativa não segmentada. M2-1 é um antígeno interno (não exposto), que é altamente conservado entre as cepas de RSV e conhecido por ser uma fonte de muitos epítomos de células T.

Proteína de Fusão N-M2-1

[00104] Um polinucleotídeo que codifica um ligante é posicionado entre o polinucleotídeo que codifica um antígeno RSV N, ou fragmento do mesmo, e o polinucleotídeo que codifica um antígeno RSV M2.1, ou fragmento do mesmo. Dessa forma, em certos exemplos preferenciais, um cassete de expressão contém um transgene que codifica uma proteína viral RSV N-ligante-M2.1 fundida. É preferencial que o ligante seja um ligante flexível, de preferência, um ligante flexível compreendendo uma sequência de aminoácidos de acordo com a SEQ ID NO: 11 (Gly-Gly-Gly-Ser-Gly-Gly-Gly) ou SEQ ID NO: 12 (Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly).

Transgenes do papiloma (PV)

[00105] Em uma modalidade, a presente invenção fornece o uso de um vetor adenoviral derivado de símio recombinante compreendendo dois cassetes de expressão, em que cada cassete de expressão compreende um transgene imunogênico derivado de um vírus do papiloma (PV), no tratamento ou profilaxia de uma doença induzida pelo vírus do papiloma. Adequadamente, o vetor adenoviral derivado de símio recombinante da presente invenção compreende um antígeno E1 do vírus do papiloma modificado em um dos cassetes de expressão e um antígeno E2 do vírus do papiloma modificado no outro cassete de expressão.

[00106] O papilomavírus humano (HPV) é um pequeno vírus de DNA que infecta a pele e/ou mucosa cutânea e causa várias doenças, inclu-

indo neoplasia cervical, câncer cervical e outros tipos de câncer anogenital. Sabe-se que existem mais de 40 tipos de HPV que infectam o trato anogenital de seres humanos e cerca de 15 genótipos de alto risco estão associados causalmente a cânceres cervicais humanos. A maioria das infecções por HPV do epitélio cervical é subclínica e auto-resolvida em um período de dois anos. No entanto, a infecção persistente com tipos de HPV de alto risco pode causar lesões e progredir para câncer invasivo.

[00107] Antígenos adequados para HPV que são úteis como imunógenos são descritos no documento WO2018060288 e incluem, em particular, as proteínas E1 e E2 do HPV.

Transgenes da Raiva (RG)

[00108] O *Lyssavirus* é um vírus de RNA de fita simples da família *Rhabdoviridae*. Os membros do gênero *Lyssavirus* causam raiva e têm a maior taxa de mortalidade de todos os patógenos virais humanos conhecidos. A raiva é transmitida pela saliva de mamíferos infectados. Um vírus neurotrópico, entra no sistema nervoso de seu hospedeiro, causando uma encefalomielite quase invariavelmente fatal. Atualmente, existem cerca de 60.000 mortes por raiva em todo o mundo anualmente, principalmente causadas por mordidas de cães em países em desenvolvimento na Ásia e África e por animais selvagens e morcegos na América do Norte.

[00109] A raiva se apresenta de forma furiosa ou paralítica. O período de incubação varia entre cerca de cinco dias e vários anos, mas é tipicamente entre cerca de 20 e 90 dias. A doença clínica geralmente começa com queixas prodrômicas de mal-estar, anorexia, fadiga, dor de cabeça e febre, seguidas de dor ou paratesia no local da exposição. Ansiedade, agitação ou irritabilidade podem ser proeminentes durante esse período, seguidas de hiperatividade, desorientação, convulsões, hidrofobia, hipersalivação e, eventualmente, paralisia, coma e morte.

[00110] Os antígenos da raiva podem ser derivados da glicoproteína viral da raiva (RG). Por exemplo, a glicoproteína da raiva pode ser usada como um antígeno modelo.

Entrega de Vetores Adenovirais

[00111] Em algumas modalidades, o vetor adenoviral recombinante da invenção é administrado a um indivíduo por administração epicutânea, administração intradérmica, injeção intramuscular, injeção intraperitoneal, injeção intravenosa, administração nasal, administração nasal, administração oral, administração retal, injeção subcutânea, administração transdérmica ou administração intravaginal.

[00112] Em uma modalidade da invenção, os vetores podem ser administrados por via intramuscular (IM), isto é, injeção diretamente no músculo. Os músculos são bem vascularizados e a captação é tipicamente rápida.

Adjuvantes

[00113] As abordagens para estabelecer uma imunidade forte e duradoura a patógenos específicos incluem a adição de adjuvantes às vacinas. Por “adjuvante” entende-se um agente que aumenta, estimula, ativa, potencializa ou modula a resposta imune a um ingrediente ativo da composição. O efeito adjuvante pode ocorrer no nível celular ou humoral, ou em ambos. Os adjuvantes estimulam a resposta do sistema imunológico ao antígeno real, mas não têm efeito imunológico por si só. Alternativamente ou adicionalmente, as composições adjuvantes da invenção podem incluir um ou mais imunoestimulantes. Por “imunoestimulante”, entende-se um agente que induz um aumento geral e temporário na resposta imune de um indivíduo, seja administrado com o antígeno ou separadamente.

[00114] Uma composição da invenção pode ser administrada com ou sem um adjuvante. Alternativamente, ou adicionalmente, a composição pode compreender, ou ser administrada em conjunto com um ou mais

adjuvantes (por exemplo, adjuvantes de vacina), em particular a composição compreende uma quantidade imunologicamente eficaz de um vetor da invenção que codifica um transgene.

Métodos de uso/usos

[00115] São fornecidos métodos para induzir uma resposta imune contra uma doença causada por um patógeno em um indivíduo em necessidade dos mesmos, compreendendo uma etapa de administração de uma quantidade imunologicamente eficaz de um construto ou composição, como divulgado neste pedido. Em algumas modalidades, é fornecido o uso de construtos ou composições divulgados neste pedido para induzir uma resposta imune a um antígeno transgênico em um indivíduo em necessidade dos mesmos. Os vetores da invenção podem ser aplicados para a profilaxia, tratamento ou melhora de doenças devido à infecção.

[00116] Os métodos da invenção incluem o uso de um vetor da invenção em medicina. Eles incluem o uso de um vetor da invenção para o tratamento de uma doença causada por um patógeno. Um vetor da invenção pode ser utilizado na fabricação de um medicamento para o tratamento de uma doença causada por um patógeno. Um vetor da invenção pode ser utilizado na fabricação de um medicamento para a prevenção ou tratamento de uma doença, por exemplo, uma doença causada pelo vírus sincicial respiratório (RSV).

[00117] A imunização eficaz com vetores adenovirais depende da capacidade imunomoduladora intrínseca da cadeia principal do vetor adenoviral. Os adenovírus imunologicamente menos potentes induzem menos expressão de antígeno. A imunização eficaz também depende da capacidade do promotor de conduzir a expressão forte e sustentada do transgene. Por exemplo, os vetores adenovirais conduzidos pelo promotor do citomegalovírus imediato-precocce (CMV-IE) não sustentam a expressão de transgene em longo prazo porque induzem citocinas que

amortecem a expressão.

[00118] Por “indivíduo” entende-se um vertebrado, como um mamífero, por exemplo, um ser humano ou um mamífero veterinário. Em algumas modalidades, o indivíduo é um ser humano.

Geral

[00119] Os vetores da invenção são gerados usando técnicas e sequências fornecidas neste pedido, em conjunto com técnicas conhecidas pelos versados na técnica. Tais técnicas incluem técnicas convencionais de clonagem de cDNA, como as descritas nos textos, uso de sequências oligonucleotídicas sobrepostas dos genomas de adenovírus, reação de polimerase em cadeia e qualquer método adequado que forneça a sequência nucleotídica desejada.

[00120] A menos que definido de outra forma, todos os termos técnicos e científicos usados neste pedido têm o mesmo significado que comumente entendido por um versado na técnica à qual pertence esta divulgação. Os termos singulares “um”, “uma”, “o” e “a” incluem referências no plural a menos que o contexto indique claramente de outra maneira. Da mesma forma, a palavra “ou” pretende incluir “e”, a menos que o contexto indique claramente de outra maneira. O termo “pluralidade” refere-se a dois ou mais. Além disso, as limitações numéricas dadas em relação às concentrações ou níveis de uma substância, como as concentrações dos componentes da solução ou razões das mesmas, e as condições da reação, como temperaturas, pressões e tempos de ciclo, devem ser aproximadas. O termo “aproximadamente” utilizado neste pedido pretende significar a quantidade de $\pm 10\%$.

[00121] A presente invenção será agora descrita mais detalhadamente por meio dos seguintes exemplos não limitadores.

EXEMPLOS

Exemplo 1: Construção de adenovírus de chimpanzé com um cassete de expressão única

[00122] Os adenovírus selvagens de chimpanzé tipo 155 (ChAd155) (WO 2016/198621) e tipo 83 (ChAd83) (WO 2010/086189) foram isolados de chimpanzés saudáveis usando procedimentos padrão e foram construídos conforme descrito em Sci Transl Med (2012) 4:1 e WO 2010/086189.

[00123] No Exemplo 1, os vetores ChAd155 e ChAd 83 foram construídos inserindo um cassete único de expressão de transgene. Os componentes do cassete de expressão usavam o promotor clássico de CMV humano ou o promotor CASI, glicoproteína da raiva como antígeno modelo e, opcionalmente, um intensificador do WPRE. Três sítios de inserção diferentes foram testados para o cassete do transgene:

- (i) substituição da região E3 pelo cassete do transgene,
- (ii) inserção do cassete do transgene entre a fibra e a região E4 (sítio HE1), e
- (iii) inserção do cassete do transgene a jusante da ITR direita (sítio HE2).

[00124] Essa numeração desses sítios de inserção corresponde às ilustrações da Figura 1 onde:

- (i) o painel superior ilustra o vetor RC1, no qual um cassete de transgene substituiu a região E3,
- (ii) o painel intermediário ilustra o vetor RC3, no qual um cassete do transgene é inserido entre os códons de parada do gene da fibra e da região E4 (sítio HE1), e
- (iii) o painel inferior ilustra o vetor RC2, no qual um cassete do transgene é inserido a jusante da ITR direita (sítio HE2).

[00125] Nos vetores mostrados no Exemplo 1, a região E1 permanece intacta em todas as configurações.

[00126] O transgene foi inserido por técnicas de recombinação homóloga nas seguintes posições da SEQ ID NO: 1, e da SEQ ID NO: 2: HE1 ChAd155: sítio de inserção entre os pb 34611 e 34612 da SEQ ID

NO: 1;

HE2 ChAd155: sítio de inserção entre os pb 37662 e 37663 da SEQ ID

NO: 1;

HE1 ChAd83: sítio de inserção entre os pb 33535 e 33536 da SEQ ID

NO: 2;

HE2 ChAd83: sítio de inserção entre os pb 36387 e 36388 da SEQ ID

NO: 2.

[00127] Quando o cassete do transgene foi inserido no sítio HE1, ChAd155 falhou ao replicar. No entanto, a inserção de um cassete do transgene no sítio HE1 de ChAd83 produziu um vetor viável.

Exemplo 2: Produção de Vírus, Título do Vetor e Expressão de Vetores do Exemplo 1

[00128] Para identificar um modelo animal no qual avaliar a replicação do vetor, os vetores do adenovírus tipo C ChAd155 RC2 e adenovírus tipo E ChAd83 RC2 do Exemplo 1 foram avaliados quanto à sua capacidade de replicação, medida pelo título do vetor e número de cópias do genoma, em células de várias origens animais. Os resultados são mostrados na Tabela 1.

Tabela 1. Replicação e Expressão de ChAd155 RC2 e ChAd83 RC2 do Exemplo 1

<u>Linhagem celular:</u> <u>Espécies</u>	<u>Vetor</u>	<u>Título do</u> <u>Vetor</u>	<u>Cópia do</u> <u>Genoma</u>	<u>Expressão</u>	
				<u>Dia 2</u>	<u>Dia 7</u>
MRC5: Ser humano	ChAd155	+++	+++	++	++++
	ChAd83	+++++	+++++	+++	+++++
PK15: Suíno	ChAd155	+++++	+++++	NA	NA
	ChAd83	+++	++++	NA	NA
NMuLi: Camundongo	ChAd155	++	+++	+++	+++
	ChAd83	ND	+	++	++
Vero: Primata não humano	ChAd155	++	++++	+++	+++
	ChAd83	ND	+	+	+

ND = não detectado; NA = não disponível

[00129] Como mostrado na Tabela 1, as células MRC5 humanas e as células PK15 de suínos produziram altos títulos de vetores e altos

números de cópias do genoma de ChAd155 e ChAd83. As células NMuLi murinas e as células Vero de primatas não humanos também produziram RC ChAd155, mas em menor grau do que as células humanas ou suínas. RC ChAd83 falhou em crescer bem nas células NMuLi murinas e, surpreendentemente, nas células Vero de primatas não humanos.

[00130] Células MRC5 humanas, células NMuLi de camundongo e células Vero de primatas não humanos suportaram a expressão de RC ChAd155 até o dia 7. As células MRC5 humanas suportaram a expressão de RC ChAd83 até o dia 7, assim como as células NMuLi de camundongo e as células Vero de primatas não humanos, mas em menor extensão que as células humanas.

Produção de vírus

[00131] A Figura 2A mostra a quantidade de vírus produzido por células MRC5 primárias humanas infectadas com ChAd155 ou ChAd83, cada uma compreendendo o construto do vetor RC1 ou RC2 do Exemplo 1. As células foram coletadas sete dias após a infecção e o título do vetor foi avaliado em lisados celulares obtidos após três ciclos de congelamento e descongelamento. Os títulos de vetor foram medidos por análise quantitativa de PCR (QPCR) com iniciadores projetados para as respectivas regiões promotoras. A multiplicidade de infecção (moi) foi de 1250 partículas de vírus por célula. A produção de vírus é indicada no número de partículas de vírus por célula (vp/célula) acima das barras.

[00132] As células MRC5 humanas suportaram a produção de ChAd155 compreendendo RC1 ($2,17 \times 10^3$ vp/célula) ou RC2 ($4,40 \times 10^3$ vp/célula) e também suportaram a produção de ChAd83 compreendendo RC1 ($1,18 \times 10^4$ vp/célula) ou RC2 ($1,06 \times 10^5$ vp/célula). Como mostrado na Figura 2A, ChAd83 foi produzido em um nível superior ao de ChAd155; o vetor ChAd83 compreendendo RC2 foi a mais robusta das quatro combinações vírus/vetor.

[00133] A Figura 2B mostra a quantidade de vírus produzido por células MRC5 primárias humanas infectadas com ChAd83 compreendendo o construto do vetor RC1, RC2 ou RC3 do Exemplo 1. As células foram coletadas dois e sete dias após a infecção. Como na Figura 2A, os títulos de vetor foram medidos por análise quantitativa de PCR (QPCR) com iniciadores projetados para as respectivas regiões promotoras. A multiplicidade de infecção (moi) foi de 250 ou 1250 partículas de vírus por célula. A produção de vírus é indicada no número de partículas de vírus por célula (vp/célula) acima das barras.

[00134] As células MRC5 humanas suportaram a produção de ChAd83 compreendendo RC1, RC2 ou RC3. Como mostrado na Figura 2B, houve maior produção de vírus para os vetores RC2 e RC3 de ChAd83 do que para o vetor RC1. Também houve maior produção de vírus para o vetor ChAd83 RC2 HE2 do que o vetor RC3 HE1.

Número de cópias do genoma de vetor

[00135] Após a infecção, o vetor é replicado na célula e o número de cópias do genoma do vetor pode ser medido por QPCR. A replicação do DNA vetorial pode ocorrer mesmo em células não totalmente permissivas à replicação e propagação viral. QPCR do DNA do vetor fornece uma medida de replicação do vetor dentro da célula infectada, independentemente da capacidade do vírus de completar o ciclo de replicação e ser liberado como progênie viral madura. A replicação vetorial pode, assim, ser quantificada em espécies animais, tipos de tecidos e tipos de células que não são permissivos à replicação ou propagação do vírus ChAd.

[00136] O número de cópias do genoma do vetor foi medido em paralelo com o título do vetor e os resultados mostrados na Figura 3A e na Figura 3B.

[00137] Como acontece com a produção de vírus mostrada na Figura 2A, células MRC5 humanas foram infectadas com ChAd155 ou ChAd83,

cada um compreendendo o construto do vetor RC1 ou RC2 do Exemplo 1. As células foram coletadas sete dias após a infecção, o DNA total extraído, o genoma viral quantificado por QPCR e os resultados expressos como cópia do genoma do vetor por célula. A multiplicidade de infecção (moi) foi de 250 partículas de vírus por célula e o número de partículas de vírus por célula é indicado acima das barras, indicando cópias do genoma viral por célula. O número de cópias é diretamente proporcional ao nível de expressão do transgene.

[00138] Como mostrado na Figura 3A, a quantidade de replicação de DNA viral de RC1 ($6,21 \times 10^3$ vp/célula) e RC2 ($6,71 \times 10^3$ vp/célula) por ChAd155 foi similar. ChAd83 produziu mais RC1 ($2,76 \times 10^4$ vp/célula) e RC2 ($9,19 \times 10^4$ vp/célula) de DNA viral de ChAd155. O nível mais alto de replicação do DNA viral foi observado por ChAd83 RC2.

[00139] Como acontece com a produção de vírus mostrada na Figura 2B, células MRC5 humanas foram infectadas com ChAd83, compreendendo o construto do vetor RC1, RC2 ou RC3 do Exemplo 1. As células foram coletadas dois e sete dias após a infecção, o DNA total extraído, o genoma viral quantificado por QPCR e os resultados expressos como cópia do genoma do vetor por célula. A multiplicidade de infecção (moi) foi de 250 e 1250 partículas de vírus por célula e o número de partículas de vírus por célula é indicado acima das barras, indicando cópias do genoma viral por célula. O número de cópias é diretamente proporcional ao nível de expressão do transgene.

[00140] Conforme mostrado na Figura 3B, a quantidade de replicação do DNA viral foi maior para os vetores RC2 e RC3 de ChAd83 do que para o vetor RC1. Houve replicação do DNA viral comparável entre os vetores ChAd83 RC2 e RC3.

Exemplo 3: Número de Cópias do Genoma Adenoviral de Vetores do Exemplo 1

[00141] A eficiência dos vetores adenovirais competentes para replicação com os construtos do Exemplo 1, expressa como cópias de vetores por célula, foi avaliada em culturas de células derivadas de camundongos e primatas não humanos.

[00142] A Figura 4(a) mostra o número de cópias do genoma de vetores competentes para replicação desenvolvidos em células NMuLi hepáticas murinas cultivadas em monocamadas e infectadas com ChAd155 RC1, ChAd155 RC2, ChAd83 RC1 ou ChAd83 RC2 em uma multiplicidade de infecção de 250 partículas de vírus por célula. O DNA total foi extraído cinco dias após a infecção e a replicação do vetor foi medida por QPCR usando iniciadores de recozimento na região promotora do vetor.

[00143] Os resultados, expressos como cópias de vetores por célula, são mostrados na Figura 4(a). ChAd155 amplificou o vetor RC1 e RC2 com alta eficiência nas células NMuLi. ChAd155 replicou os vetores RC1 ($1,73 \times 10^4$) e RC2 ($1,92 \times 10^4$) aproximadamente no mesmo grau. ChAd83 foi menos eficiente que ChAd155 na replicação dos vetores RC1 e RC2. ChAd83 replicou o DNA do vetor apenas em pequenas quantidades nas células murinas. O vetor RC1 replicou em um nível de $5,47 \times 10^2$ cópias por célula e o vetor RC2 em um nível de $6,74 \times 10^2$ cópias por célula.

[00144] As células Vero de primatas não humanos também foram cultivadas em monocamadas e infectadas com ChAd155 RC1, ChAd155 RC2, ChAd83 RC1 ou ChAd83 RC2 (Figura 4(b)). Duas multiplicidades diferentes de infecção foram utilizadas: 50 e 250 partículas de vírus por célula. O DNA total foi extraído cinco dias após a infecção e a replicação do vetor foi medida por QPCR usando iniciadores de recozimento na região promotora do vetor.

[00145] Os resultados, expressos como cópias de vetores por célula, são mostrados na Figura 4(b). A linhagem de células Vero de primatas

era permissiva para ChAd155 RC1 ($3,71 \times 10^3$ cópias por célula em uma moi de 50 e $4,93 \times 10^4$ cópias por célula em uma moi de 250) e ChAd155 RC2 ($8,15 \times 10^3$ cópias por célula em uma moi de 50 e $7,05 \times 10^4$ cópias por célula em uma moi de 250). A linhagem de células de primatas Vero era pouco, se é que era, permissiva para ChAd83 RC1 ou ChAd83 RC2. Não foram detectados vetores ChAd83 RC1 ou ChAd83 RC2 expressos a partir de células Vero a um moi de 50. Em um moi de 250, ChAd83 replicou o vetor RC1 no nível de $1,13 \times 10^2$ cópias por célula e o vetor RC2 no nível de $1,29 \times 10^3$ cópias por célula.

Exemplo 4: Expressão do Transgene a Partir de Células Murinas e de Primatas não Humanos de Vetores do Exemplo 1

[00146] A análise Western blot foi realizada para comparar o nível de expressão do transgene por ChAd155 RC1 e ChAd155 RC2 em células murinas NMuLi (Figura 5(a)). As células foram infectadas com ChAd155 RC1 ou ChAd155 RC2 a uma multiplicidade de infecção de 50, 250 ou 1250 partículas virais por célula. As células foram coletadas dois e cinco dias após a infecção, extratos preparados usando métodos padrão e uma quantidade equivalente de extrato celular total carregado em géis de SDS-PAGE. Após a separação eletroforética, as proteínas foram transferidas para membranas de nitrocelulose, que foram testadas com um anticorpo monoclonal disponível comercialmente para o transgene da glicoproteína da raiva.

[00147] A Figura 5(a) demonstra que ChAd155 RC1 e ChAd155 RC2 expressam um transgene em células NMuLi murinas. A expressão foi observada dois e cinco dias após a infecção, indicada pela banda de cerca de 51 kDa, que corresponde ao peso molecular esperado da glicoproteína da raiva (RG). O vetor ChAd155 RC2 produziu um nível mais alto de expressão de transgene do que o vetor ChAd155 RC1 dois e cinco dias após a infecção.

[00148] A análise Western blot foi então realizada para comparar o

nível de expressão do transgene por ChAd155 RC1, ChAd155 RC2, ChAd83 RC1 e ChAd83 RC2 em células NMuLi murinas (Figura 5(b)). As células foram infectadas com ChAd155 RC1, ChAd155 RC2, ChAd83 RC1 ou ChAd83 RC2 com uma multiplicidade de infecção de 50, 250 ou 1250 partículas virais por célula (250 e 1250 para ChAd83 RC1). As células foram processadas para western blot. As células foram coletadas dois e sete dias após a infecção, os extratos preparados usando métodos padrão e uma quantidade equivalente de extrato carregado em géis de SDS-PAGE. Após a separação eletroforética, as proteínas foram transferidas para membranas de nitrocelulose, que foram testadas com um anticorpo monoclonal disponível comercialmente para o transgene da glicoproteína da raiva.

[00149] A Figura 5(b) demonstra que ChAd155 RC1, ChAd155 RC2, ChAd83 RC1 e ChAd83 RC2 expressam um transgene em células NMuLi murinas. A expressão foi observada dois e cinco dias após a infecção, indicada pela banda de cerca de 51 kDa, que corresponde ao peso molecular esperado da glicoproteína da raiva (RG). ChAd155 demonstrou expressão mais eficiente do transgene que ChAd83. Dois dias após a infecção, a expressão robusta de transgene por ChAd155 RC2 foi observada mesmo na baixa multiplicidade de 50 vp/célula, enquanto a expressão robusta de transgene por ChAd155 RC1 foi observada pela primeira vez em mois mais altas. Além disso, RC2 demonstrou expressão de transgene mais eficiente que o RC1 nos sorotipos virais ChAd155 e ChAd83. RC2 foi expresso com mais robustez que RC1 em cada uma das comparações diretas.

[00150] A análise Western blot foi realizada para comparar o nível de expressão do transgene por ChAd83 RC1, RC2 e RC3 em células MRC5 (Figura 5(c)). As células foram infectadas com ChAd83 RC1, RC2 ou RC3 em uma multiplicidade de infecção de 250 ou 1250 partículas

virais por célula. As células foram coletadas dois e sete dias após a infecção, os extratos preparados usando métodos padrão e uma quantidade equivalente de extrato celular total carregado em géis de SDS-PAGE. Após a separação eletroforética, as proteínas foram transferidas para membranas de nitrocelulose, que foram testadas com um anticorpo monoclonal disponível comercialmente para o transgene da glicoproteína da raiva.

[00151] A Figura 5(c) demonstra que todos os ChAd83 RC1, RC2 e RC3 expressam um transgene nas células MRC5. A expressão foi observada dois e sete dias após a infecção, indicada pela banda de cerca de 51 kDa, que corresponde ao peso molecular esperado da glicoproteína da raiva (RG). O vetor ChAd83 RC2 produziu um nível mais alto de expressão de transgene do que os vetores ChAd83 RC1 e RC3 dois e sete dias após a infecção. Não houve detecção de glicoproteína da raiva para os vetores RC1 e RC3 nos 7 dias.

Exemplo 5: Construção de Adenovírus Alternativos de Chimpanzé com um Cassete de Expressão Única

[00152] Como no Exemplo 1, adenovírus selvagens de chimpanzé tipo 155 (ChAd155) (WO 2016/198621) isolados de chimpanzés saudáveis usando procedimentos padrão foram construídos como vírus com defeito de replicação, conforme descrito em Sci Transl Med (2012) 4:1 e em WO 2010/086189.

[00153] No Exemplo 5, ChAd155 é construído inserindo um cassete de expressão único do transgene. Este cassete de expressão compreende o promotor clássico de CMV humano (hCMV), antígenos RSV F0ΔTM, N e M2-1 e, opcionalmente, um intensificador de WPRE. Este vetor é mostrado na Figura 6. O cassete de expressão é inserido na região E1 do adenovírus (após a região E1 ter sido deletada).

[00154] ChAd155 mostrado na Figura 6 compreende um transgene que codifica todos os antígenos RSV F0ΔTM, M2-1 e N, em que um sítio

de auto-clivagem ("2A") é incluído entre o antígeno RSV F0ΔTM e o antígeno RSV N.M2-1 composto, no qual um ligante flexível está incluído entre os antígenos RSV M2-1 e N.

[00155] O vetor ChAd155 RSV do Exemplo 5 compõe o polinucleotídeo da SEQ ID NO: 4, e codifica o polipeptídeo da SEQ ID NO: 5.

Exemplo 6: Construção de Adenovírus de Chimpanzé com um Cassete de Expressão Duplo

[00156] Novamente, adenovírus selvagens de chimpanzé tipo 155 (ChAd155) (WO 2016/198621) isolados de chimpanzés saudáveis usando procedimentos padrão foram construídos como vírus com defeito de replicação, conforme descrito em Sci Transl Med (2012) 4:1 e em WO 2010/086189.

[00157] O ChAd155 do Exemplo 6 é construído inserindo dois cassetes de expressão de transgene em dois sítios diferentes no adenovírus:

(1) Os primeiros componentes do cassete de expressão compreendem o promotor clássico de CMV humano (hCMV) e o antígeno RSV N.M2-1 composto. Este primeiro cassete de expressão é inserido na região E1 do adenovírus (após a região E1 ter sido deletada).

(2) O segundo cassete de expressão compreende um promotor de CMV humano clássico aprimorado (hCMV aprimorado), o antígeno RSV F0ΔTM e um intensificador de WPRE. Este primeiro cassete de expressão é inserido na região HE2 do adenovírus (após a região HE2 ter sido deletada).

[00158] Este vetor compreendendo um cassete de expressão duplo é mostrado na Figura 7.

[00159] No construto da Figura 7, Ad5E4orf6 foi substituído na região inicial do gene 4 (E4). A substituição é necessária para aumentar a produtividade nas células HEK 293.

Exemplo 7: Expressão de transgene do Cassete de Expressão Duplo

do Exemplo 6

[00160] A análise Western blot foi realizada para comparar o nível de expressão do transgene no vetor ChAd155 do Exemplo 6 (marcado como “Duplo” ou “Cassete Duplo” nas Figuras) nas células MRC5 com:

- (i) um vetor compreendendo um cassete de expressão F único (ChAd155-F0 Δ TM, marcado como “F0 Δ Tm”),
- (ii) um vetor compreendendo um cassete de expressão NM2 único (ChAd155-NM2, marcado como “NM2-1”) e
- (iii) o vetor do Exemplo 5 compreendendo um cassete de expressão único contendo os antígenos F e N.M2-1 RSV (ChAd155-F0 Δ TM.NM2, também marcado como “RSV”)

[00161] A análise Western blot é mostrada na Figura 8 e na Figura 9.

[00162] Como mostrado na Figura 8, as células foram infectadas com ChAd155-F0 Δ TM, ChAd155-F0 Δ TM.NM2 (“RSV”) ou o cassete duplo ChAd155 do Exemplo 6 com uma multiplicidade de infecção de 500 partículas virais por célula. Além disso, as células foram infectadas com ChAd155-F0 Δ TM.NM2 (“RSV”) em uma multiplicidade de infecção de 500 ou 1250 partículas virais por célula. As células foram coletadas 48 horas e 96 horas após a infecção, os extratos preparados usando métodos padrão e uma quantidade equivalente de extrato celular total carregado em géis de SDS-PAGE.

[00163] A Figura 8 mostra que o cassete duplo ChAd155 fornece um nível de expressão do antígeno F que é comparável ao ChAd155F0 Δ TM e superior ao ChAd155-F Δ TM.NM2 nas células MRC5.

[00164] Como mostrado na Figura 9, as células foram infectadas com ChAd155-NM2, ChAd155-F0 Δ TM.NM2 (“RSV”) ou o cassete duplo ChAd155 do Exemplo 6 com uma multiplicidade de infecção de 250 e 1250 partículas virais por célula. As células foram coletadas 48 horas após a infecção, os extratos preparados usando métodos padrão e uma quantidade equivalente de extrato celular total carregado em géis de

SDS-PAGE.

[00165] Na Figura 9, o cassete duplo ChAd155 fornece o nível de expressão de NM2-1 pelo menos comparável ao vetorChAd155-NM2 único e superior ao ChAd155-F Δ TM.NM2 (“RSV”) nas células MRC5.

Exemplo 8: Imunogenicidade do Cassete de Expressão Duplo do Exemplo 6

[00166] A imunogenicidade do cassete de expressão duplo do Exemplo 6 foi avaliada em camundongos CD1 (10 por grupo). O experimento foi realizado por injeção de 5×10^8 partículas virais por via intramuscular em camundongos. A resposta das células B foi medida 4 e 8 semanas após a imunização, medindo os títulos neutralizadores de RSV. Cada ponto representa a resposta em um único camundongo e a linha corresponde à média de cada grupo de doses. Os resultados desta análise são mostrados na Figura 10.

[00167] A Figura 10 mostra que o cassete duplo ChAd155 fornece uma resposta de células B comparável a ChAd155F0 Δ TM e superior à produzida por ChAd155-F0 Δ TM.NM2 (“RSV”).

[00168] A imunogenicidade do cassete de expressão duplo do Exemplo 6 também foi avaliada em ratos BALB/c (48, 11 ou 8 por grupo). O experimento foi realizado injetando 10^7 ou 10^6 partículas virais por via intramuscular. A resposta das células T foi medida 3 semanas após a imunização por Immunospot ligado à enzima IFN-gama *ex vivo* (ELISPOT) usando um epítipo de célula T do peptídeo M2 mapeado em camundongos BALB/c. Os resultados são mostrados na Figura 11, expressos como células formadoras de pontos IFN-gama (SFC) por milhão de esplenócitos. Cada ponto representa a resposta em um único camundongo e a linha corresponde à média de cada grupo de doses. A dose injetada em número de partículas virais é mostrada no eixo x. Os resultados são mostrados na Figura 11.

[00169] A Figura 11 mostra que o cassete duplo ChAd155 fornece

uma resposta de células T superior à produzida pelo cassete único ChAd155-F0ΔTM.NM2 (“RSV triplo”, cujos resultados são obtidos a partir de dados históricos). Essa diferença na resposta é maior para a dose de 10^6 vp.

[00170] A Figura 11 refere-se a “#camundongos positivos”, isto é, o número de camundongos que responderam à vacina.

Exemplo 9: Imunogenicidade do Cassete de Expressão Duplo do Exemplo 6 em vacas

[00171] O desenho do estudo está detalhado na Tabela 2 abaixo:

<u>Grupo</u>	<u>Nº de Vacas</u>	<u>Vacina</u>	<u>Via</u>	<u>Dose</u>	<u>Imunização</u>	<u>Fim do estudo</u>
Gp1	4	RSV ChAd155 único	Intramuscular (IM)	1×10^{11}	D0	D90
Gp2	4	RSV ChAd155 duplo	Intramuscular (IM)	1×10^{11}	D0	D90
Gp3	4	Salina	Intramuscular (IM)	N/A	D0	D90

[00172] “ChAd155 RSV único” é o ChAd155 do Exemplo 5 e “ChAd155 RSV duplo” é o ChAd155 do Exemplo 6.

[00173] Um total de 12 vacas adultas foram incluídas no estudo. As vacas variaram de 2,7 a 7,8 anos e tinham uma média de 4,8 anos.

[00174] Antes de serem inscritas no estudo, as vacas foram pré-triadas para anticorpos para RSV bovino (BRSV) por ELISA. Isso permitiu o estabelecimento de grupos de estudo com distribuição semelhante e título médio de Ab BRSV (para não influenciar nenhum dos grupos).

[00175] As amostras foram coletadas das vacas antes da vacinação (D-5 ou D0) e após a vacinação (D7, 10, 14, 28, 60, 90). Neste estudo, as vacas foram vacinadas com 1×10^{11} partículas virais de uma das duas vacinas ou com soro fisiológico no dia zero (D0).

[00176] Uma lavagem broncoalveolar (BAL) foi realizada no dia -5, 7, 10 ou 28 após a vacinação para isolar as células T nos pulmões da vaca. Em seguida, a produção de citocinas IFN-gama das células CD4+ e CD8+ após a estimulação com antígenos RSV (na forma de agrupamentos de peptídeos) codificados nas vacinas foi detectada usando coloração intracelular de citocinas (ICS) (ou seja, ICS para IFN γ foi utilizado

para detectar respostas de células T pulmonares nos animais). Os resultados do experimento são mostrados nas Figuras 12A e 12B. Pode-se concluir a partir deste experimento que RSV ChAd155-duplo induz respostas consistentes CD4⁺ e CD8⁺ específicas para RSV consistentes no lavado broncoalveolar (LBA).

[00177] Também foram coletadas amostras de sangue das vacas nos dias 0, 14, 28, 60 e 90 após a vacinação para a produção de citocinas IFN-gama de respostas CD4⁺ e CD8⁺ específicas para RSV das células mononucleares do sangue periférico (PBMC) a ser detectada usando coloração intracelular de citocinas (ICS) (isto é, ICS para IFN γ foi utilizado para detectar as respostas periféricas das células T). Os resultados do experimento são mostrados nas Figuras 13A e 13B. Com base nesses resultados, pode-se concluir que RSV ChAd155-duplo expande consistentemente as respostas preexistentes CD4⁺ e CD8⁺ específicas para RSV em PBMC.

[00178] As amostras de sangue também foram utilizadas para detectar anticorpos neutralizantes (nAbs) para RSV no soro (isto é, a resposta humoral periférica foi detectada). Os resultados do experimento são mostrados nas Figuras 14A e 14B. Estes resultados mostram que RSV ChAd155-duplo aumenta nAbs de RSV no soro, os quais são mantidos em níveis superiores aos valores basais 3 meses após a vacinação.

Exemplo 10: Imunogenicidade de ChAd155 duplo que codifica as proteínas G da raiva e RS2 NM2

[00179] Três vetores ChAd155 diferentes foram construídos neste experimento:

- ChAd155 que codifica as proteínas G da raiva (RG) e RSV NM2 (denominado “ChAd155 duplo” neste exemplo e ChAd155 duplo hCMV NM 2-1 - CASI RG WPRE);
- ChAd155 que codifica apenas a proteína G da raiva (RG) (denominado “ChAd155 RG” neste exemplo e ChAd155(Δ E4)CASI RG

WPRE); e

- O vetor ChAd155 mostrado na Figura 6, ou seja, o vetor com transgene que codifica todos os antígenos RSV F0ΔTM, M2-1 e N (denominado “ChAd155 RSV”).

[00180] Três doses diferentes do adenovírus ChAd155 duplo foram administradas a camundongos: uma dose mais alta de 10^7 partículas virais, e uma dose média de 10^6 partículas virais e uma dose mais baixa de 10^5 partículas virais.

[00181] Duas doses diferentes dos vetores ChAd155 RG e ChAd155 RSV foram administradas aos camundongos. Para ChAd155 RSV, essa era uma dose mais alta de 10^7 partículas de vacina e uma dose mais baixa de 10^6 partículas de vacina. Para ChAd155 RG, essa era uma dose mais alta de 10^6 partículas de vacina e uma dose mais baixa de 10^5 partículas de vacina. Os camundongos foram sacrificados três semanas depois e os esplenócitos testados por IFN γ ELISpot para resposta da célula T aos antígenos codificados pela vacina.

[00182] Os resultados deste experimento são mostrados na Figura 15A, 15B e 15C. Como pode ser visto da Figura 15A, 15B e 15C, o vetor ChAd155 duplo RG-NM2 mostra respostas imunológicas comparáveis globais aos vetores que codificam cada um dos antígenos RG e NM2 sozinhos.

[00183] A Figura 15C compara a resposta cumulativa a todos os antígenos codificados na dosagem comum de 10^6 vp usada para todos os três vetores diferentes. A proteína G da raiva é listada duas vezes (G1 e G2), pois dois conjuntos de peptídeos sobrepostos foram utilizados para cobrir toda a sequência da proteína

[00184] Portanto, a colocação dos dois antígenos no mesmo vetor produz ainda uma resposta imune comparável, permitindo que antígenos para diferentes patógenos sejam fornecidos no mesmo vetor.

Exemplo 11: Expressão de ChAd155 duplo que codifica as proteínas G

da raiva e RSV NM2 em células HeLa

[00185] Nos experimentos do Exemplo 11, as células HeLa foram infectadas com o “ChAd155 duplo” purificado, “ChAd155 RG” e “ChAd155 RSV” usados no Exemplo 10.

[00186] Multiplicidades de infecção (MOI) de 50, 250 e 1250 foram utilizadas neste experimento.

[00187] Para obter o Western Blot mostrado na Figura 16A (obtido sob condições de redução), o lisado celular foi coletado 48 horas após a infecção. O tamanho estimado de NM2-1 é de 65 kDa. A Figura 16A mostra um nível de expressão comparável para o cassete ChAd155 duplo e ChAd155 NM2-1. Além disso, o nível de expressão NM2-1 foi maior para o cassete ChAd155 duplo do que o vetor RSV ChAd155.

[00188] Para obter o Western Blot mostrado na Figura 16B, o sobrenadante foi coletado 48 horas após a infecção. O tamanho estimado da glicoproteína da raiva é 57,6 kDa. A Figura 16B mostra um nível de expressão comparável para os adenovírus ChAd155 duplo e ChAd155 RG.

[00189] Além disso, os dados de infectividade também foram coletados usando os quatro vetores diferentes. A infecciosidade do vírus purificado foi avaliada em células Procell 92 aderentes por Hexon Immunostaining. Os resultados são apresentados na Tabela 3 abaixo (vp = partícula de vírus, ifu = unidade infecciosa e R é a razão entre esses dois números). Os resultados de infecciosidade indicam que todos os vetores têm infecciosidade similar. Além disso, como todos os valores de R estavam abaixo de 300, a infecciosidade de todos os vetores foi considerada dentro da faixa de aceitabilidade.

Tabela 3:

	Vp/ml	Ifu/ml	R (vp/ifu)
ChAd155 hCMV NM 2-1 - CASI RG WPRE	5,51E+11	4,53E+09	122
ChAd155(Δ E4)hCMV-RSV	1,12E+11	1,05E+09	107

ChAd155(Δ E4)hCMV NM2-1	5,68E+11	4,26E+09	133
ChAd155(Δ E4)CASI RG WPRE	3,48E+11	3,35E+09	104

Exemplo 12: Imunogenicidade de ChAd155 duplo que codifica as proteínas G da raiva e RS2 E1

[00190] Dois vetores ChAd155 diferentes foram construídos neste experimento:

- ChAd155 que codifica uma proteína CRPV E2 modificada em um primeiro cassete de expressão e uma proteína CRPV E1 modificada em um segundo cassete de expressão (chamado “CRPV Duplo”); e
- ChAd155 que codifica uma fusão das proteínas E2 e E1 de CRPV modificadas em um cassete de expressão único (chamado “Fusão de CRPV”)

[00191] Duas doses diferentes dos dois vetores de adenovírus foram administradas a camundongos: uma dose mais elevada de 10^7 partículas virais, e uma dose mais baixa de 10^6 partículas virais. Os resultados deste experimento são mostrados na Figura 17. A Figura 17 é um IFN γ ELISpot em esplenócitos 3 semanas após a vacinação. Uma análise estatística foi realizada sobre os resultados e as diferenças entre a resposta dos diferentes vetores não foram consideradas estatisticamente significativas. No entanto, como pode ser visto na Figura 17, os vetores ChAd155 CRPV Duplo mostram frequência aumentada de camundongos que respondem na dose mais baixa que os vetores de fusão CRPV (6/6 camundongos com resposta positiva para a dose de 10^6 do vetor CRPV duplo, mas apenas 4/6 camundongos com resposta positiva para dose de 10^6 do vetor de fusão de CRPV).

Exemplo 13: Expressão de ChAd155 duplo que codifica as proteínas E2 e E1 de CRPV

[00192] Os dois vetores diferentes ChAd155 usados no Exemplo 12 também foram usados no Exemplo 13.

[00193] Multiplicidades de infecção (MOI) de 250 e 1250 foram utilizadas neste experimento. O lisado celular foi coletado 48 horas após a infecção. O tamanho estimado da proteína E1 modificada é de 48 kDa, a proteína E2 modificada é de 35 kDa e a proteína de fusão que contém as proteínas E1 e E2 modificadas é de 88 kDa.

[00194] A Figura 18 mostra um Western blot (obtido sob condições de redução) ilustrando que houve melhor expressão das proteínas E1 e E2 modificadas pelo vetor de CRPV dual do que o vetor de fusão de CRPV. As colunas “Pvj” mostradas na Figura 18 são os controles usados.

[00195] Além disso, os dados de infectividade foram coletados usando dois vetores diferentes. A infecciosidade do vírus purificado foi avaliada em células Procell 92 aderentes por Hexon Immunostaining. Os resultados são apresentados na Tabela 4 abaixo (vp = partícula de vírus, ifu = unidade infecciosa e R é a razão entre esses dois números). Os resultados de infecciosidade indicam que os dois vetores têm infecciosidade similar. Além disso, como todos os valores de R estavam abaixo de 300, a infecciosidade de todos os vetores foi considerada dentro da faixa de aceitabilidade.

Tabela 4:

	Vp/ml	Ifu/ml	R (vp/ifu)
ChAd155 hCMV CRPV DE2DE1 Fusão HA WPRE “Fusão CRPV”	4,52E+11	2,97E+09	153
ChAd155 hCMV CRPV DE2 HA WPRE - CMV CRPV DE1 HA WPRE aprim. “CRPV Duplo”	6,20E+11	4,39E+09	143

DESCRIÇÃO DAS SEQUÊNCIAS

SEQ ID NO: 1 Sequência de polinucleotídeos que codifica

ChAd155 selvagem

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGA-
TGGGCGGCG-
CGGGGCGGGAGGCGGGTCCGGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGG

AAGTGGACTTTGTAAGTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGG-
TAAAAGTGACG-
TTTTCCGTGCGCGACAACGCCCACGGGAAGTGACATTTTTCCCGCGGTTTTT
ACCGGATGTTGTAGTGAATTTGGGCGTAACCAAGTAAGATTTGGCCATTTT-
CGCGGGAAAAC-
TGAAACGGGGAAGTGAAATCTGATTAATTTTCGCGTTAGTCATACCGCGTAAT
ATTTGTGCGAGGGCCGAGGGACTTTGGCCGATTACGTGGAGGACTCGCCCA-
GGTG-
TTTTTTGAGGTGAATTTCCGCGTTCCGGGTCAAAGTCTCCGTTTTATTATTA
TAGTCAGCTGACGCGGAGTGTATTTATACCCTCTGATCTCGTCAAGTGG-
CCACTCTTGAGTG-
CCAGCGAGTAGAGTTTTCTCCTCTGCCGCTCTCCGCTCCGCTCCGCTCGGCT
CTGACACCGGGGAAAAAATGAGACATTTACCTACGATGGCGGTGTGCT-
CACCGGCCAG-
CTGGCTGCTGAAGTCCTGGACACCCTGATCGAGGAGGTATTGGCCGATAATT
ATCCTCCCTCGACTCCTTTTGAGCCACCTACACTTCACGAATC-
TACGATCTGGATGTGG-
TGGGGCCCAGCGATCCGAACGAGCAGGCGGTTTCCAGTTTTTTTCCAGAGTC
CATGTTGTTGGCCAGCCAGGAGGGGGTCGAACTTGAGAC-
CCCTCCTCCGATCGTGGATT-
CCCCCGATCCGCCGCAGCTGACTAGGCAGCCCGAGCGCTGTGCGGGACCTGA
GACTATGCCCCAGCTGCTACCTGAGGTGATCGATCTCACCTGTAATGAG-
TCTGGTTTTCCAC-
CCAGCGAGGATGAGGACGAAGAGGGTGAGCAGTTTGTGTTAGATTCTGTGGA
ACAACCCGGGCGAGGATGCAGGTCTTGTCAATATCACCGGAAAAACACAG-
GAGACTCCCAGA-
TTATGTGTTCTCTGTGTTATATGAAGATGACCTGTATGTTTATTTACAGTAA
GTTTATCATCTGTGGGCAGGTGGGCTATAGTGTGGGTGGTGG-
TCTTTGGGGGGTTTTTTAA-
TATATGTCAGGGGTATGCTGAAGACTTTTTTATTGTGATTTTTTAAAGGTCC

AGTGTCTGAGCCCGAGCAAGAACCTGAACCGGAGCCTGAGCCTTCTCG-
CCCCAGGAGAAAG-
CCTGTAATCTTAACTAGACCCAGCGCACCGGTAGCGAGAGGCCTCAGCAGCG
CGGAGACCACCGACTCCGGTGCTTCCTCATCACCCCCGGAGATTCAC-
CCCCTGGTGCCCCTG-
TGTCCCGTTAAGCCCGTTGCCGTGAGAGTCAGTGGGCGGCGGTCTGCTGTGG
AGTGCATTGAGGACTTGCTTTTTGATTCACAGGAACCTTTGGACTTGAG-
CTTGAAACG-
CCCCAGGCATTAAACCTGGTCACCTGGACTGAATGAGTTGACGCCTATGTTT
GCTTTTGAATGACTTAATGTGTATAGATAATAAAGAGTGAGATAATGTTT-
TAATTGCATGG-
TGTGTTTAACTTGGGCGGAGTCTGCTGGGTATATAAGCTTCCCTGGGCTAAA
CTTGGTTACACTTGACCTCATGGAGGCCTGGGAGTGTTTGGAGAACTTTG-
CCGGAGTTCGTG-
CCTTGCTGGACGAGAGCTCTAACAATACCTCTTGGTGGTGGAGGTATTTGTG
GGGCTCTCCCCAGGGCAAGTTAGTTTGTAGAATCAAGGAGGATTACAAG-
TGGGAA-
TTTGAAGAGCTTTTGAAATCCTGTGGTGAGCTATTGGATTCTTTGAATCTAG
GCCACCAGGCTCTCTTCCAGGAGAAGGTCATCAGGACTTTGGATTTTTCCA-
CACCGGGGCG-
CATTGCAGCCGCGGTTGCTTTTCTAGCTTTTTTTGAAGGATAGATGGAGCGAA
GAGACCCACTTGAGTTCGGGCTACGTCCTGGATTTTCTGGCCATGCAACTG-
TGGAGAGCATG-
GATCAGACACAAGAACAGGCTGCAACTGTTGTCTTCCGTCCGCCCCGTTGCTG
ATTCCGGCGGAGGAGCAACAGGCCGGGTGAGAGGACCGGGCCCGTCGGGA-
TCCGGAGGA-
GAGGGCACCGAGGCCGGGCGAGAGGAGCGCGCTGAACCTGGGAACCGGGCTG
AGCGGCCATCCACATCGGGAGTGAATGTGCGGCAGGTGGTGGATCTTTTT-
CCAGAACTGCGG-
CGGATTTTGACTATTAGGGAGGATGGGCAATTTGTTAAGGGTCTTAAGAGGG

AGAGGGGGGCTTCTGAGCATAACGAGGAGGCCAGTAATTTAGCTTTTAG-
CTTGATGACCAGA-
CACCGTCCAGAGTGCATCACTTTTCAGCAGATTAAGGACAATTGTGCCAATG
AGTTGGATCTGTTGGGTCAGAAGTATAGCATAGAGCAGCTGACCACTTAC-
TGGCTGCAG-
CCGGGTGATGATCTGGAGGAAGCTATTAGGGTGTATGCTAAGGTGGCCCTGC
GGCCCGATTGCAAGTACAAGCTCAAGGGGCTGGTGAATATCAGGAATTG-
TTGCTACATTT-
CTGGCAACGGGGCGGAGGTGGAGATAGAGACCGAAGACAGGGTGGCTTTTCAG
ATGCAGCATGATGAATATGTGGCCGGGGGTGCTGGGCATGGACGGGGTGG-
TGATTATGAATG-
TGAGGTTTCACGGGGCCCAACTTTAACGGCACGGTGTTTTTGGGGAACACCAA
CCTGGTCCTGCACGGGGTGAGCTTCTATGGGTTTAAACAACACCTGTGTGGA-
GGCCTGGAC-
CGATGTGAAGGTCCGCGGTTGCGCCTTTTATGGATGTTGGAAGGCCATAGTG
AGCCGCCCTAAGAGCAGGAGTTCCATTAAGAAATGCTTGTTTGAGAGGTG-
CACCTTGGGGA-
TCCTGGCCGAGGGCAACTGCAGGGTGCGCCACAATGTGGCCTCCGAGTGCGG
TTGCTTCATGCTAGTCAAGAGCGTGGCGGTAATCAAGCATAATATGGTGTG-
CGGCAACAG-
CGAGGACAAGGCCTCACAGATGCTGACCTGCACGGATGGCAACTGCCACTTG
CTGAAGACCATCCATGTAACCAGCCACAGCCGGAAGGCCTGGCCCGTGTT-
CGAGCACAAC-
TTGCTGACCCGCTGCTCCTTGCATCTGGGCAACAGGCGGGGGGTGTTCTCTGC
CCTATCAATGCAACTTTAGTCACACCAAGATCTTGCTAGAGCCCGAGAG-
CATGTCCAAGG-
TGAAC TTGAACGGGGTGTTTGACATGACCATGAAGATCTGGAAGGTGCTGAG
GTACGACGAGACCAGGTCCCGGTGCAGACCCTGCGAGTGCGGGGGCAAGCA-
TATGAGGAAC-
CAGCCCGTGATGCTGGATGTGACCGAGGAGCTGAGGACAGACCACTTGGTTC

TGGCCTGCACCAGGGCCGAGTTTGGTTCTAGCGATGAAGACACAGA-
TTGAGGTGGGTGAG-
TGGGCGTGGCCTGGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGT
CTCTTTATTTGTGTTGCAGAGACCGCCGGAGCCATGAGCGGGAGCAGCAG-
CAGCAGCAGTAG-
CAGCAGCGCCTTGGATGGCAGCATCGTGAGCCCTTATTTGACGACGCGGATG
CCCCACTGGGCCGGGGTGCCTCAGAATGTGATGGGCTCCAGCATCGACGG-
CCGACCCG-
TCCTGCCCCGCAAATTCCGCCACGCTGACCTATGCGACCGTCGCGGGGACGCC
GTTGGACGCCACCGCCGCCGCCGCCGCCACCGCAGCCGCCTCGGCCGTGCG-
CAGCCTGG-
CCACGGACTTTGCATTCCTGGGACCACTGGCGACAGGGGCTACTTCTCGGGC
CGCTGCTGCCGCCGTTTCGCGATGACAAGCTGACCGCCCTGCTGGCGCAG-
TTGGATGCGCT-
TACTCGGGAACCTGGGTGACCTTTCTCAGCAGGTCATGGCCCTGCGCCAGCAG
GTCTCCTCCCTGCAAGCTGGCGGGAATGCTTCTCCACAAATGCCGTTTAA-
GATAAA-
TAAAACCAGACTCTGTTTGGATTAAAGAAAAGTAGCAAGTGCATTGCTCTCT
TTATTTTATAATTTTCCGCGCGCGATAGGCCCTAGACCAGCGTTCTCGG-
TCGTTGAGGGTG-
CGGTGTATCTTCTCCAGGACGTGGTAGAGGTGGCTCTGGACGTTGAGATACA
TGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACCACTGCAGAGCTT-
CATGCTCCGGGG-
TGGTGTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGGTGCCATAA
AATGTCCTTCAGCAGCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAG-
TGTTTA-
CAAAACGGTTAAGTTGGGAAGGGTGCATTCGGGGAGAGATGATGTGCATCTT
GGACTGTATTTTTAGATTGGCGATGTTTCCGCCAGATCCCTTCTGGGATT-
CATGTTGTGCA-
GGACCACCAGTACAGTGTATCCGGTGCACCTGGGGAATTTGTCATGCAGCTT

AGAGGGAAAAGCGTGGAAGAACTTGGAGACGCCTTTGTGGCCTCCCAGA-
TTTTCCATGCATT-
CGTCCATGATGATGGCAATGGGCCCCGCGGGAGGCAGCTTGGGCAAAGATATT
TCTGGGGTCGCTGACGTCGTAGTTGTGTTCCAGGGTGAGGTCGTCATAGG-
CCATTTTTTA-
CAAAGCGCGGGCGGAGGGTGCCCGACTGGGGGATGATGGTCCCCTCTGGCCC
TGGGGCGTAGTTGCCCTCGCAGATCTGCATTTCCCAGGCCTTAATCTCGGA-
GGGGGGAATCA-
TATCCACCTGCGGGGCGATGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAA
CTGGGATGAGAGCAGGTTTCTAAGCAGCTGTGATTTTCCACAACCGGTGGG-
CCCATAAATAA-
CACCTATAACCGGTTGCAGCTGGTAGTTTAGAGAGCTGCAGCTGCCGTCGTC
CCGGAGGAGGGGGGCCACCTCGTTGAGCATGTCCCTGACGCGCATGTT-
CTCCCCGACCAGA-
TCCGCCAGAAGGCGCTCGCCGCCAGGGACAGCAGCTCTTGCAAGGAAGCAA
AGTTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATGTTTTTCAGGG-
TCTGGCTCAGCAG-
CTCCAGGCGGTCCCAGAGCTCGGTGACGTGCTCTACGGCATCTCTATCCAGC
ATATCTCCTCGTTTCGCGGGTTGGGGCGACTTTCGCTGTAGGGCACCAAG-
CGGTGGTCG-
TCCAGCGGGGCCAGAGTCATGTCCTTCCATGGGCGCAGGGTCCTCGTCAGGG
TGGTCTGGGTCACGGTGAAGGGGTGCGCTCCGGGCTGAGCGCTTGCCAAGG-
TGCGCTTGAGG-
CTGGTTCTGCTGGTGCTGAAGCGCTGCCGGTCTTCGCCCTGCGCGTCGGCCA
GGTAGCATTTGACCATGGTGTCATAGTCCAGCCCCTCCGCGGCGTG-
TCCCTTGCGCGCAG-
CTTGCCCTTGAGGTTGGCGCCGCACGAGGGGCAGAGCAGGCTCTTGAGCGCG
TAGAGCTTGGGGGCGAGGAAGACCGATTTCGGGGGAGTAGGCGTCCGCGCCG-
CAGACCCCGCA-
CACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCGCGCCGGGTCAAAA

ACCAGGTTTCCCCCATGCTTTTTGATGCGTTTCTTACCTCGGG-
TCTCCATGAGGTGGTG-
TCCCCGCTCGGTGACGAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGG
GGTCTTTTCTCCAGGGGGGTCCCTCGGTCTTCCTCGTAGAGGAACTCGGAC-
CACTCTGAGA-
CGAAGGCCCGCGTCCAGGCCAGGACGAAGGAGGCTATGTGGGAGGGGTAGCG
GTCGTTGTCCACTAGGGGGTCCACCTTCTCCAAGGTGTGAAGACACATG-
TCGCCTTCCTCGG-
CGTCCAGGAAGGTGATTGGCTTGTAGGTGTAGGCCACGTGACCGGGGGTTCC
TGACGGGGGGGTATAAAAGGGGGTGGGGGCGCGCTCGTCGTCACTCTCTT-
CCGCATCGCTG-
TCTGCGAGGGCCAGCTGCTGGGGTGAGTATTCCTCTCGAAGGCGGGCATGA
CCTCCGCGCTGAGGTTGTCAGTTTCCAAAAACGAGGAGGATTTGATGTT-
CACCTG-
TCCCAGAGGTGATACCTTTGAGGGTACCCGCGTCCATCTGGTCAGAAAACACG
ATCTTTTTTATTGTCCAGCTTGGTGGCGAACGACCCGTAGAGGGCGTTGGA-
GAGCAGCTTGG-
CGATGGAGCGCAGGGTCTGGTTCTTGTCCCTGTCGGCGCGCTCCTTGGCCGC
GATGTTGAGCTGCACGTACTCGCGCGCGACGCAGCGCCACTCGGGGAAGA-
CGGTGGTGCG-
CTCGTCGGGCACCAGGCGCACGCGCCAGCCGCGGTTGTGCAGGGTGACCAGG
TCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTTGGTCCAGCAGAGA-
CGGCCG-
CCCTTGCGCGAGCAGAAGGGGGGCAGGGGGTTCGAGCTGGGTCTCGTCCGGGG
GGTCCGCGTCCACGGTGAAAACCCCGGGGCGCAGGCGCGCGTCTGAAGTAG-
TCTATCTTG-
CAACCTTGCAATGTCCAGCGCCTGCTGCCAGTCGCGGGCGGCGAGCGCGCGCT
CGTAGGGGTTGAGCGGCGGGCCCCAGGGCATGGGGTGGGTGAGTGCGGAGG-
CGTACATGCCG-
CAGATGTCATAGACGTAGAGGGGCTCCCGCAGGACCCCGATGTAGGTGGGGT

AGCAGCGGCCGCCGCGGATGCTGGCGCGCACGTAGTCATACAGCTCGTG-
CGAGGGGGCGAG-
GAGGTCGGGGCCAGGTTGGTGCGGGCGGGGCGCTCCGCGCGGAAGACGATC
TGCCTGAAGATGGCATGCGAGTTGGAAGAGATGGTGGGGCGCTGGAAGACG-
TTGAAGCTGG-
CGTCCTGCAGGCCGACGGCGTCGCGCACGAAGGAGGCGTAGGAGTCGCGCAG
CTTGTGTACCAGCTCGGCGGTGACCTGCACGTCGAGCGCGCAGTAG-
TCGAGGGTCTCGCGGA-
TGATGTCATATTTAGCCTGCCCCCTTCTTTTTCCACAGCTCGCGGTTGAGGAC
AAACTCTTCGCGGTCTTTCAGTACTCTTGATCGGGAAACCGTCCGGTT-
CCGAACGG-
TAAGAGCCTAGCATGTAGAACTGGTTGACGGCCTGGTAGGCGCAGCAGCCCT
TCTCCACGGGGAGGGCGTAGGCCTGCGCGGCCTTGCGGAGCGAGGTGTGGG-
TCAGGG-
CGAAGGTGTCCCTGACCATGACTTTGAGGTACTGGTGCTTGAAGTCGGAGTC
GTCGCAGCCGCCCCGCTCCCAGAGCGAGAAGTCGGTGCGCTTCTTGGAG-
CGGGGGTTGGGCA-
GAGCGAAGGTGACATCGTTGAAGAGGATTTTGCCCGCGCGGGGCATGAAGTT
GCGGGTGATGCGGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGAC-
CTGGGCGGCGAG-
CACGATCTCGTCGAAGCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGG
AAGCGGGGCGGGCCCTTTACGGTGGGCAGCTTCTTTAGCTCTTCGTAGG-
TGAGCTCCTCGGG-
CGAGGCGAGGCCGTGCTCGGCCAGGGCCCAGTCCGCGAGGTGCGGGTTGTCT
CTGAGGAAGGACTTCCAGAGGTCGCGGGCCAGGAGGGTCTGCAGGCGG-
TCTCTGAAGG-
TCCTGAACTGGCGGGCCACGGCCATTTTTTCGGGGGTGATGCAGTAGAAGGT
GAGGGGGTCTTGCTGCCAGCGGTCCCAGTCGAGCTGCAGGGCGAGGTCGCG-
CGCGGCGGTGA-
CCAGGCGCTCGTCGCCCCCGAATTTTCATGACCAGCATGAAGGGCACGAGCTG

CTTTCCGAAGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGA-
CAAAGAGGCGCTCCG-
TGCGAGGATGCGAGCCGATCGGGAAGAACTGGATCTCCCGCCACCAGTTGGA
GGAGTGGCTGTTGATGTGGTGGAAAGTAGAAGTCCCGTCGCCGGGCCGAACA-
CTCGTGCTGG-
CTTTTGTAAAAGCGAGCGCAGTACTGGCAGCGCTGCACGGGCTGTACCTCAT
GCACGAGATGCACCTTTCGCCCCGCGCACGAGGAAGCCGAGGGGAAA-
TCTGAGCCCCCG-
CCTGGCTCGCGGCATGGCTGGTTCTCTTCTACTTTGGATGCGTGTCCGTCTC
CGTCTGGCTCCTCGAGGGGTGTTACGGTGGAGCGGACCACCACGCCGCG-
CGAGCCGCAGG-
TCCAGATATCGGCGCGCGGGCGGTTCGGAGTTTGATGACGACATCGCGCAGCTG
GGAGCTGTCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTCAGCCGGGAG-
TTCTTGCAGG-
TTCACCTCGCAGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTGA
TCTCTAGGGGCGTGTTGGTGGCGGCGTCGATGGCTTGCAGGAGCCCGCAG-
CCCCGGGGGG-
CGACGACGGTGCCCCGCGGGGTGGTGGTGGTGGTGGCGGTGCAGCTCAGAAG
CGGTGCCGCGGGCGGGCCCCCGAGGTAGGGGGGGCTCCGGTCCCGCGGG-
CAGGGGCGGCAG-
CGGCACGTCGGCGTGGAGCGCGGGCAGGAGTTGGTGCTGTGCCCGGAGGTTG
CTGGCGAAGGCGACGACGCGGCGGTTGATCTCCTGGATCTGGCGCCTCTG-
CGTGAAGACGA-
CGGGCCCGGTGAGCTTGAACCTGAAAGAGAGTTCGACAGAATCAATCTCGGT
GTCATTGACCGCGGCCTGGCGCAGGATCTCCTGCACGTCTCCCGAGTTG-
TCTTGGTAGG-
CGATCTCGGCCATGAACTGCTCGATCTCTTCCTCCTGGAGGTCTCCGCGTCC
GGCGCGTTCCACGGTGGCCGCCAGGTCGTTGGAGATGCGCCCCATGAGCTG-
CGAGAAGGCG-
TTGAGTCCGCCCTCGTTCCAGACTCGGCTGTAGACCACGCCCCCTGGTCAT

CGCGGGCGCGCATGACCACCTGCGCGAGGTTGAGCTCCACGTGCCGCG-
CGAAGACGGCGTAG-
TTGCGCAGACGCTGGAAGAGGTAGTTGAGGGTGGTGGCGGTGTGCTCGGCCA
CGAAGAAGTTCATGACCCAGCGGCGCAACGTGGATTTCGTTGATG-
TCCCCCAAGGCCTCCAG-
CCGTTCCATGGCCTCGTAGAAGTCCACGGCGAAGTTGAAAACTGGGAGTTG
CGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGGCGA-
CGGTGTCGCGCAC-
CTCGCGCTCGAAGGCTATGGGGATCTCTTCCTCCGCTAGCATCACCACCTCC
TCCTCTTCCTCCTCTTCTGGCACTTCCATGATGGCTTCCTCCTCTT-
CGGGGGGTGGCGGCGG-
CGGCGGTGGGGGAGGGGGCGCTCTGCGCCGGCGGGCGGCACCGGGAGGCGG
TCCACGAAGCGCGCGATCATCTCCCCGCGGCGGCGGCATGGTCTCGG-
TGACGGCGCGG-
CCGTTCTCCCGGGGGCGCAGTTGGAAGACGCCGCCGGACATCTGGTGCTGGG
GCGGGTGGCCGTGAGGCAGCGAGACGGCGCTGACGATGCATCTCAACAA-
TTGCTGCGTAGG-
TACGCCGCCGAGGGACCTGAGGGAGTCCATATCCACCGGATCCGAAAACCTT
TCGAGGAAGGCGTCTAACCAGTCGCAGTCGCAAGGTAGGCTGAGCACCG-
TGGCGGGCGG-
CGGGGGGTGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGATGTAATTGAAG
TAGGCGGACTTGACACGGCGGATGGTCGACAGGAGCACCATGTCCTTGGG-
TCCGGCCTGCTG-
GATGCGGAGGCGGTCTGGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGG
TCCTTGTAGTAGTCTTGTCATGAGCCTTTCCACCGGCACCTCTTCTCCTT-
CCTCTTCTGCTT-
CTTCCATGTCTGCTTCGGCCCTGGGGCGGCGCCGCGCCCCCTGCCCCCAT
GCGCGTGACCCCGAACCCCTGAGCGGTTGGAGCAGGGCCAGGTCTGGCGA-
CGACGCGCTCGG-
CCAGGATGGCCTGCTGCACCTGCGTGAGGGTGGTTTGAAGTCATCCAAGTC

CACGAAGCGGTGGTAGGCGCCCGTGTTGATGGTGTAGGTGCAGTTGG-
CCATGACGGACCAG-
TTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTACCTGAGTCGCGAGT
AGGCGCGGGAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAGGTACTGG-
TAGCCCACCAG-
GAAGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCT
CCGGGGGCCAGGTCTTCCAGCATGAGGCGGTGGTAGGCGTAGATGTACCTG-
GACATCCAGG-
TGATACCCGCGGCGGTGGTGGAGGCGCGCGGGAAGTCGCGCACCCGGTTCCA
GATGTTGCGCAGGGGCAGAAAGTGCTCCATGGTAGGCGTGCTCTGTCCAGT-
CAGACGCGCG-
CAGTCGTTGATACTCTAGACCAGGGAAAACGAAAGCCGGTCAGCGGGCACTC
TTCCGTGGTCTGGTGAATAGATCGCAAGGGTATCATGGCGGAGGGCCTCGG-
TTCGAG-
CCCCGGGTCCGGGCCGGACGGTCCGCCATGATCCACGCGGTTACCGCCCGCG
TGTCGAACCCAGGTGTGCGACGTCAGACAACGGTGGAGTGTTTCCTTTTGG-
CGTTTTTCTGG-
CCGGGCGCCGGCGCCGCGTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTAAG
TGGCTCGCTCCCCGTAGCCGGAGGGATCCTTGCTAAGGGTTGCGTTGCGG-
CGAACCCCGGTT-
CGAATCCCGTACTCGGGCCGGCCGGACCCGCGGCTAAGGTGTTGGATTGGCC
TCCCCCTCGTATAAAGACCCCGCTTGCGGATTGACTCCGGACACGGGGA-
CGAGCCCCTTTTA-
TTTTTGCTTTCCCCAGATGCATCCGGTGCTGCGGCAGATGCGCCCCCGCCC
CAGCAGCAGCAACAACACCAGCAAGAGCGGCAGCAACAGCAGCGGGAGT-
CATGCAGGG-
CCCCCTCACCACCCCTCGGCGGGCCGGCCACCTCGGCGTCCGCGGCCGTGTC
TGGCGCCTGCGGCGGCGGCGGGGGGCGGCTGACGACCCCGAGGAG-
CCCCCGCGGCGCAGGG-
CCAGACACTACCTGGACCTGGAGGAGGGCGAGGGCCTGGCGCGGCTGGGGGC

GCCGTCTCCCGAGCGCCACCCGCGGGTGCAGCTGAAGCGCGACTCGCG-
CGAGGCGTACGTG-
CCTCGGCAGAACCTGTTTCAGGGACCGCGCGGGCGAGGAGCCCGAGGAGATGC
GGGACAGGAGGTTTCAGCGCAGGGCGGGAGCTGCGGCAGGGGCTGAACCG-
CGAGCGGCTGCTG-
CGCGAGGAGGACTTTGAGCCCGACGCGCGGACGGGGATCAGCCCCGCGCGCG
CGCACGTGGCGGCCCGCCGACCTGGTGACGGCGTACGAGCAGACGGTGAAC-
CAGGAGATCAAC-
TTCCAAAAGAGTTTCAACAACCACGTGCGCACGCTGGTGGCGCGCGAGGAGG
TGACCATCGGGCTGATGCACCTGTGGGACTTTGTAAGCGCGCTGGTGCA-
GAACCCCAACAG-
CAAGCCTCTGACGGCGCAGCTGTTCTGATAGTGCAGCACAGCAGGGACAAC
GAGGCGTTTAGGGACGCGCTGCTGAACATCACCGAGCCCGAGGGTCGGTGG-
CTGCTGGAC-
CTGATTAACATCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCCTGG
CCGACAAGGTGGCGGCCATCAACTACTCGATGCTGAGCCTGGGCAAGTTT-
TACGCGCGCAA-
GATCTACCAGACGCCGTACGTGCCCATAGACAAGGAGGTGAAGATCGACGGT
TTTTACATGCGCATGGCGCTGAAGGTGCTCACCCTGAGCGACGACCTGGG-
CGTGTACCG-
CAACGAGCGCATCCACAAGGCCGTGAGCGTGAGCCGGCGGCGCGAGCTGAGC
GACCGCGAGCTGATGCACAGCCTGCAGCGGGCGCTGGCGGGCGCCGGCAG-
CGGCGACAGGGA-
GGCGGAGTCCTACTTCGATGCGGGGGCGGACCTGCGCTGGGCGCCCAGCCGG
CGGGCCCTGGAGGCCGCGGGGGTCCGCGAGGACTATGACGAGGACGGCGAG-
GAGGATGAGGA-
GTACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCGGGTGGTGTTC
GGTAGATGCAAGACCCGAACGTGGTGGACCCGGCGCTGCGGGCGGCTCTG-
CAGAGCCAGCCG-
TCCGGCCTTAACTCCTCAGACGACTGGCGACAGGTCATGGACCGCATCATGT

CGCTGACGGCGCGTAACCCGGACGCGTTCCGGCAGCAGCCGCAGGCCAACA-
GGCTCTCCG-
CCATCCTGGAGGCGGTGGTGCCTGCGCGCTCGAACCCACGCACGAGAAGGT
GCTGGCCATAGTGAACGCGCTGGCCGAGAACAGGGCCATCCGCCCCGGA-
CGAGGCCGGGCTGG-
TGTACGACGCGCTGCTGCAGCGCGTGGCCCGCTACAACAGCGGCAACGTGCA
GACCAACCTGGACCGGCTGGTGGGGGACGTGCGCGAGGCGGTGGCGCAGCG-
CGAGCGCGCG-
GATCGGCAGGGCAACCTGGGCTCCATGGTGGCGCTGAATGCCTTCCTGAGCA
CGCAGCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAACTTTG-
TGAGCGCGCTGCGG-
CTGATGGTGACCGAGACCCCCAGAGCGAGGTGTACCAGTCGGGCCCGGACT
ACTTCTTCCAGACCAGCAGACAGGGCCTGCAGACGGTGAACCTGAGCCAGG-
CTTTCAAGAAC-
CTGCGGGGGCTGTGGGGCGTGAAGGCGCCACCGGCGACCGGGCGACGGTGT
CCAGCCTGCTGACGCCCAACTCGCGCCTGCTGCTGCTGCTGATCGCGCCG-
TTCACGGACAG-
CGGCAGCGTGTCCCGGGACACCTACCTGGGGCACCTGCTGACCCTGTACCGC
GAGGCCATCGGGCAGGCGCAGGTGGACGAGCACACCTTCCAGGAGATCAC-
CAGCGTGAGCCG-
CGCGCTGGGGCAGGAGGACACGAGCAGCCTGGAGGCGACTCTGAACTACCTG
CTGACCAACCGGCGGCAGAAGATTCCCTCGCTGCACAGCCTGACCTCCGAG-
GAGGAGCG-
CATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAACCTGATGCGCGACGGG
GTGACGCCCAGCGTGGCGCTGGACATGACCGCGCGCAACATGGAACCGGG-
CATGTACGCCG-
CGCACCGGCCTTACATCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGC
CGTGAACCCCGAGTACTTTACCAACGCCATCCTGAACCCGCACTGG-
CTCCCGCCGCCCGGG-
TTCTACAGCGGGGGCTTCGAGGTCCCGGAGACCAACGATGGCTTCCTGTGGG

ACGACATGGACGACAGCGTGTCTCCCCGCGGCCGCAGGCGCTGGCGGAAG-
CGTCCCTGCTG-
CGTCCCAAGAAGGAGGAGGAGGAGGAGGCGAGTCGCCGCCGCGGCAGCAGCG
GCGTGGCTTCTCTGTCCGAGCTGGGGGCGGCAGCCGCCGCGCGCCCCGGG-
TCCCTGGGCGG-
CAGCCCCCTTTCCGAGCCTGGTGGGGTCTCTGCACAGCGAGCGCACCACCCGC
CCTCGGCTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCTGCAG-
CCGGTGCGGGA-
GAAAAACCTGCCTCCCGCCTTCCCCAACAACGGGATAGAGAGCCTGGTGGAC
AAGATGAGCAGATGGAAGACCTATGCGCAGGAGCACAGGGACGCGCCTGCG-
CTCCGGCCG-
CCCACGCGGCGCCAGCGCCACGACCGGCAGCGGGGGCTGGTGTGGGATGACG
AGGACTCCGCGGACGATAGCAGCGTGCTGGACCTGGGAGGGAGCGGCAAC-
CCGTTCGCGCAC-
CTGCGCCCCCGCCTGGGGAGGATGTTTTAAAAAAAAAAAAAAAAAAGCAAGAA
GCATGATGCAAAAATTAAATAAAACTCACCAAGGCCATGGCGACCGAGCG-
TTGGTTTCTTG-
TGTTCCCTTCAGTATGCGGCGCGCGGCGATGTACCAGGAGGGACCTCCTCCC
TCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGGCGGCCCTCTT-
CTCCCTTTGCGTCGCAG-
CTGCTGGAGCCCGCGTACGTGCCTCCGCGCTACCTGCGGCCTACGGGGGGGA
GAAACAGCATCCGTTACTCGGAGCTGGCGCCCCCTGTTGACACCACCCGGG-
TGTACCTGGTG-
GACAACAAGTCGGCGGACGTGGCCTCCCTGAACTACCAGAACGACCACAGCA
ATTTTTTGACCACGGTCATCCAGAACAATGACTACAGCCCGAGCGAGGCCA-
GCACCCAGAC-
CATCAATCTGGATGACCGGTGCGCACTGGGGCGGCGACCTGAAAACCATCCTG
CACACCAACATGCCCAACGTGAACGAGTTCATGTTCCACCAATAAGTT-
CAAGGCGCGGG-
TGATGGTGTGCGCTCGCACACCAAGGAAGACCGGGTGGAGCTGAAGTACGA

GTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGAGACCATGAC-
CATTGAC-
CTGATGAACAACGCGATCGTGGAGCACTATCTGAAAGTGGGCAGGCAGAACG
GGGTCCTGGAGAGCGACATCGGGGTCAAGTTCGACACCAGGAACTTCCG-
CCTGGGGCTGGAC-
CCCGTGACCGGGCTGGTTATGCCCCGGGTGTACACCAACGAGGCCTTCCATC
CCGACATCATCCTGCTGCCCCGGCTGCGGGGTGGACTTCACTTACAGCCG-
CCTGAGCAAC-
CTCCTGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGGCTTCAGGATCACCT
ACGAGGACCTGGAGGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCC-
TACCAGGATAG-
CTTGAAGGAAAATGAGGCGGGACAGGAGGATACCGCCCCCGCCGCTCCGCC
GCCGCCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCGCGGACGGGGCA-
GAGGCCGAC-
CCCGCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGAATGACAGTG
CGGTGCGCGGAGACACCTTCGTCACCCGGGGGGAGGAAAAGCAAGCGGAGG-
CCGAGGCCG-
CGGCCGAGGAAAAGCAACTGGCGGCAGCAGCGGCGGCGGCGGCGTTGGCCGC
GGCGGAGGCTGAGTCTGAGGGGACCAAGCCCGCCAAGGAGCCCGTGAT-
TAAGCCCCTGAC-
CGAAGATAGCAAGAAGCGCAGTTACAACCTGCTCAAGGACAGCACCAACACC
GCGTACCGCAGCTGGTACCTGGCCTACAACCTACGGCGACCCGTCGA-
CGGGGGTGCGCTCCTG-
GACCCTGCTGTGCACGCCGGACGTGACCTGCGGCTCGGAGCAGGTGTACTGG
TCGCTGCCCCGACATGATGCAAGACCCCGTGACCTTCCGCTCCACGCGGCA-
GGTCAGCAACTT-
CCCGGTGGTGGGCGCCGAGCTGCTGCCCCGTGCACTCCAAGAGCTTCTACAAC
GACCAGGCCGTCTACTCCCAGCTCATCCGCCAGTTCACCTCTCTGAC-
CCACGTGTTCAATCG-
CTTTCCTGAGAACCAGATTCTGGCGCGCCCGCCCCGCCCCCACCATCACCACC

GTCAGTGAAAACGTTCTCTCACAGATCACGGGACGCTACCGCTGCG-
CAACAGCATCG-
GAGGAGTCCAGCGAGTGACCGTTACTGACGCCAGACGCCGCACCTGCCCCTA
CGTTTACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCCTTTCCAGCCGCAC-
TTTTTGAGCAA-
CACCACCATCATGTCCATCCTGATCTCACCAGCAATAACTCCGGCTGGGGA
CTGCTGCGCGCGCCCAGCAAGATGTTTCGGAGGGGCGAGGAAGCGTTCCGAG-
CAGCACCCCG-
TGCGCGTGCGCGGGCACTTCCGCGCCCCCTGGGGAGCGCACAAACGCGGCCG
CGCGGGGCGCACCAACCGTGGACGACGCCATCGACTCGGTGGTGGAGCAGG-
CGCGCAACTACA-
GGCCCGCGGTCTCTACCGTGGACGCGGCCATCCAGACCGTGGTGCGGGGCGC
GCGGCGGTACGCCAAGCTGAAGAGCCGCCGAAGCGCGTGGCCCGCCG-
CCACCGCCGCCGAC-
CCGGGGCCCGCCGCCAAACGCGCCGCGCGGCCCTGCTTCGCCGGGCCAAGCG
CACGGGCCCGCCGCGCCGCCATGAGGGCCGCGCGCCGCTTGGCCGCCGGCAT-
CACCGCCG-
CCACCATGGCCCCCGTACCCGAAGACGCGCGGCCGCCGCCGCCGCCGCCGC
CATCAGTGACATGGCCAGCAGGCGCCGGGGCAACGTGTACTGGGTGCGCGA-
CTCGGTGAC-
CGGCACGCGCGTGCCCGTGCGCTTCCGCCCCCGCGGACTTGAGATGATGTG
AAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCCAGCGGCGGCGG-
CGCGCGCAGCGT-
CATGTCCAAGCGCAAATCAAAGAAGAGATGCTCCAGGTCGTCGCGCCGGAG
ATCTATGGGCCCCCGAAGAAGGAAGAGCAGGATTGAAGCCCCGCAAGA-
TAAAGCGGGT-
CAAAAAGAAAAAGAAAGATGATGACGATGCCGATGGGGAGGTGGAGTTCCTG
CGCGCCACGGCGCCCAGGCGCCCGGTGCAGTGGAAGGGCCGGCGCGTAAAG-
CGCGTCCTGCG-
CCCCGGCACCGCGGTGGTCTTCACGCCCGGCGAGCGCTCCACCCGGACTTTC

AAGCGCGTCTATGACGAGGTGTACGGCGACGAAGACCTGCTGGAGCAGG-
CCAACGAGCGCTT-
CGGAGAGTTTGCTTACGGGAAGCGTCAGCGGGCGCTGGGGAAGGAGGACCTG
CTGGCGCTGCCGCTGGACCAGGGCAACCCACCCCCAGTCTGAAGCCCG-
TGACCCTGCAGCA-
GGTGCTGCCGAGCAGCGCACCTCCGAGGCGAAGCGGGGTCTGAAGCGCGAG
GGCGGCGACCTGGCGCCACCGTGCAGCTCATGGTGCCCAAGCGGCAGAGG-
CTGGAGGATG-
TGCTGGAGAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCAGGGTCCG
CCCCATCAAGCAGGTGGCGCCGGGCCTCGGCGTGCAGACCGTGGACGTGGT-
CATCCCCAC-
CGGCAACTCCCCCGCCGCCGCGCCACCACTACCGCTGCCTCCACGGACATGGAG
ACACAGACCGATCCCGCCGCAGCCGCAGCCGCAGCCGCCGCCGCCGCGAC-
CTCCTCGGCGGAGG-
TGCAGACGGACCCCTGGCTGCCGCCGGCGATGTCAGCTCCCCGCGCGCGTCTG
CGGGCGCAGGAAGTACGGCGCCGCCAACGCGCTCCTGCCCCGAGTACG-
CCTTGCATCCTT-
CCATCGCGCCCACCCCGGCTACCGAGGCTATACCTACCGCCCGCGAAGAGC
CAAGGGTTCCACCCGCCGTCCCCGCCGACGCGCCGCCGCCACCACCCGCCG-
CCGCCGCCGCA-
GACGCCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCGCGCGACGG
ACACACCCTGGTGCTGCCCAGGGCGCGCTACCACCCCAGCATCGTT-
TAAAAGCCTGTTGTGG-
TTCTTGCAGATATGGCCCTCACTTGCCGCCTCCGTTTCCCGGTGCCGGGATA
CCGAGGAGGAAGATCGCGCCGCAGGAGGGGTCTGGCCGGCCGCGGCCTGAG-
CGGAGGCAG-
CCGCCGCGCGCACCGGCGGCGACGCGCCACCAGCCGACGCATGCGCGGCGGG
GTGCTGCCCCCTGTTAATCCCCCTGATCGCCGCGGCGATCGGCGCCGTG-
CCCGGGATCG-
CCTCCGTGGCCTTGCAAGCGTCCCAGAGGCATTGACAGACTTGCAAACCTTGC

AAATATGGAAAAAAAAACCCCAATAAAAAAGTCTAGACTCTCACGCTCG-
CTTGGTCCTGTGA-
CTATTTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCCCGCGTCACG
GCTCGCGCCCGTTCTTGGGACACTGGAACGATATCGGCACCAGCAACA-
TGAGCGGTGGCG-
CCTTCAGTTGGGGCTCTCTGTGGAGCGGCATTAAAAGTATCGGGTCTGCCGT
TAAAAATTACGGCTCCCGGGCCTGGAACAGCAGCACGGGCCAGATG-
TTGAGAGACAAG-
TTGAAAGAGCAGAACTTCCAGCAGAAGGTGGTGGAGGGCCTGGCCTCCGGCA
TCAACGGGGTGGTGGACCTGGCCAACCAGGCCGTGCAGAATAAGATCAACA-
GCAGACTGGAC-
CCCCGGCCGCCGGTGGAGGAGGTGCCGCCGGCGCTGGAGACGGTGTCCCCCG
ATGGGCGTGGCGAGAAGCGCCCGCGGCCCGATAGGGAAGAGACCACTCTGG-
TCACGCAGAC-
CGATGAGCCGCCCCCGTATGAGGAGGCCCTGAAGCAAGGTCTGCCACCACG
CGGCCCATCGCGCCCATGGCCACCGGGGTGGTGGGCCGCCACACCCCCG-
CCACGCTGGAC-
TTGCCTCCGCCCCGCGATGTGCCGCAGCAGCAGAAGGCGGCACAGCCGGGCC
CGCCCGCGACCGCCTCCCGTTCTCCGCCGGTCCTCTGCGCCGCGCGGCCA-
GCGGCCCCCG-
CGGGGGGGTTCGCGAGGCACGGCAACTGGCAGAGCACGCTGAACAGCATCGTG
GGTCTGGGGGTGCGGTCCGTGAAGCGCCGCCGATGCTACTGAATAGCTTAG-
CTAACGTGTTG-
TATGTGTGTATGCGCCCTATGTCGCCGCCAGAGGAGCTGCTGAGTCGCCGCC
GTTTCGCGCGCCCAACCACCGCCACTCCGCCCTCAAGATGGCGAC-
CCCATCGATGATG-
CCGCAGTGGTCGTACATGCACATCTCGGGCCAGGACGCCTCGGAGTACCTGA
GCCCCGGGCTGGTGCAGTTCGCCCGCGCCACCGAGAGCTACTTCAG-
CCTGAGTAACAAGTT-
TAGGAACCCACGGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCAG

CGCCTGACGCTGCGGTTTCATTCCCGTGGACCGCGAGGACACCGCGTACTCG-
TACAAGGCG-
CGGTTACCCCTGGCCGTGGGCGACAACCGCGTGCTGGACATGGCCTCCACCT
ACTTTGACATCCGCGGGGTGCTGGACCGGGGTCCCACCTTCAAGCCCTAC-
TCTGGCACCGCC-
TACAACTCCCTGGCCCCCAAGGGCGCTCCCAACTCCTGCGAGTGGGAGCAAG
AGGAACTCAGGCAGTTGAAGAAGCAGCAGAAGAGGAAGAAGAAGATG-
CTGACGGTCAAG-
CTGAGGAAGAGCAAGCAGCTACCAAAAAGACTCATGTATATGCTCAGGCTCC
CCTTTCTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACGGACGC-
TACAGCTACA-
GAACAAAAACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCCAAATCG
GGGAGTCCCAGTGGAATGAGGCAGATGCTACAGTCGCCGGCGGTAGAGTGC-
TAAAGAAATC-
TACTCCCATGAAACCATGCTATGGTTCCTATGCAAGACCCACAAATGCTAAT
GGAGGTCAGGGTGTACTAACGGCAAATGCCAGGGACAGCTAGAATCTCA-
GGTTGAAATGCA-
ATTCTTTTCAACTTCTGAAAACGCCCCGTAACGAGGCTAACAAACATTCAGCCC
AAATTGGTGCTGTATAGTGAGGATGTGCACATGGAGACCCCGGATACGCAC-
CTTTCTTA-
CAAGCCCGCAAAAAGCGATGACAATTCAAAAATCATGCTGGGTCAGCAGTCC
ATGCCCAACAGACCTAATTACATCGGCTTCAGAGACAACCTTTATCGGCCT-
CATGTATTACAA-
TAGCACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAGTTGAATGCA
GTGGTGGACTTGCAAGACAGAAACACAGAACTGTCCTACCAGCTCTTG-
CTTGATTCCATGGG-
TGACAGAACCAGATACTTTTCCATGTGGAATCAGGCAGTGGACAGTTATGAC
CCAGATGTTAGAATTATTGAAAATCATGGAAGTGAAGACGAGCTCCCCAAC-
TATTGTTT-
CCCTCTGGGTGGCATAGGGGTAACTGACACTTACCAGGCTGTTAAAACCAAC

AATGGCAATAACGGGGGCCAGGTGACTTGGACAAAAGATGAAACTTTTGCA-
GATCGCAA-
TGAAATAGGGGTGGGAAACAATTTTCGCTATGGAGATCAACCTCAGTGCCAAC
CTGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCAGACAAG-
CTTAAGTA-
CAACCCCTCCAATGTGGACATCTCTGACAACCCCAACACCTACGATTACATG
AACAAAGCGAGTGGTGGCCCCGGGGCTGGTGGACTGCTACATCAACCTGGG-
CGCGCGCTGG-
TCGCTGGACTACATGGACAACGTCAACCCCTTCAACCACCACCGCAATGCGG
GCCTGCGCTACCGCTCCATGCTCCTGGGCAACGGGCGCTACGTGCCCTT-
CCACATCCAGGTG-
CCCCAGAAGTTCTTTGCCATCAAGAACCTCCTCCTCCTGCCGGGCTCCTACA
CCTACGAGTGGAACCTCAGGAAGGATGTCAACATGGTCCTCCAGAG-
CTCTCTGGG-
TAACGATCTCAGGGTGGACGGGGCCAGCATCAAGTTCGAGAGCATCTGCCTC
TACGCCACCTTCTTCCCCATGGCCCACAACACGGCCTCCACGCTCGAGG-
CCATGCTCAGGA-
ACGACACCAACGACCAGTCCTTCAATGACTACCTCTCCGCCGCCAACATGCT
CTACCCCATACCCGCCAACGCCACCAACGTCCCCATCTCCATCCCCTCGCG-
CAACTGGGCGG-
CCTTCCGCGGCTGGGCCTTCACCCGCCTCAAGACCAAGGAGACCCCTCCCT
GGGCTCGGGATTGACCCCTACTACACCTACTCGGGCTCCATTCCCTAC-
CTGGACGGCAC-
CTTCTACCTCAACCACACTTTCAAGAAGGTCTCGGTACCTTCGACTCCTCG
GTCAGCTGGCCGGGCAACGACCGTCTGCTCACCCEAACGAGTTCGAGAT-
CAAGCGCTCGG-
TCGACGGGGAGGGCTACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTT
CCTGGTCCAGATGCTGGCCAACATAACATCGGCTACCAGGGCTTCTACA-
TCCCAGAGAGC-
TACAAGGACAGGATGTACTCCTTCTTCAGGAACCTTCCAGCCCATGAGCCGGC

AGGTGGTGGACCAGACCAAGTACAAGGACTACCAGGAGGTGGGCAT-
CATCCACCAGCACAA-
CAACTCGGGCTTCGTGGGCTACCTCGCCCCACCATGCGCGAGGGACAGGCC
TACCCCGCCAACTTCCCCTATCCGCTCATAGGCAAGACCGCGGTTCGACAG-
CATCACCCA-
GAAAAAGTTCCTCTGCGACCGCACCTCTGGCGCATCCCCTTCTCCAGCAAC
TTCATGTCCATGGGTGCGCTCTCGGACCTGGGCCAGAACTTGCTCTACG-
CCAACTCCG-
CCCACGCCCTCGACATGACCTTCGAGGTTCGACCCCATGGACGAGCCCACCCT
TCTCTATGTTCTGTTCGAAGTCTTTGACGTGGTCCGGGTCCACCAGCCG-
CACCGCGGCGT-
CATCGAGACCGTGTACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACC
TAAAGAAGCAAGCCGCAGTCATCGCCGCCTGCATGCCGTCGGGTTCAC-
CGAGCAAGAGCT-
CAGGGCCATCGTCAGAGACCTGGGATGCGGGCCCTATTTTTTGGGCACCTTC
GACAAGCGCTTCCCTGGCTTTGTCTCCCCACACAAGCTGGCCTGCG-
CCATCGTCAACACGG-
CCGGCCGCGAGACCGGGGGCGTGCACTGGCTGGCCTTCGCCTGGAACCCGCG
CTCCAAAACATGCTTCCTCTTTGACCCCTTCGGCTTTTCGGACCAGCGGCT-
CAAGCAAATC-
TACGAGTTCGAGTACGAGGGCTTGCTGCGTCGCAGCGCCATCGCCTCCTCGC
CCGACCGCTGCGTCACCCTCGAAAAGTCCACCCAGACCGTGCAGGGG-
CCCGACTCGGCCG-
CCTGCGGTCTCTTCTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCA
GAGTCCCATGGACCGCAACCCCACCATGAACTTGCTGACGGGGGTG-
CCCAACTCCATG-
CTCCAGAGCCCCCAGGTTCGAGCCCACCCTGCGCCGCAACCAGGAGCAGCTCT
ACAGCTTCCTGGAGCGCCACTCGCCTTACTTCCGCCGCCACAGCGCACAGA-
TCAGGAGGG-
CCACCTCCTTCTGCCACTTGCAAGAGATGCAAGAAGGGTAATAACGATGTAC

ACACTTTTTTCTCAATAAATGGCATCTTTTTATTTATACAAG-
CTCTCTGGGGTATTCATTT-
CCCACCACCACCCGCCGTTGTCGCCATCTGGCTCTATTTAGAAATCGAAAGG
GTTCTGCCGGGAGTCGCCGTGCGCCACGGGCAGGGACACGTTGCGATAC-
TGGTAGCGGGTG-
CCCCACTTGAACTCGGGCACCACCAGGCGAGGCAGCTCGGGGAAGTTTTTCGC
TCCACAGGCTGCGGGTCAGCACCAGCGCGTTCATCAGGTCGGGCGCCGAGA-
TCTTGAAGTCG-
CAGTTGGGGCCGCCGCCCTGCGCGCGCAGTTGCGGTACACCGGGTTGCAGC
ACTGGAACACCAACAGCGCCGGGTGCTTACGCTGGCCAGCACGCTGCGG-
TCGGAGATCAG-
CTCGGCGTCCAGGTCCTCCGCGTTGCTCAGCGCGAACGGGGTCATCTTGGGC
ACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGTTTCGAGTTGCAGTCGCAG-
CGCAGCGGGAT-
CAGCAGGTGCCCCGTGCCCGGACTCGGCGTTGGGGTACAGCGCGCGCATGAAG
GCCTGCATCTGGCGGAAGGCCATCTGGGCCTTGGCGCCCTCCGAGAAGAA-
CATGCCGCAGGA-
CTTGCCCGAGAACTGGTTTGCGGGGCAGCTGGCGTCGTGCAGGCAGCAGCGC
GCGTCGGTGTTGGCGATCTGCACCACGTTGCGCCCCCACCAGTTCTT-
CACGATCTTGG-
CCTTGGACGATTGCTCCTTCAGCGCGCGCTGCCCGTTCTCGCTGGTCACATC
CATCTCGATCACATGTTCTTGTTCACCATGCTGCTGCCGTGCAGACACTT-
CAGCTCG-
CCCTCCGTCTCGGTGCAGCGGTGCTGCCACAGCGCGCAGCCCGTGGGCTCGA
AAGACTTGTAGGTCACCTCCGCGAAGGACTGCAGGTACCCCTGCAAAAAG-
CGGCCCAT-
CATGGTCACGAAGGTCTTGTTGCTGCTGAAGGTCAGCTGCAGCCCGCGGTGC
TCCTCGTTCAGCCAGGTCTTGACACGCGCCGAGCGCCTCCACCTGG-
TCGGGCAG-
CATCTTGAAGTTCACCTTCAGCTCATTCTCCACGTGGTACTTGTCCATCAGC

GTGCGCGCCGCCTCCATGCCCTTCTCCCAGGCCGACACCAGCGGCAGGCT-
CACGGGGTTCTT-
CACCATCACCGTGGCCGCCGCCTCCGCCGCGCTTTCGCTTTCGCCCCCGCTG
TTCTCTTCCTCTTCCTCCTCTTCCTCGCCGCCGCCCACTCGCAGCCCCCG-
CACCACGGGG-
TCGTCTTCCTGCAGGCGCTGCACCTTGCGCTTGCCGTTGCGCCCCTGCTTGA
TGCGCACGGGCGGGTTGCTGAAGCCCACCATCACAGCGCGGCCTCTT-
CTTGCTCGTCCTCG-
CTGTCCAGAATGACCTCCGGGGAGGGGGGGTTGGTCATCCTCAGTACCGAGG
CACGCTTCTTTTTCTTCCTGGGGGCGTTCGCCAGCTCCGCGGCTGCGGCCG-
CTGCCGAGG-
TCGAAGGCCGAGGGCTGGGCGTGCGCGGCACCAGCGCGTCCTGCGAGCCGTC
CTCGTCCTCCTCGGACTCGAGACGGAGGCGGGCCCGCTTCTTCGGGGGCG-
CGCGGGGCGGCG-
GAGGCGGCGGCGGCGACGGAGACGGGGACGAGACATCGTCCAGGGTGGGTGG
ACGGCGGGCCGCGCCGCGTCCGCGCTCGGGGGTGGTCTCGCGCTGG-
TCCTCTTCCCGACTGG-
CCATCTCCCAGTCTCCTTCTCCTATAGGCAGAAAGAGATCATGGAGTCTCT
CATGCGAGTCGAGAAGGAGGAGGACAGCCTAACCGCCCCCTCTGAG-
CCCTCCACCACCGCCG-
CCACCACCGCCAATGCCGCCGCGGACGACGCGCCACCGAGACCACCGCCAG
TACCACCCTCCCCAGCGACGCACCCCCGCTCGAGAATGAAGTG-
CTGATCGAGCAGGAC-
CCGGGTTTTGTGAGCGGAGAGGAGGATGAGGTGGATGAGAAGGAGAAGGAGG
AGGTCGCCGCCTCAGTGCCAAAAGAGGATAAAAAGCAAGACCAGGACGACG-
CAGATAAGGA-
TGAGACAGCAGTCGGGCGGGGGAACGGAAGCCATGATGCTGATGACGGCTAC
CTAGACGTGGGAGACGACGTGCTGCTTAAGCACCTGCACCGCCAGTGCGT-
CATCGTCTGCGA-
CGCGCTGCAGGAGCGCTGCGAAGTGCCCCCTGGACGTGGCGGAGGTCAGCCGC

GCCTACGAGCGGCACCTCTTCGCGCCGCACGTGCCCCCAAGCGCCGGGA-
GAACGGCACCTG-
CGAGCCCAACCCGCGTCTCAACTTCTACCCGGTCTTCGCGGTACCCGAGGTG
CTGGCCACCTACCACATCTTTTTTCCAAAAGTCAAGATCCCCCTCTCCTG-
CCGCGCCAACCG-
CACCCGCGCCGACAAAACCTGACCCTGCGGCAGGGCGCCACATACCTGAT
ATCGCCTCTCTGGAGGAAGTGCCCAAGATCTTCGAGGGTCTCGGTGCGGA-
CGAGAAACGGG-
CGGCGAACGCTCTGCACGGAGACAGCGAAAACGAGAGTCACTCGGGGGTGCT
GGTGGAGCTCGAGGGCGACAACGCGCGCCTGGCCGTACTCAAGCGCAGCA-
TAGAGGTCAC-
CCACTTTGCCTACCCGGCGCTCAACCTGCCCCCAAGGTCATGAGTGTGGTC
ATGGGCGAGCTCATCATGCGCCGCGCCCAGCCCCTGGCCGCGGATGCAAAC-
TTGCAAGAG-
TCCTCCGAGGAAGGCCTGCCCCGCGGTCAGCGACGAGCAGCTGGCGCGCTGGC
TGGAGACCCGCGACCCCGCGCAGCTGGAGGAGCGGCGCAAGCTCATGATGG-
CCGCGGTG-
CTGGTCACCGTGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCGAGA
TGCAGCGCAAGCTCGAGGAGACCCTGCACTACACCTTCCGCCAGGGCTACG-
TGCGCCAGG-
CCTGCAAGATCTCCAACGTGGAGCTCTGCAACCTGGTCTCCTACCTGGGCAT
CCTGCACGAGAACCGCCTCGGGCAGAACGTCCTGCACTCCACCCT-
CAAAGGGGAGGCGCG-
CCGCGACTACATCCGCGACTGCGCCTACCTCTTCCTCTGCTACACCTGGCAG
ACGGCCATGGGGGTCTGGCAGCAGTGCCTGGAGGAGCGCAACCTCAAGGAG-
CTGGAAAAG-
CTCCTCAAGCGCACCCCTCAGGGACCTCTGGACGGGCTTCAACGAGCGCTCGG
TGGCCGCCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCAAGAC-
CCTGCAGCAGGG-
CCTGCCCCGACTTCACCAGCCAGAGCATGCTGCAGAACTTCAGGACTTTCATC

CTGGAGCGCTCGGGCATCCTGCCGGCCACTTGCTGCGCGCTGCCCAGCGAC-
TTCGTGCCCAT-
CAAGTACAGGGAGTGCCCGCCGCGCTCTGGGGCCACTGCTACCTCTTCCAG
CTGGCCAACTACCTCGCCTACCACTCGGACCTCATGGAAGACGTGAGCGG-
CGAGGGCCTG-
CTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCCCACCGCTCTCTAGTCT
GCAACCCGCAGCTGCTCAGCGAGAGTCAGATTATCGGTACCTTCGAGCTG-
CAGGGTCCCTCG-
CCTGACGAGAAGTCCGCGGCTCCAGGGCTGAAACTCACTCCGGGGCTGTGGA
CTTCCGCCTACCTACGCAAATTTGTACCTGAGGACTACCACGCCCACGAGA-
TCAGGTTC-
TACGAAGACCAATCCCGCCCCGCCAAGGCGGAGCTCACCGCCTGCGTCATCA
CCCAGGGGCGACATCCTGGGCCAATTGCAAGCCATCAACAAAGCCCG-
CCGAGAGTTCTTG-
CTGAAAAAGGGTCGGGGGGTGTACCTGGACCCCCAGTCCGGCGAGGAGCTAA
ACCCGCTACCCCCGCCGCGCCCGCCCGCCAGCAGCGGGACCTTGCTTCCCAGGA-
TGGCACCCA-
GAAAGAAGCAGCAGCCGCGCCGCGCCGCGCAGCCATACATGCTTCTGGAGGA
AGAGGAGGAGGACTGGGACAGTCAGGCAGAGGAGGTTTCGGACGAGGAGCA-
GGAGGAGA-
TGATGGAAGACTGGGAGGAGGACAGCAGCCTAGACGAGGAAGCTTCAGAGGC
CGAAGAGGTGGCAGACGCAACACCATCGCCCTCGGTGCGAGCCCCCTCG-
CCGGGG-
CCCCTGAAATCCTCCGAACCCAGCACCGAGCGCTATAACCTCCGCTCCTCCGG
CGCCGGCGCCACCCGCCCCGCGAGACCCAACCGTAGATGGGACACCACAGGA-
ACCGGGGTCGG-
TAAGTCCAAGTGCCCGCCGCGCCGCGCCAGCAGCAGCAGCAGCAGCGCCAG
GGCTACCGCTCGTGGCGCGGGCACAAGAACGCCATAGTCGCCTGCTTGCAA-
GACTGCGGGGG-
CAACATCTCTTTGCCCCGCGCTTCCTGCTATTCCACCACGGGGTCGCCTTT

CCCCGCAATGTCCTGCATTACTACCGTCATCTCTACAGCCCCTACTGCAG-
CGGCGACCCA-
GAGGCGGCAGCGGCAGCCACAGCGGCGACCACCACCTAGGAAGATATCCTCC
GCGGGCAAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCAGCGG-
CGGGAGCGGTGGG-
CGCACTGCGCCTCTCGCCCAACGAACCCCTCTCGACCCGGGAGCTCAGACAC
AGGATCTTCCCCACTTTGTATGCCATCTTCCAACAGAGCAGAGGCCAGGAG-
CAGGAG-
CTGAAAATAAAAAACAGATCTCTGCGCTCCCTCACCCGCAGCTGTCTGTATC
ACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAGGACGCGGAGGCAC-
TCTTCAGCAAATAC-
TGCGCGCTCACTCTTAAAGACTAGCTCCGCGCCCTTCTCGAATTTAGGCGGG
AGAAAACACTACGTCATCGCCGGCCGCCGCCAGCCCGCCCAGCCGAGATGAG-
CAAAGAGATT-
CCCACGCCATACATGTGGAGCTACCAGCCGCAGATGGGACTCGCGGCGGGAG
CGGCCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGGGACCCACAC-
TGATCTCACA-
GGTCAACGGGATCCGCGCCCAGCGAAACCAAATACTGCTGGAACAGGCGGCC
ATCACCGCCACGCCCCGCCATAATCTCAACCCCCGAAATTGGCCCGCCG-
CCCTCGTGTACCA-
GGAAACCCCCTCCGCCACCACCGTACTACTTCCGCGTGACGCCAGGCCGAA
GTCCAGATGACTAACTCAGGGGCGCAGCTCGCGGGCGGCTTTCGT-
CACGGGGCGCGGCCG-
CTCCGACCAGGTATAAGACACCTGATGATCAGAGGCCGAGGTATCCAGCTCA
ACGACGAGTCGGTGAGCTCTTCGCTCGGTCTCCGTCCGGACGGAACCTTT-
CCAGCTCGCCGGA-
TCCGGCCGCTCTTCGTTACGCCCCGCCAGGCGTACCTGACTCTGCAGACCT
CGTCCTCGGAGCCCCGCTCCGGCGGCATCGGAACCCCTCCAGTTCGTGGAG-
GAGTTCGTG-
CCCTCGGTCTACTTCAACCCCTTCTCGGGACCTCCCGGACGCTACCCCGACC

AGTTCATTCCGAACTTTGACGCGGTGAAGGACTCGGCGGACGGCTACGAC-
TGAATGTCAGG-
TGTCGAGGCAGAGCAGCTTCGCCTGAGACACCTCGAGCACTGCCGCCGCCAC
AAGTGCTTCGCCCCGCGTTCTGGTGAGTTCTGCTACTTTCAGCTACCCGAG-
GAGCATAC-
CGAGGGGCGCGGCACGGCGTCCGCCTGACCACCCAGGGCGAGGTTACCTGT
TCCCTCATCCGGGAGTTTACCCTCCGTCCCCTGCTAGTGAGCGGGAG-
CGGGGTCCCTGTG-
TCCTAACTATCGCCTGCAACTGCCCTAACCTGGATTACATCAAGATCTTTG
CTGTCACTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAATCTAC-
TGGGGCTCCTGTG-
CCATCCTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAGGCGAAC
CTCACCTGCGGTCTGCATCGGAGGGCCAAGAAGTACCTCACCTGGTACTT-
CAACGGCAC-
CCCCTTTGTGGTTTACAACAGCTTCGACGGGGACGGAGTCTCCCTGAAAGAC
CAGCTCTCCGGTCTCAGCTACTCCATCCACAAGAACACCACCCTCCAAC-
TCTTCCCTCCC-
TACCTGCCGGGAACCTACGAGTGCGTCACCGGCCGCTGCACCCACCTCACCC
GCCTGATCGTAAACCAGAGCTTTCGGGAACAGATAACTCCCTCTTCCCCA-
GAACAGGAGG-
TGAGCTCAGGAACTCCCCGGGGACCAGGGCGGAGACGTACCTTCGACCCTT
GTGGGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTTAATCAAAG-
TTTCCTTGAGA-
TTTGTTCTTTCCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACC
CTTTCTTCGGAATCAGGTGACTTCTCTGAAATCGGGCTTGGTGTGCTGCT-
TACTCTG-
TTGATTTTTTTCCTTATCATACTCAGCCTTCTGTGCCTCAGGCTCGCCGCCT
GCTGCGCACACATCTATATCTACTGCTGGTTGCTCAAGTGCAGGGGTCG-
CCACCCAAGA-
TGAACAGGTACATGGTCCTATCGATCCTAGGCCTGCTGGCCCTGGCGGCCTG

CAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCTTGCAATGTAAC-
TTTCAAG-
CCCGAGGGTGACCAATGCACCACCCTCGTCAAATGCGTTACCAATCATGAGA
GGCTGCGCATCGACTACAAAAACAAAACCTGGCCAGTTTGCGGTCTATAGTG-
TGTTTACG-
CCCGGAGACCCCTCTAACTACTCTGTCACCGTCTTCCAGGGCGGACAGTCTA
AGATATTCAATTACACTTTCCCTTTTTATGAGTTATGCGATGCGGTCATG-
TACATGTCAAAA-
CAGTACAACCTGTGGCCTCCCTCTCCCCAGGCGTGTGTGGAAAATACTGGGT
CTTACTGCTGTATGGCTTTCGCAATCACTACGCTCGCTCTAATCTGCACGG-
TGCTATACA-
TAAAATTCAGGCAGAGGGCGAATCTTTATCGATGAAAAGAAAATGCCTTGATC
GCTAACACCGGCTTTCTATCTGCAGAATGAATGCAATCACCTCCCTACTAA-
TCACCACCAC-
CCTCCTTGCGATTGCCCATGGGTGACACGAATCGAAGTGCCAGTGGGGTCC
AATGTCACCATGGTGGGCCCCGCCGGCAATTCCACCCTCATGTGGGAAAAA-
TTTGTCCGCAA-
TCAATGGGTTCATTTCTGCTCTAACCGAATCAGTATCAAGCCCAGAGCCATC
TGCGATGGGCAAAATCTAACTCTGATCAATGTGCAAATGATGGATGCTGGG-
TACTAT-
TACGGGCAGCGGGGAGAAATCATTAATTACTGGCGACCCCACAAGGACTACA
TGCTGCATGTAGTCGAGGCACTTCCCCTACTACCACCCCCACTACCAC-
CTCTCCCACCACCAC-
CACCCTACTACTACTACTACTACTACTACTACTACTACCCTACCGCTGCC
CGCCATACCCGCAAAAGCACCATGATTAGCACAAAGCCCCCTCGTGCTCAC-
TCCCACGCCGG-
CGGGCCCATCGGTGCGACCTCAGAAACCAACCGAGCTTTGCTTCTGCCAATGC
ACTAACGCCAGCGCTCATGAACTGTTGACCTGGAGAATGAGGATGTCCAG-
CAGAGCTCCG-
CTTGCCCTGACCCAGGAGGCTGTGGAGCCCGTTGCCCTGAAGCAGATCGGTGA

TTCAATAATTGACTCTTCTTCTTTTGCCACTCCCGAATACCCTCCCGATTCT-
TACTTTCCACA-
TCACGGGTACCAAAGACCCTAACCTCTCTTTCTACCTGATGCTGCTGCTCTG
TATCTCTGTGGTCTCTTCCGCGCTGATGTTACTGGGGATGTTCTGCTG-
CCTGATCTGCCGCA-
GAAAGAGAAAAGCTCGCTCTCAGGGCCAACCACTGATGCCCTTCCCCTACCC
CCCGGATTTTGCAGATAACAAGATATGAGCTCGCTGCTGACACTAACCG-
CTTTACTAGCCTG-
CGCTCTAACCTTGTCGCTTGCGACTCGAGATTCCACAATGTCACAGCTGTG
GCAGGAGAAAATGTTACTTTCAACTCCACGGCCGATACCCAGTGGTCGTG-
GAGTGGCTCAGG-
TAGCTACTTAACTATCTGCAATAGCTCCACTTCCCCCGGCATATCCCCAACC
AAGTACCAATGCAATGCCAGCCTGTTACCCCTCATCAACGCTTCCACCCTG-
GACAATGGAC-
TCTATGTAGGCTATGTACCCTTTGGTGGGCAAGGAAAGACCCACGCTTACAA
CCTGGAAGTTCGCCAGCCCAGAACCACTACCCAAGCTTCTCCCACCACCAC-
CACCACCAC-
CACCATCACCAGCAGCAGCAGCAGCAGCAGCCACAGCAGCAGCAGCAGATTA
TTGACTTTGGTTTTGGCCAGCTCATCTGCCGCTACCCAGGCCATCTACAG-
CTCTGTG-
CCCGAAACCACTCAGATCCACCGCCCAGAAACGACCACCGCCACCACCCTAC
ACACCTCCAGCGATCAGATGCCGACCAACATCACCCCCTTGGCTCTTCAAA-
TGGGACTTA-
CAAGCCCCACTCCAAAACCAGTGGATGCGGCCGAGGTCTCCGCCCTCGTCAA
TGACTGGGCGGGGCTGGGAATGTGGTGGTTCGCCATAGGCATGATGGCG-
CTCTGCCTGCTT-
CTGCTCTGGCTCATCTGCTGCCTCCACCGCAGGCGAGCCAGACCCCCCATCT
ATAGACCCATCATTGTCCTGAACCCCGATAATGATGGGATCCATAGATTG-
GATGG-
CCTGAAAAACCTACTTTTTTTCTTTTACAGTATGATAAATTGAGACATGCCTC

GCATTTTCTTGTACATGTTTCCTTCTCCACCTTTTCTGGGGTGTTCCTACG-
CTGGCCGCTGTG-
TCTCACCTGGAGGTAGACTGCCTCTCACCCCTTCACTGTCTACCTGCTTTACG
GATTGGTCACCCTCACTCTCATCTGCAGCCTAATCACAGTAATCATCG-
CCTTCATCCAGTG-
CATTGATTACATCTGTGTGCGCCTCGCATACTTCAGACACCACCCGCAGTAC
CGAGACAGGAACATTGCCCAACTTCTAAGACTGCTCTAATCATGCATAAGA-
CTGTGATCTG-
CCTTCTGATCCTCTGCATCCTGCCCACCCTCACCTCCTGCCAGTACACCACA
AAATCTCCGCGCAAAAGACATGCCTCCTGCCGCTTCACCCAACTGTGGAA-
TATACCCAAATG-
CTACAACGAAAAGAGCGAGCTCTCCGAAGCTTGGCTGTATGGGGTCATCTGT
GTCTTAGTTTTCTGCAGCACTGTCTTTGCCCTCATAATCTACCCCTAC-
TTTGATTTGGGATG-
GAACGCGATCGATGCCATGAATTACCCACCTTTCCCGCACCCGAGATAATT
CCACTGCGACAAGTTGTACCCGTTGTCTGTTAATCAACGCCCCCATCCCC-
TACGCCCCAC-
TGAAATCAGCTACTTTAACCTAACAGGCGGAGATGACTGACGCCCTAGATCT
AGAAATGGACGGCATCAGTACCGAGCAGCGTCTCCTAGAGAGGCGCAGGCA-
GGCGGCTGAG-
CAAGAGCGCCTCAATCAGGAGCTCCGAGATCTCGTTAACCTGCACCAGTGCA
AAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCACCTACGAGAAGA-
CCGGCAACAG-
CCACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCTGGTGCTCATG
GTGGGTGAGAATCCCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGG-
TGTCTGCAC-
TCCCCCTGTCGGGGTCCAGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCG
GTCTCAGAGATTTAGTCCCCTTTAACTAATCAAACACTGGAATCAA-
TAAAAAGAATCACT-
TACTTAAAATCAGACAGCAGGTCTCTGTCCAGTTTATTCAGCAGCACCTCCT

TCCCCTCCTCCCAACTCTGGTACTCCAAACGCCTTCTGGCGGCAAACCTT-
CCTCCACAC-
CCTGAAGGGAATGTCAGATTCTTGCTCCTGTCCCTCCGCACCCACTATCTTC
ATGTTGTTGCAGATGAAGCGCACCAAAACGTCTGACGAGAGCTTCAAC-
CCCGTGTACCCCTA-
TGACACGGAAAGCGGCCCTCCCTCCGTCCCTTTCCCTCACCCCTCCCTTCGTG
TCTCCCGATGGATTCCAAGAAAGTCCCCCGGGGTCCTGTCTCTGAAC-
CTGGCCGAG-
CCCCTGGTCACTTCCCACGGCATGCTCGCCCTGAAAATGGGAAGTGGCCTCT
CCCTGGACGACGCTGGCAACCTCACCTCTCAAGATATCACCACCGCTAG-
CCCTCCCCT-
CAAAAAACCAAGACCAACCTCAGCCTAGAAACCTCATCCCCCTAACTGTG
AGCACCTCAGGCGCCCTCACCGTAGCAGCCGCCGCTCCCCTGGCGGTGG-
CCGGCACCTCCCT-
CACCATGCAATCAGAGGCCCCCTGACAGTACAGGATGCAAACTCACCCCTG
GCCACCAAAGGCCCCCTGACCGTGTCTGAAGGCAAACCTGGCCTTGCAAACA-
TCGGCCCCG-
CTGACGGCCGCTGACAGCAGCACCCCTCACAGTCAGTGCCACACCACCCCTTA
GCACAAGCAATGGCAGCTTGGGTATTGACATGCAAGCCCCCATTTACAC-
CACCAATGGA-
AAACTAGGACTTAACTTTGGCGCTCCCCTGCATGTGGTAGACAGCCTAAATG
CACTGACTGTAGTTACTGGCCAAGGTCTTACGATAAACGGAACAGCCCTA-
CAAACCTAGAG-
TCTCAGGTGCCCTCAACTATGACACATCAGGAAACCTAGAATTGAGAGCTGC
AGGGGGTATGCGAGTTGATGCAAATGGTCAACTTATCCTTGATGTAGCT-
TACCCATTTGATG-
CACAAAACAATCTCAGCCTTAGGCTTGGACAGGGACCCCTGTTTGTTAACTC
TGCCCACAACTTGGATGTAACTACAACAGAGGCCTCTACCTGTTTACA-
TCTGGAAATAC-
CAAAAAGCTAGAAGTTAATATCAAAACAGCCAAGGGTCTCATTTATGATGAC

ACTGCTATAGCAATCAATGCGGGTGATGGGCTACAGTTTGACTCAGGCTCA-
GATACAAA-
TCCATTAAAACTAACTTGGATTAGGACTGGATTATGACTCCAGCAGAGCC
ATAATTGCTAACTGGGAAGTGGCCTAAGCTTTGACAACACAGGTGCCAT-
CACAGTAGGCAA-
CAAAAATGATGACAAGCTTACCTTGTGGACCACACCAGACCCATCCCCTAAC
TGTTAGAATCTATTCAGAGAAAGATGCTAAATTCACACTTGTTTTGACTAAA-
TGCGGCAGTCA-
GGTGTGGCCAGCGTTTCTGTTTTATCTGTAAAAGGTAGCCTTGCGCCCATC
AGTGGCACAGTAACTAGTGCTCAGATTGTCCTCAGATTTGATGAAAATGGA-
GTTCTACTAAG-
CAATTCTTCCCTTGACCCTCAATACTGGAAGTACAGAAAAGGTGACCTTACA
GAGGGCACTGCATATACCAACGCAGTGGGATTTATGCCCAACCTCACAGCA-
TACCCAAAAA-
CACAGAGCCAACTGCTAAAAGCAACATTGTAAGTCAGGTTTACTTGAATGG
GGACAAATCCAAACCCATGACCCTCACCATTACCCTCAATGGAAGTAA-
TGAAACAGGAGATG-
CCACAGTAAGCACTTACTCCATGTCATTCTCATGGAAGTGAATGGAAGTAA
TTACATTAATGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCG-
CCCAAGAA-
TAAAAAGCATGACGCTGTTGATTTGATTCAATGTGTTTCTGTTTTATTTTCA
AGCACAACAAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACACA-
GTAGCTTAATA-
GACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAGATCAGACAGTGATAATT
AACCACCACCACCACCATACCTTTTGATTGAGGAAATCATGATCATCACAG-
GATCCTAG-
TCTTCAGGCCGCCCCCTCCCTCCCAAGACACAGAATACACAGTCCTCTCCCC
CCGACTGGCTTTAAATAACACCATCTGGTTGGTCACAGACATGTTCT-
TAGGGGTATATT-
CCACACGGTCTCCTGCCGCGCCAGGCGCTCGTCGGTGATGTTGATAAACTCT

CCCGGCAGCTCGCTCAAGTTCACGTCGCTGTCCAGCGGCTGAACCTCCGG-
CTGACGCGA-
TAACTGTGCGACCGGCTGCTGGACGAACGGAGGCCGCGCCTACAAGGGGGTA
GAGTCATAATCCTCGGTCAGGATAGGGCGGTGATGCAGCAGCAGCGAG-
CGAAACATCTGCTG-
CCGCCGCCGCTCCGTCCGGCAGGAAAACAACACGCCGGTGGTCTCCTCCGCG
ATAATCCGCACCGCCCGCAGCATCAGCTTCCTCGTTCTCCGCGCGCAGCAC-
CTCACCCCTTA-
TCTCGCTCAAATCGGCGCAGTAGGTACAGCACAGCACACCACGATGTTATTCAT
GATCCCACAGTGCAGGGCGCTGTATCCAAAGCTCATGCCGGGAACCACCG-
CCCCACGTGG-
CCATCGTACCACAAGCGCACGTAAATCAAGTGTCGACCCCTCATGAACGCGC
TGGACACAAACATTACTTCCTTGGGCATGTTGTAATTCACCACCTCCCGG-
TACCAGATAAAC-
CTCTGGTTGAACAGGGCACCTTCCACCACCATCCTGAACCAAGAGGCCAGAA
CCTGCCCACCGGCTATGCACTGCAGGGAACCCGGGTGGAACAATGACAA-
TGCAGAC-
TCCAAGGCTCGTAACCGTGGATCATCCGGCTGCTGAAGGCATCGATGTTGGC
ACAACACAGACACACGTGCATGCACTTTCTCATGATTAGCAGCTCTT-
CCCTCGTCAGGATCA-
TATCCCAAGGAATAACCCATTCTTGAATCAACGTAAAACCCACACAGCAGGG
AAGGCCTCGCACATAACTCACGTTGTGCATGGTCAGCGTGTTGCATTCCG-
GAAACAGCGGA-
TGATCCTCCAGTATCGAGGCGCGGGTCTCCTTCTCACAGGGAGGTAAAGGGT
CCCTGCTGTACGGACTGCGCCGGGACGACCGAGATCGTGTTGAGCGTAGTG-
TCATGGA-
AAAGGGAACGCCGGACGTGGTCATACTTCTTGAAGCAGAACCAGGTTTCGCGC
GTGGCAGGCCTCCTTGCGTCTGCGGTCTCGCCGTCTAGCTCGCTCCGTG-
TGATAGTTGTAG-
TACAGCCACTCCCGCAGAGCGTCGAGGCGCACCCCTGGCTTCCGGATCTATGT

AGACTCCGTCTTGCACCGCGGCCCTGATAATATCCACCACCGTAGAATAAG-
CAACACCCAG-
CCAAGCAATACACTCGCTCTGCGAGCGGCAGACAGGAGGAGCGGGCAGAGAT
GGGAGAACCATGATAAAAACTTTTTTTAAAGAATATTTTCCAATTCTT-
CGAAAGTAAGATC-
TATCAAGTGGCAGCGCTCCCCTCCACTGGCGCGGTCAAACCTCTACGGCCAAA
GCACAGACAACGGCATTCTAAGATGTTCTTAATGGCGTCCAAAAGACA-
CACCGCTCT-
CAAGTTGCAGTAAACTATGAATGAAAACCCATCCGGCTGATTTTCCAATATA
GACGCGCCGGCAGCGTCCACCAAACCCAGATAATTTTCTTCTCTCCAGCGG-
TTTACGATCTG-
TCTAAGCAAATCCCTTATATCAAGTCCGACCATGCCAAAAATCTGCTCAAGA
GCGCCCTCCACCTTCATGTACAAGCAGCGCATCATGATTGCAAAAATTCA-
GGTTCTTCAGA-
GACCTGTATAAGATTCAAAATGGGAACATTAACAAAAATTCCTCTGTGCGCGC
AGATCCCTTCGCGAGGGCAAGCTGAACATAATCAGACAGGTCCGAACGGAC-
CAGTGAGGCCAA-
ATCCCCACCAGGAACCAGATCCAGAGACCCTATACTGATTATGACGCGCATA
CTCGGGGCTATGCTGACCAGCGTAGCGCCGATGTAGGCGTGCTGCATGGG-
CGGCGAGATAAA-
ATGCAAAGTGCTGGTTAAAAAATCAGGCAAAGCCTCGCGCAAAAAAGCTAAC
ACATCATAATCATGCTCATGCAGGTAGTTGCAGGTAAGCTCAGGAAC-
CAAAACGGAATAACA-
CACGATTTTCTCTCAAACATGACTTCGCGGATACTGCGTAAAACAAAAAAT
TATAAATAAAAAATTAATTAAATAACTTAAACATTGGAAGCCTGTCTCA-
CAACAGGAAAAAC-
CACTTTAATCAACATAAGACGGGCCACGGGCATGCCGGCATAGCCGTAAAAA
AATTGGTCCCCGTGATTAACAAGTACCACAGACAGCTCCCCGGTCATG-
TCGGGGGTCAT-
CATGTGAGACTCTGTATACACGTCTGGATTGTGAACATCAGACAAACAAAGA

AATCGAGCCACGTAGCCCGGAGGTATAATCACCCGCAGGCGGAGGTACAG-
CAAAACGAC-
CCCCATAGGAGGAATCACAAAATTAGTAGGAGAAAAAATACATAAACACCA
GAAAAACCTGTTGCTGAGGCAAAATAGCGCCCTCCCGATCCAAAACAACA-
TAAAGCGCTT-
CCACAGGAGCAGCCATAACAAAGACCCGAGTCTTACCAGTAAAAGAAAAAG
ATCTCTCAACGCAGCACCAGCACCAACACTTCGCAGTGTAAGGCCAAG-
TGCCGAGAGAG-
TATATATAGGAATAAAAAGTGACGTAAACGGGCAAAGTCCAAAAACGCCCA
GAAAAACCGCACGCGAACCTACGCCCCGAAACGAAAGCCAAAAACACTA-
GACACTCCCTT-
CCGGCGTCAACTTCCGCTTTCCCACGCTACGTCACTTCCCCCGGTCAAACAA
ACTACATATCCCGAACTTCCAAGTCGCCACGCCCAAAACACCGCCTACAC-
CTCCCCGCCCG-
CCGGCCCCGCCCCGGACCCGCCTCCCGCCCCGCGCCGCCCATCTCATTATCA
TATTGGCTTCAATCCAAAATAAGGTATATTATTGATGATG

SEQ ID NO: 2 Sequência de polinucleotídeos que codifica
ChAd83 selvagem

CATCATCAATAATATACCTCAAACCTTTTGGTGCGCGTTAATATGCAAA-
TGAGCTGTTTGAA-
TTTGGGGATGCGGGGCGCTGATTGGCTGCGGGAGCGGCGACCGTTAGGGGCG
GGGCGGGTGACGTTTTGATGACGTGGCCGTGAGGCGGAGCCGGTTTGCAAG-
TTCTCGTGGA-
AAAGTGACGTCAAACGAGGTGTGGTTTGAACACGGAAATACTCAATTTTCCC
GCGCTCTCTGACAGGAAATGAGGTGTTTCTGGGCGGATGCAAG-
TGAAAACGGGCCATTTTCG-
CGCGAAAACCTGAATGAGGAAGTGAAAATCTGAGTAATTTGCGGTTTATGGCA
GGGAGGAGTATTTGCCGAGGGCCGAGTAGACTTTGACCGATTACGTGGGGG-
TTTCGATTAC-
CGTATTTTTTCACCTAAATTTCCGCGTACGGTGTCAAAGTCCGGTGTTTTTAC

GTAGGCGTCAGCTGATCGCCAGGGTATTTAAACCTGCGCTCACTAGT-
CAAGAGGCCAC-
TCTTGAGTGCCAGCGAGTAGAGTTTTCTCCTCCGCGCCGCGAGTCAGATCTA
CACTTTGAAAGATGAGGCACTTGAGAGACCTGCCCGGTAATGTTTTCTTGG-
CTACTGGGA-
ACGAGATTCTGGAATTGGTGGTGGACGCCATGATGGGTGACGACCCTCCCGA
GCCCCCTACCCCATTTGAGGCGCCTTCGCTGTACGATTTGTATGATCTGGA-
GGTGGATGTG-
CCCGAGAACGACCCCAACGAGGAGGCGGTGAATGATTTGTTTAGCGATGCCG
CGCTGCTGGCTGCCGAGCAGGCTAATACGGACTTTGGCTCAGACAGCGATT-
CTTCTCTCCA-
TACCCCGAGACCCGGCAGAGGTGAGAAAAAGATCCCGAGCTTAAAGGGGAA
GAGCTCGACCTGCGCTGCTATGAGGAATGCTTGCCTCCGAGCGATGATGAG-
GAGGACGAGGA-
GGCGATTCGAGCTGCAGCGAACCAGGGAGTGAAAGCTGCGGGCGAAAGCTTT
AGCCTGGACTGTCCTACTCTGCCCGGACACGGCTGTAAGTCTTGTGAATTT-
CATCGCATGAA-
TACTGGAGATAAGAATGTGATGTGTGCCCTGTGCTATATGAGAGCTTACAAC
CATTGTGTTTACAGTAAGTGTGATTAACCTTTAGTTGGGAAGGCAGAGGG-
TGACTGGGTG-
CTGACTGGTTTATTTATGTATATGTTTTTTATGTGTAGGTCCCGTCTCTGAC
GCAGATGAGACCCCCACTTCAGAGTGCATTTTCATCACCCCCAGAAATTGG-
CGAGGAACCG-
CCCGAAGATATTATTCATAGACCAGTTGCAGTGAGAGTCACCGGGCGGAGAG
CAGCTGTGGAGAGTTTGGATGACTTGCTACAGGGTGGGGATGAACCTTTG-
GACTTGTGTAC-
CCGGAACGCCCCAGGCACTAAGTGCCACACATGTGTGTTTACTTAAGGTGA
TGTCAGTATTTATAGGGTGTGGAGTGCAATAAAATCCGTGTTGACTTTAAG-
TGCGTGGTTTA-
TGACTCAGGGGTGGGGACTGTGGGTATATAAGCAGGTGCAGACCTGTGTGGT

CAGTTCAGAGCAGGACTCATGGAGATCTGGACGGTCTTGGAAGACTTTCAC-
CAGACTAGACA-
GCTGCTAGAGAACTCATCGGAGGGAGTCTCTTACCTGTGGAGATTCTGCTTC
GGTGGGCCTCTAGCTAAGCTAGTCTATAGGGCCAAGCAGGATTATAAGGAT-
CAATTTGAGGA-
TATTTTGAGAGAGTGTCCTGGTATTTTGGACTCTCTCAACTTGGGCCATCAG
TCTCACTTTAACCAGAGTATTCTGAGAGCCCTTGACTTTTCCACTCCTGG-
CAGAACTACCG-
CCGCGGTAGCCTTTTTTGCCTTTATCCTTGACAAATGGAGTCAAGAAACCCA
TTTCAGCAGGGATTACCGTCTGGACTGCTTAGCAGTAGCTTTGTGGAGAA-
CATGGAGGTG-
CCAGCGCCTGAATGCAATCTCCGGCTACTTGCCAGTACAGCCGGTAGACACG
CTGAGGATCCTGAGTCTCCAGTCACCCCAGGAACACCAACGCCGCCAGCAG-
CCGCAGCAGGA-
GCAGCAGCAAGAGGAGGACCGAGAAGAGAACCCGAGAGCCGGTCTGGACCCT
CCGGTGGCGGAGGAGGAGGAGTAGCTGACTTGTTTCCCGAGCTGCGCCGGG-
TGCTGACTAGG-
TCTTCCAGTGACGGGAGAGGGGGATTAAGCGGGAGAGGCATGAGGAGACTA
GTCACAGAACTGAACTGACTGTCAGTCTGATGAGCCGCAGGCGCCCAGAA-
TCGGTGTGGTGG-
CATGAGGTGCAGTCGCAGGGGATAGATGAGGTCTCGGTGATGCATGAGAAAT
ATTCCCTAGAACAAGTCAAGACTTGTTGGTTGGAGCCTGAGGATGATTGG-
GAGGTAGCCAT-
CAGGAATTATGCCAAGCTAGCTCTGAAGCCAGACAAGAAGTACAAGATTACC
AAACTGATTAAATATCAGAAATTCCTGCTACATTTTCAGGGAATGGGG-
CCGAGGTGGAGATCAG-
TACCCAGGAGAGGGTGGCCTTCAGATGCTGCATGATGAATATGTACCCGGGG
GTGGTGGGCATGGAGGGAGTCACCTTTATGAACGCGAGGTTCAGGGG-
CGATGGGTATAA-
TGGGGTGGTCTTTATGGCCAACACCAAGCTGACAGTGCACGGATGCTCCTTC

TTTGGCTTCAATAACATGTGCATCGAGGCCTGGGGCAGTGTTTCAG-
TGAGGGGATGCAG-
TTTTTTCAGCCAACCTGGATGGGGGTCGTGGGCAGAACCAAGAGCAAGGTGTCA
GTGAAGAAATGCCTGTTTCGAGAGGTGCCACCTGGGGGTGATGAGCGAGGG-
CGAAGCCAAAGT-
CAAACACTGCGCCTCTACTGAGACGGGCTGCTTTGTGCTGATCAAGGGCAAT
GCCCCAAGTCAAGCATAACATGATCTGTGGGGCCTCGGATGAGCGCGGCTAC-
CAGATGCTGAC-
CTGCGCCGGTGGGAACAGCCATATGCTGGCCACCGTGCATGTGACCTCGCAC
CCCCGCAAGACATGGCCCCGAGTTCGAGCACAAACGTCATGACCCGCTGCAA-
TGTGCACCTGGG-
CTCCCGCCGAGGCATGTTTCATGCCCTACCAGTGCAACATGCAATTTGTGAAG
GTGCTGCTGGAGCCCGATGCCATGTCCAGAGTGAGCCTGACGGGGGTG-
TTTGACATGAATG-
TGGAGATGTGGAAAATTCTGAGATATGATGAATCCAAGACCAGGTGCCGGGC
CTGCGAATGCGGAGGCAAGCACGCCAGGCTTCAGCCCGTGTGTGTGGAGG-
TGACGGAGGAC-
CTGCGACCCGATCATTTGGTGTTGTCCTGCAACGGGACGGAGTTCGGCTCCA
GCGGGGAAGAATCTGACTAGAGTGAGTAGTGTTTGGGGGAGGTGGAGGG-
CCTGGATGAGGGG-
CAGAATGACTAAAATCTGTGTTTTTCTGCGCAGCAGCATGAGCGGAAGCGCC
TCCTTTGAGGGAGGGGTATTTCAGCCCTTATCTGACGGGGCG-
TCTCCCCTCCTGGGCGGGAG-
TGCGTCAGAATGTGATGGGATCCACGGTGGACGGCCGGCCCGTGCAGCCCGC
GAACTCTTCAACCCTGACCTACGCGACCCTGAGCTCCTCGTCCGTGGACG-
CAGCTGCCGCCG-
CAGCTGCTGCTTCCGCCGCCAGCGCCGTGCGCGGAATGGCCCTGGGCGCCGG
CTACTACAGCTCTCTGGTGGCCAACTCGAGTTCCACCAATAATCCCGCCAG-
CCTGAACGAG-
GAGAAGCTGCTGCTGCTGATGGCCCAGCTCGAGGCCCTGACCCAGCGCCTGG

GCGAGCTGACCCAGCAGGTTGCTCAGCTGCAGGCGGAGACGCGGGCCGCGG-
TTGCCACGG-
TGAAAACCAAATAAAAAATGAATCAATAAATAAACGGAGACGGTTGTTGATT
TTAACACAGAGTCTTGAATCTTTATTTGATTTTTTCGCGCGCGGTAGG-
CCCTGGACCACCGG-
TCTCGATCATTGAGCACCCGGTGGATCTTTTCCAGGACCCGGTAGAGGTGGG
CTTGGATGTTGAGGTACATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAG-
CTCCATTGCA-
GGGCCTCGTGCTCGGGGGTGGTGTGTAAATCACCCAGTCATAGCAGGGGCG
CAGGGCGTGGTGCTGCACGATGTCCTTGAGGAGGAGACTGATGGCCACGGG-
CAGCCCCTTGG-
TGTAGGTGTTGACGAACCTGTTGAGCTGGGAGGGATGCATGCGGGGGGAGAT
GAGATGCATCTTGGCCTGGATCTTGAGATTGGCGATGTTCCCGCCCAGA-
TCCCGCCGGGGG-
TTCATGTTGTGCAGGACCACCAGCACGGTGTATCCGGTGCACCTTGGGGAATT
TGTCATGCAACTTGAAGGGAAGGCGTGAAAGAATTTGGAGACGCCCTTG-
TGGCCGCCCAGG-
TTTTCCATGCACTCATCCATGATGATGGCGATGGGCCCCGTGGGCGGCGGCCT
GGGCAAAGACGTTTCGGGGGTCGGACACATCGTAGTTGTGGTCCTGGG-
TGAGCTCGTCATA-
GGCCATTTTAATGAATTTGGGGCGGAGGGTGCCCGACTGGGGGACGAAGGTG
CCCTCGATCCCGGGGGCGTAGTTGCCCTCGCAGATCTGCATCTCCCAGG-
CCTTGAGCTCGGA-
GGGGGGGATCATGTCCACCTGCGGGGCGATGAAAAAACGGTTTCCGGGGCG
GGGGAGATGAGCTGCGCCGAAAGCAGGTTCCGGAGCAGCTGGGACTTGCCG-
CAGCCGG-
TGGGGCCGTAGATGACCCCGATGACCGGCTGCAGGTGGTAGTTGAGGGAGAG
ACAGCTGCCGTCTTCGCGGAGGAGGGGGGCCACCTCGTTCATCATCTCGCG-
CACATGCATG-
TTCTCGCGCACGAGTTCCGCCAGGAGGCGCTCGCCCCCAGCGAGAGGAGCT

CTTGCAGCGAGGCGAAGTTTTTCAGCGGCTTGAGCCCGTCGGCCATGGG-
CATTTTGA-
GAGGGTCTGTTGCAAGAGTTCCAGACGGTCCCAGAGCTCGGTGATGTGCTCT
AGGGCATCTCGATCCAGCAGACCTCCTCGTTTTCGCGGGTTGGGGCGACTG-
CGGGAGTAGGG-
CACCAGGCGATGGGCGTCCAGCGAGGCCAGGGTCCGGTCCTTCCAGGGTCGC
AGGGTCCGCGTCAGCGTGGTCTCCGTCACGGTGAAGGGGTGCGCGCCGGG-
CTGGGCGCTTG-
CGAGGGTGCGCTTCAGGCTCATCCGGCTGGTTCGAGAACCGCTCCCGGTCGGC
GCCCTGTGCGTCGGCCAGGTAGCAATTGAGCATGAGTTCGTAGTTGAGCG-
CCTCGGCCGCG-
TGGCCCTTGGCGCGGAGCTTACCTTTGGAAGTGTGTCCGCAGACGGGACAGA
GGAGGGACTTGAGGGCGTAGAGCTTGGGGGCGAGGAAGACGGACTCGGGGG-
CGTAGGCG-
TCCGCGCCGCAGCTGGCGCAGACGGTCTCGCACTCCACGAGCCAGGTGAGGT
CGGGGCGGTGCGGGTCAAAAACGAGGTTTCCTCCGTGCTTTTTGATGCG-
TTTCTTAC-
CTCTGGTCTCCATGAGCTCGTGTCCCCGCTGGGTGACAAAGAGGCTGTCCGT
GTCCCCGTAGACCGACTTTATGGGCCGGTCCTCGAGCGGGGTGCCGCGG-
TCCTCGTCG-
TAGAGGAACCCCGCCCACTCCGAGACGAAGGCCCGGGTCCAGGCCAGCACGA
AGGAGGCCACGTGGGAGGGGTAGCGGTCGTTGTCCACCAGCGGGTCCAC-
CTTCTCCAGGGTA-
TGCAAGCACATGTCCCCCTCGTCCACATCCAGGAAGGTGATTGGCTTGTAAG
TGTAGGCCACGTGACCGGGGGTCCCGGCCGGGGGGGTATAAAAGGGGG-
CGGGCCCCTGCTCG-
TCCTCACTGTCTTCCGGATCGCTGTCCAGGAGCGCCAGCTGTTGGGGTAGGT
ATTCCCTCTCGAAGGCGGGCATGACCTCGGCACTCAGGTTGTCAGTTTC-
TAGAAACGAGGAG-
GATTTGATATTGACGGTGCCGTTGGAGACGCCTTTCATGAGCCCCTCGTCCA

TCTGGTCAGAAAAGACGATCTTTTTGTTGTCGAGCTTGGTGGCGAAGGAG-
CCGTAGAGGGCG-
TTGGAGAGCAGCTTGGCGATGGAGCGCATGGTCTGGTTCTTTTCCTTGTCGG
CGCGCTCCTTGGCGGCGATGTTGAGCTGCACGTACTCGCGGCCACGCAC-
TTCCATTCTGGG-
GAAGACGGTGGTGAGCTCGTCGGGCACGATTCTGACCCGCCAGCCGCGGTTG
TGCAGGGTGATGAGGTCCACGCTGGTGGCCACCTCGCCGCGCAGGGGCTCG-
TTGGTCCAGCA-
GAGGCGCCCGCCCTTGCGCGAGCAGAAGGGGGGCAGCGGGTCCAGCATGAGC
TCGTCGGGGGGGTCGGCGTCCACGGTGAAGATGCCGGGCAGGAGCTCGGGG-
TCGAAGTAG-
CTGATGCAGGTGCCCAGATCGTCCAGCGCCGCTTGCCAGTCGCGCACGGCCA
GCGCGCGCTCGTAGGGGCTGAGGGGCGTGCCCCAGGGCATGGGGTGCG-
TGAGCGCGGAGGCG-
TACATGCCGCAGATGTCGTAGACGTAGAGGGGCTCCTCGAGGACGCCGATGT
AGGTGGGGTAGCAGCGCCCCCGCGGATGCTGGCGCGCACGTAGTCGTACA-
GCTCGTG-
CGAGGGCGCGAGGAGCCCCGTGCCGAGGTTGGAGCGTTGCGGCTTTTCGGCG
CGGTAGACGATCTGGCGGAAGATGGCGTGGGAGTTGGAGGAGATGGTGGG-
CCTCTGGAAGA-
TGTTGAAGTGGGCGTGGGGCAGGCCGACCGAGTCCCTGATGAAGTGGGCGTA
GGAGTCCTGCAGCTTGGCGACGAGCTCGGCGGTGACGAGGACGTCCAGGG-
CGCAGTAG-
TCGAGGGTCTCTTGATGATGTCGTACTTGAGCTGGCCCTTCTGCTTCCACA
GCTCGCGGTTGAGAAGGAACTCTTCGCGGTCCTTCCAGTACTCTT-
CGAGGGGGAACCCG-
TCCTGATCGGCACGGTAAGAGCCCACCATGTAGAACTGGTTGACGGCCTTGT
AGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAAGCTTGCGCGGCCTTG-
CGCAGGGAGGTG-
TGGGTGAGGGCGAAGGTGTCGCGCACCATGACTTTGAGGAACTGGTGCTTGA

AGTCGAGGTCGTCGCAGCCGCCCTGCTCCCAGAGTTGGAAGTCCGTGCG-
CTTCTTGTAGG-
CGGGGTTGGGCAAAGCGAAAGTAACATCGTTGAAGAGGATCTTGCCCGCGCG
GGGCATGAAGTTGCGAGTGATGCGGAAAGGCTGGGGCACCTCGGCCCGG-
TTGTTGATGAC-
CTGGGCGGCGAGGACGATCTCGTCGAAGCCGTTGATGTTGTGCCCGACGATG
TAGAGTTCCACGAATCGCGGGCAGCCCTTGACGTGGGGCAGCTTCTTGAG-
CTCGTCGTAGG-
TGAGCTCGGCGGGGTCGCTGAGCCCGTGCTGCTCGAGGGCCCAGTCGGCGAC
GTGGGGGTTGGCGCTGAGGAAGGAAGTCCAGAGATCCACGGCCAGGGCGG-
TCTGCAAGCGG-
TCCCGGTACTGACGGAAGTCTGGCCACGGCCATTTTTTCGGGGGTGACGC
AGTAGAAGGTGCGGGGGTCGCCGTGCCAGCGGTCCCAGTTGAGTTGGAGGG-
CGAGGTCG-
TGGGCGAGCTCGACGAGCGGCGGGTCCCCGGAGAGTTTCATGACCAGCATGA
AGGGGACGAGCTGCTTGCCGAAGGACCCCATCCAGGTGTAGGTTTCCACA-
TCGTAGGTGAG-
GAAGAGCCTTTCGGTGCGAGGATGCGAGCCGATGGGGAAGAACTGGATCTCC
TGCCACCAGTTGGAGGAATGGCTGTTGATGTGATGGAAGTAGAAATGCCGA-
CGGCGCG-
CCGAGCACTCGTGCTTGTGTTTATACAAGCGTCCGCAGTGCTCGCAACGCTG
CACGGGATGCACGTGCTGCACGAGCTGTACCTGGGTTCTTTGACGAGGAA-
TTTCAGTGGG-
CAGTGGAGCGCTGGCGGCTGCATCTGGTGCTGTACTACGTCCTGGCCATCGG
CGTGGCCATCGTCTGCCTCGATGGTGGTCATGCTGACGAGCCCGCGCGGGA-
GGCAGGTCCA-
GACCTCGGCTCGGACGGGTCGGAGAGCGAGGACGAGGGCGCGCAGGCCGGAG
CTGTCCAGGGTCTTGAGACGCTGCGGAGTCAGGTCAGTGGGCAGCGGCGG-
CGCGCGGTTGAC-
TTGCAGGAGCTTTTCCAGGGCGCGCGGGAGGTCCAGATGGTACTTGATCTCC

ACGGCGCCGTTGGTGGCGACGTCCACGGCTTGCAGGGTCCCGTG-
CCCCTGGGGCGCCACCAC-
CGTGCCCCGTTTCTTCTTGGGCGGCGGCGGCTCCATGCTTAGAAGCGGCGGC
GAGGACGCGCGCCGGGCGGCAGGGGCGGCTCGGGGCCCGAGGCAGGGG-
CGGCAGGGGCACG-
TCGGCGCCGCGCGCGGGCAGGTTCTGGTACTGCGCCCGGAGAAGACTGGCGT
GAGCGACGACGCGACGGTTGACGTCCTGGATCTGACGCCTCTGGGTGAAGG-
CCACGGGAC-
CCGTGAGTTTGAACCTGAAAGAGAGTTCGACAGAATCAATTTTCGGTATCGTT
GACGGCGGCCTGCCGCAGGATCTCTTGACGTCGCCCCGAGTTGTCCTGG-
TAGGCGATCTCGG-
TCATGAACTGCTCGATCTCCTCCTCCTGAAGGTCTCCGCGGCCGGCGCGCTC
GACGGTGGCCGCGAGGTCGTTGGAGATGCGGCCCATGAGCTGCGAGAAGG-
CGTTCATGCCGG-
CCTCGTTCCAGACGCGGCTGTAGACCACGGCTCCGTTGGGGTCGCGCGCGCG
CATGACCACCTGGGCGAGGTTAAGCTCGACGTGGCGCGTGAAGACCGCG-
TAGTTGCAGAGG-
CGCTGGTAGAGGTAGTTGAGCGTGGTGGCGATGTGCTCGGTGACGAAGAAGT
ACATGATCCAGCGGCGGAGCGGCATCTCGCTGACGTCGCCCAGGGCTT-
CCAAGCG-
CTCCATGGTCTCGTAGAAGTCCACGGCGAAGTTGAAAACTGGGAGTTGCGC
GCCGAGACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGGCGATGG-
TGGCGCGCAC-
CTCGCGCTCGAAGGCCCGGGGGGCTCCTCTTCTTCCATCTCCTCCTCCTCT
TCCTCCTCCACTAACATCTCTTCTACTTCCTCCTCAGGAGGCGGCGG-
CGGGGGAGGGG-
CCCTGCGTCGCCGGCGGCGCACGGGCAGACGGTCGATGAAGCGCTCGATGGT
CTCCCCGCGCCGGCGACGCATGGTCTCGGTGACGGCGCGCCCGTCTCTCG-
CGGGGCCGCAGCG-
TGAAGACGCCGCCGCGCATCTCCAGGTGGCCGCCGGGGGGGTCTCCGTTGGG

CAGGGAGAGGGCGCTGACGATGCATCTTATCAATTGGCCCGTAGGGAC-
TCCGCGCAAGGAC-
CTGAGCGTCTCGAGATCCACGGGATCCGAAAACCGCTGAACGAAGGCTTCGA
GCCAGTCGCAGTCGCAAGGTAGGCTGAGCCCGGTTTCTTGTTCTTCGGGTA-
TTTGGTCGGGA-
GGCGGGCGGGCGATGCTGCTGGTGATGAAGTTGAAGTAGGCGGTCCTGAGAC
GGCGGATGGTGGCGAGGAGCACCAGGTCCTTGGGCCCCGGCTTGCTGGATG-
CGCAGACGG-
TCGGCCATGCCCCAGGCGTGGTCCTGACACCTGGCGAGGTCCTTGTAGTAGT
CCTGCATGAGCCGCTCTACGGGCACGTCCTCCTCGCCCGCGCGGCCGTG-
CATGCGCGTGAG-
CCCGAACCCGCGCTGCGGCTGGACGAGCGCCAGGTCGGCGACGACGCGCTCG
GCGAGGATGGCCTGCTGGATCTGGGTGAGGGTGGTCTGGAAGTCGTCGAAG-
TCGACGAAG-
CGGTGGTAGGCTCCGGTGTTGATGGTGTAGGAGCAGTTGGCCATGACGGACC
AGTTGACGGTCTGGTGGCCGGGGCGCACGAGCTCGTGGTACTTGAGGCG-
CGAGTAGGCGCG-
CGTGTCGAAGATGTAGTCGTTGCAGGTGCGCACGAGGTACTGGTATCCGACG
AGGAAGTGCGGCGGCGGCTGGCGGTAGAGCGGCCATCGCTCGGTGG-
CGGGGGCGCCGGGCG-
CGAGGTCCTCGAGCATGAGGCGGTGGTAGCCGTAGATGTACCTGGACATCCA
GGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGGAACCTCGCGGACGCGGTT-
CCAGATGTTG-
CGCAGCGGCAGGAAGTAGTTCATGGTGGCCGCGGTCTGGCCCGTGAGGCGCG
CGCAGTCGTGGATGCTCTAGACATACGGGCAAAAACGAAAGCGGTCAGCGG-
CTCGACTCCG-
TGGCCTGGAGGCTAAGCGAACGGGTGGGCTGCGCGTGTACCCCGGTTTCGAA
TCTCGAATCAGGCTGGAGCCGCAGCTAACGTGGTACTGGCACTCCCG-
TCTCGACCCAAG-
CCTGCTAACGAAACCTCCAGGATACGGAGGCGGGTCGTTTTTTGGCCTTGGT

CGCTGGTCATGAAAACTAGTAAGCGCGGAAAGCGGCCGCCCGCGATGG-
CTCGCTGCCGTAG-
TCTGGAGAAAGAATCGCCAGGGTTGCGTTGCGGTGTGCCCCGGTTCGAGCCT
CAGCGCTCGGTGCCGGCCGGATTCCGCGGCTAACGTGGGCGTGGCTG-
CCCCGTCGTTTCCAA-
GACCCCTTAGCCAGCCGACTTCTCCAGTTACGGAGCGAGCCCTCTTTTTCT
TGTGTTTTTGCCAGATGCATCCCGTACTGCGGCAGATGCGCCCCAC-
CCTCCACCACAACCG-
CCCCTACCGCAGCAGCAGCAACAGCCGGCGCTTCTGCCCCGCCCCAGCAGC
AGCAGCCAGCCACTACCGCGGCGGCCCGCCGTGAGCGGAGCCGGCGTTCAG-
TATGACCTGG-
CCTTGAAGAGGGCGAGGGGCTGGCGCGGCTGGGGGCGTCGTCGCCGGAGCG
GCACCCGCGCGTGCAGATGAAAAGGGACGCTCGCGAGGCCTACGTG-
CCCAAGCAGAACCTG-
TTCAGAGACAGGAGCGGCGAGGAGCCCGAGGAGATGCGCGCCTCCCGCTTCC
ACGCGGGGCGGGAGCTGCGGCGCGGCCTGGACCGAAAGCGGGTGCTGAGG-
GACGAGGATTT-
CGAGGCGGACGAGCTGACGGGGATCAGCCCCGCGCGCGCGCACGTGGCCGCG
GCCAACCTGGTCACGGCGTACGAGCAGACCGTGAAGGAGGAGCAACTTT-
CAAAAATCCTT-
CAACAACCACGTGCGCACGCTGATCGCGCGCGAGGAGGTGACCCTGGGCCTG
ATGCACCTGTGGGACCTGCTGGAGGCCATCGTGCAGAACCCACGAGCAAG-
CCGCTGACGG-
CGCAGCTGTTTCTGGTGGTGCAGCACAGTCGGGACAACGAGACGTTTCAGGGA
GGCGCTGCTGAATATCACCGAGCCCGAGGGCCGCTGGCTCCTGGACCTGG-
TGAACATTCTG-
CAGAGCATCGTGGTGCAGGAGCGCGGGCTGCCGCTGTCCGAGAAGCTGGCGG
CCATCAACTTCTCGGTGCTGAGCCTGGGCAAGTACTACGCTAGGAAGATC-
TACAAGACCCCG-
TACGTGCCCATAGACAAGGAGGTGAAGATCGACGGGTTTTACATGCGCATGA

CCCTGAAAGTGCTGACCCTGAGCGACGATCTGGGGGTGTACCGCAACGACA-
GGATGCACCG-
CGCGGTGAGCGCCAGCCGCCGGCGCGAGCTGAGCGACCAGGAGCTGATGCAC
AGCCTGCAGCGGGCCCTGACCGGGGCCGGGACCGAGGGGGAGAGCTAC-
TTTGACATGGGCG-
CGGACCTGCGCTGGCAGCCCAGCCGCCGGGCCTTGGAAGCTGCCGGCGGGCT
GCCCTACGTGGAGGAGGTGGACGATGAGGAGGAGGAGGGCGAGTACCTGGA-
AGACTGATGG-
CGCGACCGTATTTTTGCTAGATGCAGCAACAGCCACCGCCGCCTCCTGATCC
CGCGATGCGGGCGGCGCTGCAGAGCCAGCCGTCCGGCATTAACTCCTCGGA-
CGATTGGAC-
CCAGGCCATGCAACGCATCATGGCGCTGACGACCCGCAATCCCGAAGCCTTT
AGACAGCAGCCTCAGGCCAACCGGCTCTCGGCCATCCTGGAGGCCGTGGTG-
CCCTCGCG-
CTCGAACCCACGCACGAGAAGGTGCTGGCCATCGTGAACGCGCTGGTGGAG
AACAAGGCCATCCGCGGCGACGAGGCCGGGCTGGTGTACAACGCGCTGCTG-
GAGCGCGTGG-
CCCGCTACAACAGCACCAACGTGCAGACGAACCTGGACCGCATGGTGACCGA
CGTGCGCGAGGCGGTGTGCGAGCGCGAGCGGTTCCACCGCGAGTCGAAC-
CTGGGCTCCATGG-
TGGCGCTGAACGCCTTCCTGAGCACGCAGCCCGCCAACGTGCCCCGGGGCCA
GGAGGACTACACCAACTTCATCAGCGCGCTGCGGCTGATGGTGGCCGAGG-
TGCCCCAGAG-
CGAGGTGTACCAGTCGGGGCCGGACTACTTCTTCCAGACCAGTCGCCAGGGC
TTGCAGACCGTGAACCTGAGCCAGGCTTTCAAGAACTTGCAGGGACTG-
TGGGGCGTGCAGG-
CCCCGGTCGGGGACCGCGCGACGGTGTGAGCCTGCTGACGCCGAACCTCGCG
CCTGCTGCTGCTGCTGGTGGCGCCCTTCACGGACAGCGGCAGCGTGAGCCG-
CGACTCGTAC-
CTGGGCTACCTGCTTAACCTGTACCGCGAGGCCATCGGGCAGGCGCACGTGG

ACGAGCAGACCTACCAGGAGATCACCCACGTGAGCCGCGCGCTGGGCCAG-
GAGGACCCGGG-
CAACCTGGAGGCCACCCTGAACTTCCTGCTGACCAACCGGTCGCAGAAGATC
CCGCCCCAGTACGCGCTGAGCACCGAGGAGGAGCGCATCCTGCGCTACGTG-
CAGCAGAGCG-
TGGGGCTGTTCTGATGCAGGAGGGGGCCACGCCCAGCGCCGCGCTCGACAT
GACCGCGCGCAACATGGAGCCCAGCATGTACGCCCCGAACCGCCCGTTTCAT-
CAATAAG-
CTGATGGACTACTTGCATCGGGCGGCCGCCATGAACTCGGACTACTTTACCA
ACGCCATCTTGAACCCGCACTGGCTCCCGCCGCCCGGGTTCTACACGGG-
CGAGTACGACATG-
CCCGACCCCAACGACGGGTTCTGTGGGATGACGTGGACAGCAGCGTGTCT
CGCCGCGTCCCACCACCACCGTGTGGAAGAAAGAGGGCGGGGACCGGCGG-
CCGTCTCTGGCG-
CTGTCCGGTCGCGCGGGTGCTGCCGCGGCGGTGCCCGAGGCCGCCAGCCCCT
TTCCGAGCCTGCCCTTTTCGCTGAACAGCGTGCGCAGCAGCGAGCTGGG-
TCGGCTGACGCGG-
CCGCGCCTGCTGGGCGAGGAGGAGTACCTGAACGACTCCTTGTTGAGGCCCG
AGCGCGAAAAGAACTTCCCCAATAACGGGATAGAGAGCCTGGTGGACAAGA-
TGAGCCGCTG-
GAAGACGTACGCGCACGAGCACAGGGACGAGCCCCGAGCTAGCAGCGCAGGC
ACCCGTAGACGCCAGCGGCACGACAGGCAGCGGGGTCTGGTGTGGGA-
CGATGAGGATTCCG-
CCGACGACAGCAGCGTGTTGGACTTGGGTGGGAGTGGTGGTGGTAACCCGTT
CGCTCACTTGCGCCCCCGTATCGGGCGCCTGATGTAAGAATCTGAAAAA-
TAAAAAACGGTAC-
TCACCAAGGCCATGGCGACCAGCGTGCGTTCTTCTCTGTTGTTTGTAGTAGT
ATGATGAGGCGCGTGTACCCGGAGGGTCCTCCTCCCTCGTACGAGAGCG-
TGATGCAGCAGG-
CGGTGGCGGGCGGCGATGCAGCCCCCGCTGGAGGCGCCTTACGTGCCCCCGCG

GTACCTGGCGCCTACGGAGGGGCGGAACAGCATTCGTTACTCGGAGCTGG-
CACCCTTG-
TACGATACCACCCGGTTGTACCTGGTGGACAACAAGTCGGCGGACATCGCCT
CGCTGAACTACCAGAACGACCACAGCAACTTCCTGACCACCGTGGTGCA-
GAACAACGATTT-
CACCCCCACGGAGGCCAGCACCCAGACCATCAACTTTGACGAGCGCTCGCGG
TGGGGCGGCCAGCTGAAAACCATCATGCACACCAACATGCCCAACG-
TGAACGAGTTCATGTA-
CAGCAACAAGTTCAAGGCGCGGGTGATGGTCTCGCGCAAGACCCCCAACGGG
GTCACAGTAACAGATGGTAGTCAGGACGAGCTGACCTACGAGTGGGTGGAG-
TTTGAGCTG-
CCCGAGGGCAACTTCTCGGTGACCATGACCATCGATCTGATGAACAACGCCA
TCATCGACAACTACTTGCGGTGGGGCGGCAGAACGGGGTGCTGGAGAG-
CGACATCGGCG-
TGAAGTTCGACACGCGCAACTTCCGGCTGGGCTGGGACCCCGTGACCGAGCT
GGTGATGCCGGGCGTGACACCAACGAGGCCTTCCACCCCGACATCG-
TCCTGCTGCCCCG-
CTGCGGCGTGGAATTCACCGAGAGCCGCCTCAGCAACCTGCTGGGCATCCGC
AAGCGGCAGCCCTTCCAGGAGGGCTTCCAGATCCTGTACGAGGACCTGGA-
GGGGGGCAACA-
TCCCCGCGCTCTTGATGTCGAAGCCTACGAGAAAAGCAAGGAGGATAGCAC
CGCCGTGGCTACCGCCGCGACTGTGGCAGATGCCACTGTCACCAGGGGCGA-
TACATTCG-
CCACCCAGGCGGAGGAAGCAGCCGCCCTAGCGGCGACCGATGATAGTGAAAG
TAAGATAGTTATCAAGCCGGTGGAGAAGGACAGCAAGGACAGGAGCTA-
CAACGTTCTATCG-
GATGGAAAGAACACCGCCTACCGCAGCTGGTACCTGGCCTACAACCTACGGCG
ACCCCGAGAAGGGCGTGCGCTCCTGGACGCTGCTCACCACCTCGGACGT-
CACCTGCGGCGTG-
GAGCAAGTCTACTGGTCGCTGCCCGACATGATGCAAGACCCGGTCACCTTCC

GCTCCACGCGTCAAGTTAGCAACTACCCGGTGGTGGGCGCCGAGCTCCTG-
CCCGTCTAC-
TCCAAGAGCTTCTTCAACGAGCAGGCCGTCTACTCGCAGCAGCTGCGCGCCT
TCACCTCGCTCACGCACGTCTTCAACCGCTTCCCCGAGAACCAGATCCTCG-
TCCGCCCCG-
CCCGCGCCCAACCATTACCACCGTCAGTGAAAACGTTTCCTGCTCTCACAGATC
ACGGGACCCTGCCGCTGCGCAGCAGTATCCGGGGAGTCCAGCGCGTGACCG-
TCACTGACG-
CCAGACGCCGCACCTGCCCCCTACGTCTACAAGGCCCTGGGCGTAGTCGCGCC
GCGCGTCCTCTCGAGCCGCACCTTCTAAAAAATGTCCATTCTCATCTCG-
CCCAGTAATAACA-
CCGGTTGGGGCCTGCGCGCGCCAGCAAGATGTACGGAGGCGCTCGCCAACG
CTCCACGCAACACCCCGTGCGCGTGCGCGGGCACTTCCGCGCTCCCTGGGG-
CGCCCT-
CAAGGGTCGCGTGCGCTCGCGCACCAACCGTCGACGACGTGATCGACCAGGTG
GTGGCCGACGCGCGCAACTACACGCCCGCCGCGCGCCCGCCTCCACCGTG-
GACGCCGT-
CATCGACAGCGTGGTGGCCGACGCGCGCCGGTACGCCCCGCGCCAAGAGCCGG
CGGCGGCGCATCGCCCGGCGGCACCGGAGACCCCCGCCATGCGCGCGGGCG-
CGAGCCTTG-
CTGCGCAGGGCCAGGCGCACGGGACGCAGGGCCATGCTCAGGGCGGCCAGAC
GCGCGGCCTCCGGCAGCAGCAGCGCCGGCAGGACCCGCAGACGCGCGG-
CCACGGCGGGCGG-
CGGCGGCCATCGCCAGCATGTCCCGCCCGCGGGCGCGGCAACGTGTACTGGGT
GCGCGACGCCGCCACCGGTGTGCGCGTGCCCGTGCGCACCCGCCCCCCTCG-
CACTTGAAGA-
TGCTGACTTCGCGATGTTGATGTGTCCCAGCGGCGAGGAGGATGTCCAAGCG
CAAATACAAGGAAGAGATGCTCCAGGTCATCGCGCCTGAGATCTACGG-
CCCCGCGGCGGGCGG-
TGAAGGAGGAAAGAAAGCCCCGCAAACTGAAGCGGGTCAAAAAGGACAAAAA

GGAGGAGGAAGATGTGGACGGACTGGTGGAGTTTGTGCGCGAGTTCG-
CCCCCGGCGGGCGCG-
TGCAGTGGCGCGGGCGGAAAGTGAAACCGGTGCTGCGGCCCCGGCACCACGGT
GGTCTTCACGCCCCGGCGAGCGTTCCGGCTCCGCCTCCAAGCGCTCCTACGA-
CGAGGTG-
TACGGGGACGAGGACATCCTCGAGCAGGCGGCGAGCGTCTGGGCGAGTTTG
CTTACGGCAAGCGCAGCCGCCCCGCGCCCTTGAAAGAGGAGGCGGTG-
TCCATCCCGCTGGAC-
CACGGCAACCCACGCCGAGCCTGAAGCCGGTGACCCTGCAGCAGGTGCTGC
CGAGCGCGGCGCCGCGCCGGGGCTTCAAGCGCGAGGGCGGCGAGGATCTG-
TACCCGACCATG-
CAGCTGATGGTGCCCAAGCGCCAGAAGCTGGAGGACGTGCTGGAGCACATGA
AGGTGGACCCCGAGGTGCAGCCCGAGGTCAAGGTGCGGCCCATCAAGCAGG-
TGGCCCCGGG-
CCTGGGCGTGCGAGACCGTGGACATCAAGATCCCCACGGAGCCCATGGAAACG
CAGACCGAGCCCGTGAAGCCCAGCACCAGCACCATGGAGGTGCAGACGGA-
TCCCTGGATG-
CCGGCGCCGGCTTCCACCACCACTCGCCGAAGACGCAAGTACGGCGCGGCCA
GCCTGCTGATGCCCAACTACGCGCTGCATCCTTCCATCATCCCCACG-
CCGGGCTACCGCGG-
CACGCGCTTCTACCGCGGCTACAGCAGCCGCGCAAGACCACCACCCGCCGC
CGCCGTCGCCGCACCCGCCGCAGCACCACCGCGACTTCCGCCGCCG-
CCTTGGTGCGGAGAG-
TGTACCGCAGCGGGCGTGAGCCTCTGACCCTGCCGCGCGCGCTACCACCC
GAGCATCGCCATTTAACTCTGCCGTCGCCTCCTTGCAGATATGGCCCTCA-
CATGCCG-
CCTCCGCGTCCCCATTACGGGCTACCGAGGAAGAAAGCCGCGCCGTAGAAGG
CTGACGGGGGAACGGGCTGCGTCGCCATCACCACCGGCGGGCGCGCCAT-
CAGCAAGCGG-
TTGGGGGGAGGCTTCCTGCCCGCGCTGATCCCCATCATCGCCGCGGCGATCG

GGGCGATCCCCGGCATAGCTTCCGTGGCGGTGCAGGCCTCTCAGCGCCAC-
TGAGACACAG-
CTTGAAAAATTTGTAATAAAAAAATGGACTGACGCTCCTGGTCCTGTGATGT
GTGTTTTTAGATGGAAGACATCAATTTTTTCGTCCCTGGCACCGCGACACGG-
CACGCGGCCG-
TTTATGGGCACCTGGAGCGACATCGGCAACAGCCAACTGAACGGGGGCGCCT
TCAATTGGAGCAGTCTCTGGAGCGGGCTTAAGAATTTTCGGGTCCACGCT-
CAAAACCTATGG-
CAACAAGGCGTGGAACAGCAGCACAGGGCAGGCGCTGAGGGAAAAGCTGAAA
GAGCAGAACTTCCAGCAGAAGGTGGTCGATGGCCTGGCCTCGGGCAT-
CAACGGGGTGGTGGA-
CCTGGCCAACCAGGCCGTGCAGAAACAGATCAACAGCCGCCTGGACGCGGTC
CCGCCCCGCGGGGTCCGTGGAGATGCCCCAGGTGGAGGAGGAGCTG-
CCTCCCCTGGACAAGCG-
CGGCGACAAGCGACCGCGTCCCGACGCGGAGGAGACGCTGCTGACGCACACG
GACGAGCCGCCCCCGTACGAGGAGGCGGTGAAACTGGGTCTGCCCACCACG-
CGGCCCCGTGG-
CGCCTCTGGCCACCGGGGTGCTGAAACCCAGCAGCAGCAGCAGCCAGCCCGC
GACCCTGGACTTGCCCTCCACCTCGCCCCTCCACAGTGGCTAAGCCCCTG-
CCGCCGGTGGCCG-
TCGCGTCGCGCGCCCCCGAGGCCGCCCCCAGGCGAACTGGCAGAGCACTCT
GAACAGCATCGTGCGGTCTGGGAGTGCAGAGTGTGAAGCGCCGCGCTGCTA-
TTAAAAGACAC-
TGTAGCGCTTAACTTGCTTGCTGTGTGTATATGTATGTCCGCCGACCAGAA
GGAGGAGGAAGAGGCGCGTCGCCGAGTTGCAAGATGGCCACCCCATCGATG-
CTGCCCCAG-
TGGGCGTACATGCACATCGCCGGACAGGACGCTTCGGAGTACCTGAGTCCGG
GTCTGGTGCAGTTCGCCCGCGCCACAGACACCTACTTCAGTCTGGGGAA-
CAAGTTTAGGAAC-
CCCACGGTGGCACCCACGCACGATGTGACCACCGACCGCAGCCAGCGGCTGA

CGCTGCGCTTCGTGCCCCGTGGACCGCGAGGACAACACCTACTCGTACAAAG-
TGCCTACACG-
CTGGCCGTGGGCGACAACCGCGTGCTGGACATGGCCAGCACCTACTTTGACA
TCCGCGGCGTGCTGGATCGGGGCCCCAGCTTCAAACCCTACTCCGGCACCG-
CCTACAACAG-
CCTGGCTCCCAAGGGAGCGCCCAACACCTCACAGTGGATAACCAAAGACAAT
GGAAGTATAAGACATACAGTTTTGGAAATGCTCCAGTCAGAGGATTGGA-
CATTACA-
GAAGAGGGTCTCCAAATAGGAACCGATGAGTCAGGGGGTGAAAGCAAGAAAA
TTTTTGACAGACAAAACCTATCAGCCTGAACCTCAGCTTGGAGATGAGGAA-
TGGCATGATAC-
TATTGGAGCTGAAGACAAGTATGGAGGCAGAGCGCTTAAACCTGCCACCAAC
ATGAAACCCTGCTATGGGTCTTTCGCCAAGCCAACTAATGCTAAGGGAGGT-
CAGGCTAAAAG-
CAGAACCAAGGACGATGGCACTACTGAGCCTGATATTGACATGGCCTTCTTT
GACGATCGCAGTCAGCAAGCTAGTTTCAGTCCAGAACTTGTTTTGTATAC-
TGAGAATG-
TCGATCTGGACACCCCGGATACCCACATTATTTACAAACCTGGCACTGATGA
AACAAAGTTCTTCTTTCAACTTGGGTCAGCAGTCCATGCCCAACAGAC-
CCAACCTACATTGG-
CTTCAGAGACAACCTTTATCGGGCTCATGTACTACAACAGCACTGGCAATATG
GGTGTACTGGCCGGTCAGGCCTCCCAGCTGAATGCTGTGGTGGACTTGCAG-
GACAGAAACAC-
TGAAGTGTCTTACCAGCTCTTGCTTGACTCTCTGGGTGACAGAACCAGGTAT
TTCAGTATGTGGAATCAGGCGGTGGACAGCTATGACCCCGATGTGCGCAT-
TATTGAAAAT-
CACGGTGTGGAGGATGAACTCCCCAACTATTGCTTCCCTTTGAATGGTGTGG
GCTTTACAGATACATTCCAGGGAATTAAGGTTAAACTACAAATAACGGAA-
CAGCAAATGC-
TACAGAGTGGGAATCTGATACCTCTGTCAATAATGCTAATGAGATTGCCAAG

GGCAATCCTTTTCGCCATGGAGATCAACATCCAGGCCAACCTGTGGCGGAAC-
TTCCTCTACG-
CGAACGTGGCGCTGTACCTGCCCCGACTCCTACAAGTACACGCCGGCCAAACAT
CACGCTGCCCCACCAACACCAACACCTACGATTACATGAACGGCCGCGTGG-
TGGCGCCCTCG-
CTGGTGGACGCCTACATCAACATCGGGGCGCGCTGGTCGCTGGACCCCATGG
ACAACGTCAACCCCTTCAACCACCACCGCAACGCGGGCCTGCGCTACCG-
CTCCATG-
CTCCTGGGCAACGGGCGCTACGTGCCCTTCCACATCCAGGTGCCCCAAAAGT
TTTTCGCCATCAAGAGCCTCCTGCTCCTGCCCCGGGTCCTACACCTACGAG-
TGGAAC TTCCG-
CAAGGACGTCAACATGATCCTGCAGAGCTCCCTCGGCAACGACCTGCGCACG
GACGGGGCCTCCATCGCCTTCACCAGCATCAACCTCTACGCCACCTTCTT-
CCCCATGGCGCA-
CAACACCGCCTCCACGCTCGAGGCCATGCTGCGCAACGACACCAACGACCAG
TCCTTCAACGACTACCTCTCGGCGGCCAACATGCTCTACCCCATCCCGG-
CCAACGCCAC-
CAACGTGCCCATCTCCATCCCCTCGCGCAACTGGGCCGCCTTCCGCGGATGG
TCCTTCACGCGCCTCAAGACCCGCGAGACGCCCTCGCTCGGCTCCGGGTT-
CGACCCCTACTT-
CGTCTACTCGGGCTCCATCCCCTACCTCGACGGCACCTTCTACCTCAACCAC
ACCTTCAAGAAGGTCTCCATCACCTTCGACTCCTCCGTCAGCTGGCCCCGG-
CAACGACCG-
CCTCCTGACGCCCAACGAGTTCGAAATCAAGCGCACCGTCGACGGAGAGGGG
TACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTTCCTGGTCCAGATG-
CTGGCCCACTA-
CAACATCGGCTACCAGGGCTTCTACGTGCCCCGAGGGCTACAAGGACCGCATG
TACTCCTTCTTCCGCAACTTCCAGCCCATGAGCCGCCAGGTCGTGGA-
CGAGGTCAACTA-
CAAGGACTACCAGGCCGTCACCCTGGCCTACCAGCACAACAACCTCGGGCTTC

GTCGGCTACCTCGCGCCCACCATGCGCCAGGGCCAGCCCTACCCCGCCAAC-
TACCCCTAC-
CCGCTCATCGGCAAGAGCGCCGTCGCCAGCGTCACCCAGAAAAAGTTCCTCT
GCGACCGGGTCATGTGGCGCATCCCCTTCTCCAGCAACTTCATG-
TCCATGGGCGCGCTCAC-
CGACCTCGGCCAGAACATGCTCTACGCCAACTCCGCCCACGCGCTAGACATG
AATTTTGAAGTCGACCCCATGGATGAGTCCACCCTTCTCTATGTTGTCTT-
CGAAGTCTTCGA-
CGTCGTCCGAGTGCACCAGCCCCACCGCGGCGTCATCGAGGCCGTCTACCTG
CGCACGCCCTTCTCGGCCGGCAACGCCACCACCTAAGCCTCTTGCTTCTTG-
CAAGATGACGG-
CCTGTGGCTCCGGCGAGCAGGAGCTCAGGGCCATCCTCCGCGACCTGGGCTG
CGGGCCCTACTTCCTGGGCACCTTCGACAAGCGCTTCCCGGGATTCATGG-
CCCCGCACAAG-
CTGGCCTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGGCGAGCACT
GGCTGGCCTTCGCCTGGAACCCGCGCACCCACACCTGCTACCTCTTCGAC-
CCCTTCGGGTT-
CTCGGACGAGCGCCTCAAGCAGATCTACCAGTTCGAGTACGAGGGCCTGCTG
CGCCGCAGCGCCCTGGCCACCGAGGACCGCTGCGTCACCCTGGAAAAG-
TCCACCCAGACCG-
TGCAGGGTCCGCGCTCGGCCGCCTGCGGGCTCTTCTGCTGCATGTTCTGCA
CGCCTTCGTGCACTGGCCCGACCGCCCCATGGACAAGAACCCAC-
CATGAACTTGCTGA-
CGGGGGTGCCCAACGGCATGCTCCAGTCGCCCCAGGTGGAACCCACCCTGCG
CCGCAACCAGGAGGCGCTCTACCGCTTCTCAACGCCCACTCCGCCTAC-
TTTCGCTCCCAC-
CGCGCGCGCATCGAGAAGGCCACCGCCTTCGACCGCATGAATCAAGACATGT
AAACTGTGTGTATGTGAATGCTTTATTTCATAATAAACAGCACATGTTTATG-
CCACCTT-
CTCTGAGGCTCTGACTTTATTTAGAAATCGAAGGGGTCTGCCGGCTCTCGG

CGTGCCCCGCGGGCAGGGATACGTTGCGGAACTGGTACTTGGGCAGCCAC-
TTGAACTCGGG-
GATCAGCAGCTTCGGCACGGGGAGGTCGGGGAACGAGTCGCTCCACAGCTTG
CGCGTGAGTTGCAGGGCGCCCAGCAGGTCGGGCGCGGATATCTTGAAATCA-
CAGTTGGGAC-
CCGCGTTCTGCGCGCGAGAGTTGCGGTACACGGGGTTGCAGCACTGGAACAC
CATCAGGGCCGGGTGCTTCACGCTCGCCAGCACCGTCGCGTCGGTGATG-
CCCTCCACGTCCA-
GATCCTCGGCGTTGGCCATCCCGAAGGGGGTCATCTTGCAGGTCTGCCGCCC
CATGCTGGGCACGCAGCCGGGCTTGTGGTTGCAATCGCAGTGCAGGGGGAT-
CAGCAT-
CATCTGGGCCTGCTCGGAGCTCATGCCCCGGGTACATGGCCTTCATGAAAGCC
TCCAGCTGGCGGAAGGCCTGCTGCGCCTTGCCGCCCTCGGTGAAGAAGAC-
CCCGCAGGAC-
TTGCTAGAGAACTGGTTGGTGGCGCAGCCGGCGTCGTGCACGCAGCAGCGCG
CGTCGTTGTTGGCCAGCTGCACCACGCTGCGCCCCCAGCGGTTCTGGG-
TGATCTTGGCCCGG-
TCGGGGTTCTCCTTCAGCGCGCGCTGCCCGTTCTCGCTCGCCACATCCATCT
CGATCGTGTGCTCCTTCTGGATCATCACGGTCCCGTGCAGGCACCGCAG-
CTTGCCCTCGG-
CTTCGGTGCATCCGTGCAGCCACAGCGCGCAGCCGGTGCCTCCAGTTCTT
GTGGGCGATCTGGGAGTGCGAGTGACGAAGCCCTGCAGGAAGCGGCCCAT-
CATCGCGGTCA-
GGGTCTTGTTGCTGGTGAAGGTCAGCGGGATGCCGCGGTGCTCCTCGTTCAC
ATACAGGTGGCAGATGCGGCGGTACACCTCGCCCTGCTCGGGCATCAGCTG-
GAAGGCGGAC-
TTCAGGTCGCTCTCCACGCGGTACCGCTCCATCAGCAGCGTCATGACTTCCA
TGCCCTTCTCCCAGGCCGAAACGATCGGCAGGCTCAGGGGGTTCTTCACCG-
TTGTCATCT-
TAGTCGCCGCGCCGAGGTCAGGGGGTCGTTCTCGTCCAGGGTCTCAAACAC

TCGCTTGCCGTCCTTCTCGGTGATGCGCACGGGGGGAAAGCTGAAG-
CCCACGGCCGCCAG-
CTCCTCCTCGGCCTGCCTTTCGTCCTCGCTGTCCTGGCTGATGTCTTGCAAA
GGCACATGCTTGGTCTTGCGGGGTTTCTTTTTTGGGCGGCAGAGGCGGCGG-
CGGAGACGTG-
CTGGGCGAGCGCGAGTTCTCGCTCACCACGACTATTTCTTCTTCTTGGCCGT
CGTCCGAGACCACGCGGCGGTAGGCATGCCTCTTCTGGGGCAGAGGCGGA-
GGCGACGGG-
CTCTCGCGGTTTCGGCGGGCGGGCTGGCAGAGCCCCTTCCGCGTTCGGGGGTGC
GCTCCTGGCGGCGCTGCTCTGACTGACTTCCTCCGCGGCCGCGCCATTGTG-
TTCTCCTAGGGA-
GCAAGCATGGAGACTCAGCCATCGTCGCCAACATCGCCATCTGCCCCCGCCG
CCGCCGACGAGAACCAGCAGCAGCAGAATGAAAGCTTAACCGCCCCGCCG-
CCCAGCCCCAC-
CTCCGACGCCGCGGGCCCCAGACATGCAAGAGATGGAGGAATCCATCGAGATT
GACCTGGGCTACGTGACGCCCGCGGAGCACGAGGAGGAGCTGGCAGCGCG-
CTTTTCAG-
CCCCGGAAGAGAACCACCAAGAGCAGCCAGAGCAGGAAGCAGAGAGCGAGCA
GAGCCAGGCTGGGCTCGAGCATGGCGACTACCTGAGCGGGGCAGAGGACG-
TGCTCATCAAG-
CATCTGGCCCGCCAATGCATCATCGTCAAGGATGCGCTGCTCGACCGCGCCG
AGGTGCCCCCTCAGCGTGGCGGAGCTCAGCCGCGCCTACGAGCGCAAC-
CTCTTCTCGCCGCG-
CGTGCCCCCAAGCGCCAGCCCAACGGCACCTGCGAGCCCAACCCGCGCCTC
AACTTCTACCCGGTCTTCGCGGTGCCGAGGCCCTGGCCACCTACCAC-
CTCTTTTT-
CAAGAACCAAAGGATCCCCGTCTCCTGCCGCGCCAACCGCACCCGCGCCGAC
GCCCTGCTCAACCTGGGCCCCGGCGCCCGCCTACCTGATATCGCCTCCTTG-
GAAGAGGTT-
CCCAAGATCTTCGAGGGTCTGGGCAGCGACGAGACTCGGGCCGCGAACGCTC

TGCAAGGAAGCGGAGAGGAGCATGAGCACCACAGCGCCCTGGTGGAGTTG-
GAAGGCGA-
CAACGCGCGCCTGGCGGTCCTCAAGCGCACGGTCGAGCTGACCCACTTCGCC
TACCCGGCGCTCAACCTGCCCCCAAGGTCATGAGCGCCGTCATGGACCA-
GGTGCTCAT-
CAAGCGCGCCTCGCCCCCTCTCGGAGGAGGAGATGCAGGACCCCGAGAGCTCG
GACGAGGGCAAGCCCGTGGTCAGCGACGAGCAGCTGGCGCGCTGGCTGGGA-
GCGAGTAGCAC-
CCCCAGAGCCTGGAAGAGCGGCGCAAGCTCATGATGGCCGTGGTCCTGGTG
ACCGTGGAGCTGGAGTGTCTGCGCCGCTTCTTCGCCGACGCGGAGACCCTG-
CGCAAGG-
TCGAGGAGAACCTGCACTACCTCTTCAGGCACGGGTTCGTGCGCCAGGCCTG
CAAGATCTCCAACGTGGAGCTGACCAACCTGGTCTCCTACATGGG-
CATCCTGCACGAGAAC-
CGCCTGGGGCAGAACGTGCTGCACACCACCCTGCGCGGGGAGGCCCGCCGCG
ACTACATCCGCGACTGCGTCTACCTGTACCTCTGCCACACCTGGCAGA-
CGGGCATGGGCGTG-
TGGCAGCAGTGCCTGGAGGAGCAGAACCTGAAAGAGCTCTGCAAGCTCCTGC
AGAAGAACCTGAAGGCCCTGTGGACCGGGTTCGACGAGCGCACCAACCG-
CCTCGGACCTGG-
CCGACCTCATCTTCCCCGAGCGCCTGCGGCTGACGCTGCGCAACGGGCTGCC
CGACTTTATGAGCCAAAGCATGTTGCAAAACTTTTCGCTCTTT-
CATCCTCGAACGCTCCGGGA-
TCCTGCCCCGCCACCTGCTCCGCGCTGCCCTCGGACTTCGTGCCGCTGACCTT
CCGCGAGTGCCCCCGCCGCTCTGGAGCCACTGCTACCTGCTGCGTCTGG-
CCAACCTACCTGG-
CCTACCACTCGGACGTGATCGAGGACGTCAGCGGCGAGGGTCTGCTCGAGTG
CCTGCGCGCTGCAACCTCTGCACGCCGCACCGCTCCCTGGCCTGCAAC-
CCCCAGCTG-
CTGAGCGAGACCCAGATCATCGGCACCTTCGAGTTGCAAGGCCCGGCGAGG

AGGGCAAGGGGGGTCTGAAACTCACCCCGGGGCTGTGGACCTCGGCCTAC-
TTGCGCAAGTT-
CGTGCCCGAGGACTACCATCCCTTCGAGATCAGGTTCTACGAGGACCAATCC
CAGCCGCCCCAAGGCCGAGCTGTTCGGCCTGCGTCATCACCCAGGGGG-
CCATCCTGGCCCAA-
TTGCAAGCCATCCAGAAATCCCGCCAAGAATTTCTGCTGAAAAAGGGCCACG
GGGTCTACTTGGACCCCCAGACCGGAGAGGAGCTCAACCCAGCTT-
CCCCCAGGATGCCCA-
GAGGAAGCAGCAAGAAGCTGAAAGTGGAGCTGCCGCTGCCGCCGGAGGATTT
GGAGGAAGACTGGGAGAGCAGTCAGGCAGAGGAGGAGGAGATGGAAGAC-
TGGGACAGCACT-
CAGGCAGAGGAGGACAGCCTGCAAGACAGTCTGGAAGACGAGGTGGAGGAGG
AGGCAGAGGAAGAAGCAGCCGCCGCCAGACCGTCGTCCTCGGCGGAGAAAG-
CAAGCAGCACG-
GATACCATCTCCGCTCCGGGTCTCGGGGTCTCGGCGGCCGGGCCCACAGTAGGT
GGGACGAGACCGGGCGCTTCCCGAACCCCAACCCAGACCGGTAAGAAG-
GAGCGGCAGGGA-
TACAAGTCCTGGCGGGGGCACAAAAACGCCATCGTCTCCTGCTTGCAAGCCT
GCGGGGGCAACATCTCCTTACCCGGCGCTACCTGCTCTTCCACCGCGGGG-
TGAACCTT-
CCCCCGCAACATCTTGCATTACTACCGTCACCTCCACAGCCCCTACTACTGT
TTCCAAGAAGAGGCAGAAACCCAGCAGCAGCAGAAAACCAGCAGCAGC-
TAGAAAATCCACAG-
CGGCGGCGGCGGCAGGTGGACTGAGGATCGCGGCGAACGAGCCGGCGCAGAC
CCGGGAGCTGAGGAACCGGATCTTTCCCAACCTCTATGCCATCTTCCAGCA-
GAGTCGGGGG-
CAGGAGCAGGAACTGAAAGTCAAGAACCGTTCTCTGCGCTCGCTCACCCGCA
GTTGTCTGTATCACAAGAGCGAAGACCAACTTCAGCGCACTCTCGAGGACG-
CCGAGG-
CTCTCTTCAACAAGTACTGCGCGCTCACTCTTAAAGAGTAGCCCGCGCCCGC

CCACACACGGAAAAAGGCGGGAATTACGTCACCACCTGCGCCCTTCG-
CCCGACCATCAT-
CATGAGCAAAGAGATTCCCACGCCTTACATGTGGAGCTACCAGCCCCAGATG
GGCCTGGCCGCCGGCGCCGCCAGGACTACTCCACCCGCATGAACTGGCT-
CAGTGCCGGG-
CCCGCGATGATCTCACGGGTGAATGACATCCGCGCCCGCCGAAACCAGATAC
TCCTAGAACAGTCAGCGATCACCGCCACGCCCCGCCATCACCTTAATCCG-
CGTAATTGG-
CCCGCCGCCCTGGTGTACCAGGAAATTCCCCAGCCCACGACCGTACTACTTC
CGCGAGACGCCAGGCCGAAGTCCAGCTGACTAACTCAGGTGTCCAGCTGG-
CCGGCGGCG-
CCGCCCTGTGTCGTCACCGCCCCGCTCAGGGTATAAAGCGGCTGGTGATCCG
AGGCAGAGGCACACAGCTCAACGACGAGGTGGTGAGCTCTTCGCTGGG-
TCTGCGACCTGACG-
GAGTCTTCCAACCTCGCCGGATCGGGGAGATCTTCCTTCACGCCTCGTCAGGC
CGTCCTGACTTTGGAGAGTTCGTCCTCGCAGCCCCGCTCGGGTGGCATCGG-
CACTCTCCAG-
TTCGTGGAGGAGTTCACTCCCTCGGTCTACTTCAACCCCTTCTCCGGCTCCC
CCGGCCACTACCCGGACGAGTTCATCCCGAACTTCGACGCCATCAGCGAG-
TCGGTGGACGGC-
TACGATTGAATGTCCCATGGTGGCGCGGCTGACCTAGCTCGGCTTCGACACC
TGGACCACTGCCGCCGCTTCCGCTGCTTCGCTCGGGATCTCGCCGAGTTTG-
CCTACTTTGAG-
CTGCCCCGAGGAGCACCCCTCAGGGCCCCGGCCACGGAGTGCGGATCATCGTCG
AAGGGGGCCTCGACTCCCACCTGCTTCGGATCTTCAGCCAGCG-
TCCGATCCTGGTCGAGCG-
CGAGCAAGGACAGACCCGTCTGACCCTGTACTGCATCTGCAACCACCCCGGC
CTGCATGAAAGTCTTTGTTGTCTGCTGTGTACTGAGTATAATAAAAGCTGA-
GATCAGCGAC-
TACTCCGGACTTCCGTGTGTTCCCTGAATCCATCAACCAGTCCCTGTTCTTCA

CCGGGAACGAGACCGAGCTCCAGCTCCAGTGTAAGCCCCACAAGAAGTAC-
CTCACCTGGCTG-
TTCCAGGGCTCCCCGATCGCCGTTGTCAACCACTGCGACAACGACGGAGTCC
TGCTGAGCGGCCCTGCCAACCTTACTTTTTCCACCCGCAGAAGCAAG-
CTCCAGCTCTT-
CCAACCCTTCCTCCCCGGGACCTATCAGTGCGTCTCGGGACCCTGCCATCAC
ACCTTCCACCTGATCCCGAATACCACAGCGTCGCTCCCCGCTACTAACAAC-
CAAACCTAC-
CCACCAACGCCACCGTCGCGACCTTTCCTCTGAATCTAATACCACTACCGGA
GGTGAGCTCCGAGGTGACCAACCTCTGGGATTTACTACGGCCCCCTGGGA-
GGTGGTGGGGT-
TAATAGCGCTAGGCCTAGTTGTGGGTGGGCTTTTGGCTCTCTGCTACCTATA
CCTCCCTTGCTGTTTCGTACTTAGTGGTGCTGTGTTGCTGGTTTAAGAAA-
TGGGGCAGATCAC-
CCTAGTGAGCTGCGGTGTGCTGGTGGCGGTGGTGCTTTCGATTGTGGGACTG
GGCGGCGCGGCTGTAGTGAAGGAGAAGGCCGATCCCTGCTTGCAATTTCAA-
TCCCGACAAATG-
CCAGCTGAGTTTTTCAGCCCGATGGCAATCGGTGCGCGGTGCTGATCAAGTGC
GGATGGGAATGCGAGAACGTGAGAATCGAGTACAATAACAAGACTCGGAA-
CAATACTCTCG-
CGTCCGTGTGGCAGCCCGGGGACCCCGAGTGGTACACCGTCTCTGTCCCCGG
TGCTGACGGCTCCCCGCGCACCGTGAATAATACTTTCATTTTTTGCGCACA-
TGTGCGACACGG-
TCATGTGGATGAGCAAGCAGTACGATATGTGGCCCCCAGGAAGGAGAACAT
CGTGGTCTTCTCCATCGCTTACAGCCTGTGCACGGTGCTAATCACCGCTA-
TCGTGTG-
CCTGAGCATTCACATGCTCATCGCTATTCGCCCCAGAAATAATGCCGAAAAA
GAGAAACAGCCATAACACGTTTTTTCACACACCTTGTTTTTACAGACAATG-
CGTCTGTTAAA-
TTTTTTAAACATTGTGCTCAGTATTGCTTATGCCTCTGGCTATGCAAACATA

CAGAAAACCCCTCTATGTAGGATCTGATGATACACTAGAGGGTACCCAATCA-
CAAGCTAGGG-
TTTCATGGTATTTTTATAAAAGCTCAGATAATCCTATTACTCTTTGCAAAGG
TGATCAGGGGCGGACAACAAAGCCGCCTATCACATTTAGCTGTACCAGAA-
CAAATCTCACG-
CTTTTCTCAATTACAAAACAATATGCTGGTATTTATTACAGTACAACTTTC
ATAGTGGGCAAGATAAATATTATACTGTTAAGGTAGAAAATCCTACCAC-
TCCTAGAACTAC-
CACCACCACCACCACCACCACCCTACTGCGAAGCCCACTAAACCTAAAACT
ACCAAGAAAACCACTGTGAAAACCTACAACCTAGAACCACCACAACCTACA-
GAAACCACCACCAG-
CACAACACTTGCTGCAACTACACACACACACTGAGCTAACCTTACAGACC
ACTAATGATTTGATAGCCCTGTTGCAAAGGGGGATAACAGCACCCTT-
CCAATGAGGAGA-
TACCCAAATCCATGATTGGCATTATTGTTGCTGTAGTGGTGTGCATGTTGAT
CATCGCCTTGTCATGGTGTACTATGCCTTCTGCTACAGAAAGCACAGAC-
TGAACGACAAG-
CTGGAACACTTACTAAGTGTTGAATTTTAATTTTTTTAGAACCATGAAGATCC
TAGGCCTTTTAGTTTTTTCTATCATTACCTCTGCTCTATGCAATTCTGACA-
ATGAGGACGT-
TACTGTCGTTGTCGGATCAAATTATACACTGAAAGGTCCAGCGAAGGGTATG
CTTTCGTGGTATTGCTGGTTTGGAACCTGACACTGATCAAACCTGAGCTTTG-
CAATGCAA-
TGAAAGGTCAAATACCAACCTCAAAAATTAAACATAAATGCAATGGTACTGA
CTTAGTACTACTCAATATCACGAAATCATATGCTGGCAGCTATTCATG-
CCCTGGAGATGATG-
CTGAGAACATGATTTTTTTACAAAGTAACTGTTGTTGATCCCACTACTCCACC
ACCCACCACCACAACCTACTCACACCACACACAGAACAAACACCAGAGG-
CAGCAGAAGCA-
GAGTTGGCCTTCCAGGTTACGGAGATTCCTTTGCTGTCAATACCCCTACAC

CCGATCATCGGTGTCCGGGGCTGCTAGTCAGCGGCATTGTCCGTGTGCTTT-
CGGGATTAGCA-
GTCATAATCATCTGCATGTTCAATTTTTGCTTGCTGCTATAGAAGGCTTTACC
GACAAAAATCAGACCCACTGCTGAACCTCTATGTTTAATTTTTTCCAGAG-
CCATGAAGGCAG-
TTAGCGCTCTAGTTTTTTGTTCTTTGATTGGCATTGTTTTTTGCAATCCTAT
TACTAGAGTTAGCTTTATTAAAGATGTGAATGTTACTGAGGGGGGCAATG-
TGACACTGG-
TAGGTGTAGAGGGTGCTAAAAACACCACCTGGACAAAATACCACCTTGGGTG
GAAAGATATTTGCAATTGGAGTGTCACTGTGTACACATGTGAGGGAGTTAA-
TCTTACCATTG-
TCAATGCCACCTCAGCTCAAAATGGTAGAATTCAAGGACAAAGTGTTAGTGT
GACCAGTGATGGGTATTTTACCCAACATACTTTTATCTATGACGTTAAAGT-
CATACCACTG-
CCTACGCCTAGCCACCTAGCACCACTACACAAACAACCCACACTACACAGA
CAACCACATACAGTACATCAAATCAGCCTACCACCACTACAGCAGCAGAGG-
TTGCCAGCTCG-
TCTGGAGTTCAAGTGGCATTTTTTGTTGTTGCCCCCATCTAGCAGTCCCCTG
CTATTACCAATGAGCAGACTACTGCATTTTTGTCCACTGTCGAGAGCCACA-
CCACAGCTAC-
CTCCAGTGCCTTCTCTAGCACCGCCAATCTCTCCTCGCTTTCCTCTACACCA
ATCAGTCCCGCTACTACTACTACCCCCGCTATTCTTCCCACTCCCCTGAAG-
CAAACAGACGG-
CGGCATGCAATGGCAGATCACCTGCTCATTGTGATCGGGTTGGTCATCCTA
GCCGTGTTGCTCTACTACATCTTCTGCCGCCGCATTCCCAACGCGCACCG-
CAAGCCGGTCTA-
CAAGCCCATCATTGTCCGGCAGCCGGAGCCGCTTCAGGTGGAAGGGGGTCTA
AGGAATCTTCTCTTCTTTTTACAGTATGGTGATTGAACTATGATTCCTA-
GACAATT-
CTTGATCACTATTCTTATCTGCCTCCTCCAAGTCTGTGCCACCCTCGCTCTG

GTGGCCAACGCCAGTCCAGACTGTATTGGGCCCTTCGCCTCCTACGTG-
CTCTTTGCCTTCAT-
CACCTGCATCTGCTGCTGTAGCATAGTCTGCCTGCTTATCACCTTCTTCCAG
TTCATTGACTGGATCTTTGTGCGCATCGCCTACCTGCGCCACCACCCCCAG-
TACCGCGACCA-
GCGAGTGGCGCAGCTGCTCAGGCTCCTCTGATAAGCATGCGGGCTCTGCTAC
TTCTCGCGCTTCTGCTGTTAGTGCTCCCCCGTCCCGTTGACCCCCGG-
CCCCCCTCAG-
TCCCCCGAGGAGGTCCGCAAATGCAAATTCCAAGAACCCTGGAAATTCCTCA
AATGCTACCGCCAAAATCAGACATGCATCCCAGCTGGATCATGAT-
CATTGGGATCGTGAA-
CATTCTGGCCTGCACCCTCATCTCCTTTGTGATTTACCCCTGCTTTGACTTT
GGTTGGAACCTCGCCAGAGGCGCTCTATCTCCCGCCTGAACCTGACACAC-
CACCACAGCAAC-
CTCAGGCACACGCACTACCACCACCACAGCCTAGGCCACAATACATGCCCCAT
ATTAGACTATGAGGCCGAGCCACAGCGACCCATGCTCCCCGCTATTAGT-
TACTTCAATC-
TAACCGGCGGAGATGACTGACCCACTGGCCAACAACAACGTCAACGACCTTC
TCCTGGACATGGACGGCCGCGCCTCGGAGCAGCGACTCGCCCAACTTCG-
CATTCGCCAGCAG-
CAGGAGAGAGCCGTCAAGGAGCTGCAGGACGGCATAGCCATCCACCAGTGCA
AGAAAGGCATCTTCTGCCTGGTGAAACAGGCCAAGATCTCCTACGAGGT-
CACCCAGACCGAC-
CATCGCCTCTCCTACGAGCTCCTGCAGCAGCGCCAGAAGTTCACCTGCCTGG
TCGGAGTCAACCCCATCGTCATCACCCAGCAGTCGGGCGATACCAAGGGG-
TGCATCCACTG-
CTCCTGCGACTCCCCCGACTGCGTCCACACTCTGATCAAGACCCTCTGCGGC
CTCCGCGACCTCCTCCCCATGAACTAATCACCCACTTATCCAGTGAAA-
TAAAAAATAAT-
CATTTGATTTGAAATAAAGATACAATCATATTGATGATTTGAGTTTAACAAA

AATAAAGAATCACTTACTTGAAATCTGATACCAGGTCTCTGTCCATATTTT-
CTGCCAACAC-
CACCTCACTCCCCTCTTCCCAGCTCTGGTACTGCAGGCCCCGGCGGGCTGCA
AACTTCCTCCACACGCTGAAGGGGATGTCAAATTCCTCCTGCCCCCTCAA-
TCTTCATTTTA-
TCTTCTATCAGATGTCCAAAAAGCGCGTCCGGGTGGATGATGACTTCGACCC
CGTCTACCCCTACGATGCAGACAACGCACCGACCGTGCCCTTCATCAAC-
CCCCCCTTCG-
TCTCTTCAGATGGATTCCAAGAGAAGCCCCCTGGGGGTGTTGTCCCTGCGACT
GGCCGACCCCGTCAACCACCAAGAACGGGGAAATCACCCCTCAAGCTGGGA-
GAGGGGGTGGAC-
CTCGACTCCTCGGGAAAACATCTCCAACACGGCCACCAAGGCCGCTGCCC
CTCTCAGTTTTTCCAACAACACCATTTCCCTTAACATGGATCACCCCTTT-
TACACTAAAGA-
TGGAATAATTAGCCTTACAAGTTTCTCCACCATTAAATATACTGAGAACAAGC
ATTCTAAACACACTAGCTTTAGGTTTTGGATCAGGTTTAGGACTCCGTGG-
CTCTGCCTTGG-
CAGTACAGTTAGTCTCTCCACTTACATTTGATACTGATGGAAACATAAAGCT
TACCTTAGACAGAGGTTTGCATGTTACAACAGGAGATGCAATTGAAAGCAA-
CATAAGCTGGG-
CTAAAGGTTTAAAATTTGAAGATGGAGCCATAGCAACCAACATTGGAAATGG
GTTAGAGTTTGAAGCAGTAGTACAGAAACAGGTGTCGATGATGCTTAC-
CCAATCCAAGT-
TAAACTTGGATCTGGCCTTAGCTTTGACAGTACAGGAGCCATAATGGCTGGT
AACAAAGAAGACGATAAACTCACTTTGTGGACAACACCTGATCCATCAC-
CAAACGTCAAA-
TACTCGCAGAAAATGATGCAAACTAACACTTTGCTTGACTAAATGTGGTAG
TCAAATACTGGCCACTGTGTCAGTCTTAGTTGTAGGAAGTGGAAACC-
TAAACCCCATTAC-
TGGCACCGTAAGCAGTGCTCAGGTGTTTCTACGTTTTGATGCAAACGGTGTT

CTTTTAACAGAACATTCTACACTAAAAAATACTGGGGGTATAGGCAGGGA-
GATAGCATAGA-
TGGCACTCCATATGTCAATGCTGTAGGATTCATGCCCAATTTAAAAGCTTAT
CCAAAGTCACAAAGTTCTACTACTAAAAATAATATAGTAGGGCAAGTATA-
CATGAATGGAGA-
TGTTTCAAACCTATGCTTCTCACTATAACCCTCAATGGTACTGATGACAGC
AACAGTACATATTCAATGTCATTTTCATACACCTGGACTAATGGAAGCTA-
TGTTGGAGCAA-
CATTTGGAGCTAACTCTTATACCTTCTCCTACATCGCCCAAGAATGAATACT
GTATCCCACCCTGCATGCCCAACCCTCCCCACCTCTGTCTATATGGA-
AAACTCTGAAACA-
CAAAATAAAATAAAGTTCAAGTGTTTTATTGATTCAACAGTTTTACAGGATT
CGAGCAGTTATTTTTCCTCCACCCTCCCAGGACATGGAATACACCAC-
CCTCTCCCCCGCA-
CAGCCTTGAACATCTGAATGCCATTGGTGATGGACATGCTTTTGGTCTCCAC
GTTCCACACAGTTTCAGAGCGAGCCAGTCTCGGGTCGGTCAGGGAGA-
TGAAACCCTCCGGG-
CACTCCCGCATCTGCACCTCACAGCTCAACAGCTGAGGATTGTCCTCGGTGG
TCGGGATCACGGTTATCTGGAAGAAGCAGAAGAGCGGCGGTGGGAATCATA-
GTCCG-
CGAACGGGATCGGCCGGTGGTGTCGCATCAGGCCCCGCAGCAGTCGCTGCCG
CCGCCGCTCCGTCAAGCTGCTGCTCAGGGGGTCCGGGTCCAGGGACTCCCT-
CAGCATGATG-
CCCACGGCCCTCAGCATCAGTCGTCTGGTGCGGCGGGCGCAGCAGCGCATGC
GGATCTCGCTCAGGTCGCTGCAGTACGTGCAACACAGGACCACCAGGTTG-
TTCAACAGTCCA-
TAGTTCAACACGCTCCAGCCGAACTCATCGCGGGAAGGATGCTACCCACGT
GGCCGTCGTACCAGATCCTCAGGTAAATCAAGTGGCGCCCCCTCCAGAACA-
CGCTGCCCATG-
TACATGATCTCCTTGGGCATGTGGCGGTTACCCACCTCCCGGTACCACATCA

CCCTCTGGTTGAACATGCAGCCCCGGATGATCCTGCGGAACCACAGGGCCA-
GCACCGCCCCG-
CCCGCCATGCAGCGAAGAGACCCCGGGTCCCGGCAATGGCAATGGAGGACCC
ACCGCTCGTACCCGTGGATCATCTGGGAGCTGAACAAGTCTATGTTGGCA-
CAGCACAGGCA-
CACGCTCATGCATCTCTTCAGCACTCTCAGCTCCTCGGGGGTCAAAACCATA
TCCCAGGGCACGGGAAACTCTTGCAGGACAGCGAAGCCCGCAGAACAGGG-
CAATCCTCGCA-
CATAACTTACATTGTGCATGGACAGGGTATCGCAATCAGGCAGCACCGGGTG
ATCCTCCACCAGAGAAGCGCGGGTCTCGGTCTCCTCACAGCGTGG-
TAAGGGGGCCGGCCGA-
TACGGGTGATGGCGGGACGCGGCTGATCGTGTTGCGGACCGTGTCATGATGC
AGTTGCTTTCGGACATTTTCGTACTTGCTGAAGCAGAACCTGGTCCGGGCG-
CTGCACAC-
CGATCGCCGGCGGGCGGTCTCGGCGCTTGGAACGCTCGGTGTTGAAGTTGTAA
AACAGCCACTCTCTCAGACCGTGCAGCAGATCTAGGGCCTCAGGAG-
TGATGAAGATCCCAT-
CATGCCTGATGGCTCTGATCACATCGACCACCGTGGAATGGGCCAGACCCAG
CCAGATGATGCAATTTTGTTGGGTTTCGGTGACGGCGGGGGAGGGAAGAA-
CAGGAAGAAC-
CATGATTAACTTTTAATCCAAACGGTCTCGGAGCACTTCAAAATGAAGGTGCG
CGGAGATGGCACCTCTCGCCCCCGCTGTGTTGGTGGAAAATAACAGCCAGG-
TCAAAGGTGA-
TACGGTTCTCGAGATGTTCCACGGTGGCTTCCAGCAAAGCCTCCACGCGCAC
ATCCAGAAACAAGACAATAGCGAAAGCGGGAGGGTTCTCTAATTCCTCAAT-
CATCATGTTA-
CACTCCTGCACCATCCCCAGATAATTTTCATTTTCCAGCCTTGAATGATTC
GAACTAGTTCCTGAGGTAAATCCAAGCCAGCCATGATAAAGAGCTCGCGCA-
GAGCG-
CCCTCCACCGGCATTCTTAAGCACACCCTCATAATTCCAAGATATTCTGCTC

CTGGTTCACCTGCAGCAGATTGACAAGCGGGATATCAAAATCTCTGCCG-
CGATCCCTGAG-
CTCCTCCCTCAGCAATAACTGTAAGTACTCTTTCATATCCTCTCCGAAATTT
TTAGCCATAGGACCCCCAGGAATAAGAGAAGGGCAAGCCACATTACAGA-
TAAACCGAAG-
TCCCCCCCAGTGAGCATTGCCAAATGTAAGATTGAAATAAGCATGCTGGCTA
GACCCGGTGATATCTTCCAGATAACTGGACAGAAAATCGGGCAAGCAA-
TTTTTAAGAAAAT-
CAACAAAAGAAAAATCTTCCAGGTGCACGTTTAGGGCCTCGGGAACAACGAT
GGAGTAAGTGCAAGGGGTGCGTTCCAGCATGGTTAGTTAGCTGATCTG-
TAAAAAACAAAAA-
ATAAAACATTAAACCATGCTAGCCTGGCGAACAGGTGGGTAAATCGTTCTCT
CCAGCACCAGGCAGGCCACGGGGTCTCCGGCGCGACCCTCGTAAAAATTG-
TCGCTA-
TGATTGAAAACCATCACAGAGAGACGTTCCCGGTGGCCGGCGTGAATGATTC
GAGAAGAAGCATAACCCCCGGAACATTGGAGTCCGTGAGTGAAAAAAG-
CGGCCGAGGAAG-
CAATGAGGCACTACAACGCTCACTCTCAAGTCCAGCAAAGCGATGCCATGCG
GATGAAGCACAAAATTTTCAGGTGCGTAAAAAATGTAATTAC-
TCCCCTCCTGCACAGGCAG-
CGAAGCTCCCGATCCCTCCAGATACACATACAAAGCCTCAGCGTCCATAGCT
TACCGAGCGGCAGCAGCAGCGGCACACAACAGGCGCAAGAGTCAGAGAAAA-
GACTGAGCTC-
TAACCTGTCCGCCCCGCTCTCTGCTCAATATATAGCCCCAGATCTACACTGAC
GTAAAGGCCAAAGTCTAAAAATACCCGCCAAATAATCACACACGCCCAGCA-
CACGCCCA-
GAAACCGGTGACACACTCAAAAAAATACGCGCACTTCCTCAAACGCCCAAAC
TGCCGTCATTTCCGGGTTCACACGCTACGTCATCAAAACACGACTTTCAA-
TTCCGTCGAC-
CGTTAAAAACGTCACCCGCCCCGCCCCTAACGGTCGCCGCTCCCGCAGCCAA

TCAGCGCCCCGCATCCCCAAATTCAAACAGCTCATTTGCATATTAACGCG-
CACCAAAAGTTTGAGGTATATTATTGATGATG

SEQ ID NO: 3 Sequência polinucleotídica que codifica o promo-
tor CASI

GGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG-
CCCAACGAC-
CCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCCAC-
TTGGCAGTACAT-
CAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAAT
GGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTAC-
TTGGCAGTACATC-
TACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACGTTCTGCT
TCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTA-
TTTTTTAATTA-
TTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGC
GGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAG-
CCAATCAGAG-
CGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGGCCC
TATAAAAAGCGAAGCGCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGA-
TAGAGATCG-
TCGACGAGCTCGCGGCGGGCGGGAGTCGCTGCGCGCTGCCTTCGCCCCGTGC
CCCGCTCCGCCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGT-
TACTAAAACA-
GGTAAGTCCGGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCT
CCTCACGGCGAGCGCTGCCACGTCAGACGAAGGGCGCAGCGAGCG-
TCCTGATCCTTCCG-
CCCGGACGCTCAGGACAGCGGCCCGCTGCTCATAAGACTCGGCCTTAGAACC
CCAGTATCAGCAGAAGGACATTTTAGGACGGGACTTGGGTGACTCTAGGG-
CACTGGTTTT-
CTTTCCAGAGAGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCT

GCGGAGGGATCTCCGTGGGGCGGTGAACGCCGATGATGCCTCTACTAAC-
CATGTTTCATGTTTTCTTTTTTTTTCTACAGGTCCTGGGTGACGAACAG

**SEQ ID NO: 4 Sequência de polinucleotídeos que codifica
ChAd155/RSV**

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAA-
TGAGATGGGCGGCG-
CGGGGCGGGGCGCGGGGCGGGAGGCGGGTTTGGGGGCGGGCCGGC
GGGCGGGGCGGTGTGGCGGAAGTGGACTTTGTAAGTGTGGCGGA-
TGTGACTTGCTAGTG-
CCGGGCGCGGTAAAAGTGACGTTTTCCGTGCGCGACAACGCCCCCG
GGAAGTGACATTTTTCCCGCGGTTTTTACCGGATGTTGTAGTGAA-
TTTGGGCGTAACCAAG-
TAAGATTTGGCCATTTTCGCGGGAAACTGAAACGGGGGAAGTGAA
ATCTGATTAATTTTGCGTTAGTCATACCGCGTAATATTTGTC-
TAGGGCCGAGGGACTTTGG-
CCGATTACGTGGAGGACTCGCCCAGGTGTTTTTTGAGGTGAATTTC
CGCGTTCCGGGTCAAAGTCTGCGTTTTTATTATTATAGGATA-
TCCCATTGCATACGTTGTA-
TCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTAC
CGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAAT-
TACGGGGTCATTAGTT-
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG
GCCCCGCTGGCTGACCGCCCAACGACCCCCGCCCATGACGTCAA-
TAATGACGTATGTT-
CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG
AGTATTTACGGTAAACTGCCCCTTGGCAGTACATCAAGTGTATCA-
TATGCCAAGTACG-
CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATG
CCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATC-

TACGTATTAGTCATCGC-
TATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGA
TAGCGGTTTGACTCACGGGGATTTCOAAGTCTCCACCCCATTGACG-
TCAATGGGAGTTTG-
TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAAC
CCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG-
TCTATATAAGCAGAG-
CTCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAGAGATCG
TCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACG-
CCATCCACGCTGTTTTGAC-
CTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAA
CGGTGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGAGATCTT-
CCGTTTATCTAGGTACCA-
GATATCGCCACCATGGAAGTCTGATCCTGAAGGCCAACGCCATCA
CCACCATCCTGACCGCCGTGACCTTCTGCTTCGCCAGCGGCCAGAA-
CATCACCGAGGAATTC-
TACCAGAGCACCTGTAGCGCCGTGAGCAAGGGCTACCTGAGCGCC
CTGAGAACCGGCTGGTACACCAGCGTGATCACCATCGAGCTGAG-
CAACATCAAAGAAAA-
CAAGTGCAACGGCACCGACGCCAAAGTGAAGCTGATCAAGCAGGA
ACTGGACAAGTACAAGAACGCCGTGACCGAGCTGCAGCTG-
CTGATGCAGAGCACCCCCG-
CCACCAACAACCGGGCCAGACGGGAGCTGCCCCGGTTCATGAACT
ACACCCTGAACAACGCCAAAAAGACCAACGTGACCCTGAG-
CAAGAAGCGGAAGCGGCGGTT-
CCTGGGCTTTCTGCTGGGCGTGGGCAGCGCCATTGCCAGCGGCGTG
GCCGTGTCTAAGGTGCTGCACCTGGAAGGCGAAGTGAACAAGAT-
CAAGAGCGCCCTGCTGAG-
CACCAACAAGGCCGTGGTGTCCCTGAGCAACGGCGTGAGCGTGCT

GACCAGCAAGGTGCTGGATCTGAAGAACTACATCGACAAGCAG-
CTGCTGCCCATCGTGAA-
CAAGCAGAGCTGCAGCATCAGCAACATCGAGACAGTGATCGAGTT
CCAGCAGAAGAACAACCGGCTGCTGGAAATCACCCGGGAGTTCAG-
CGTGAACGCCGGCGTGA-
CCACCCCTGTGTCCACCTACATGCTGACCAACAGCGAGCTGCTGAG
CCTGATCAACGACATGCCCATCACCAACGACCAGAAAAAG-
CTGATGAGCAACAACGTGCAGA-
TCGTGCGGCAGCAGAGCTACTCCATCATGTCCATCATCAAAGAAGA
GGTGCTGGCCTACGTGGTGCAGCTGCCCCTGTACGGCGTGATCGA-
CACCCCCTGCTGGAAG-
CTGCACACCAGCCCCCTGTGCACCACCAACACCAAAGAGGGCAGC
AACATCTGCCTGACCCGGACCGACAGAGGCTGGTACTGCGA-
CAACGCCGGCAGCGTGTTCATT-
CTTTCCACAGGCCGAGACATGCAAGGTGCAGAGCAACCGGGTGTTT
TGCGACACCATGAACAGCCTGACCCTGCCCTCCGAAGTGAACCTG-
TGCAACGTGGACATCTT-
CAACCCCAAGTACGACTGCAAGATCATGACCTCCAAGACCGACGT
GTCCAGCTCCGTGATCACCTCCCTGGGCGCCATCGTGTCTCTGC-
TACGGCAAGACCAAGTG-
CACCGCCAGCAACAAGAACCGGGGCATCATCAAGACCTTCAGCAA
CGGCTGCGACTACGTGTCCAACAAGGGGGTGGACACCGTGTCCG-
TGGGCAACACCCTGTAC-
TACGTGAACAAACAGGAAGGCAAGAGCCTGTACGTGAAGGGCGAG
CCCATCATCAACTTCTACGACCCCCTGGTGTTCCTCCAGCGACGAG-
TTCGACGCCAGCATCAG-
CCAGGTGAACGAGAAGATCAACCAGAGCCTGGCCTTCATCCGGAA
GTCCGACGAGCTGCTGCACAATGTGAATGCCGGCAAGTCCACCAC-
CAACCGGAAGCGGAGAG-
CCCCTGTGAAGCAGACCCTGAACTTCGACCTGCTGAAGCTGGCCGG

CGACGTGGAGAGCAATCCCGGCCCTATGGCCCTGAGCAAAG-
TGAAACTGAACGATACAC-
TGAACAAGGACCAGCTGCTGTCCAGCAGCAAGTACACCATCCAGC
GGAGCACCGGCGACAGCATCGATACCCCCAACTACGACGTGCA-
GAAGCACATCAACAAGCTG-
TGCGGCATGCTGCTGATCACAGAGGACGCCAACCACAAGTTCACCG
GCCTGATCGGCATGCTGTACGCCATGAGCCGGCTGGGCGGGGAG-
GACACCATCAAGATCCTG-
CGGGACGCCGGCTACCACGTGAAGGCCAATGGCGTGGACGTGACC
ACACACCGGCAGGACATCAACGGCAAAGAAATGAAGTTCGAGG-
TGCTGACCCTGGCCAG-
CCTGACCACCGAGATCCAGATCAATATCGAGATCGAGAGCCGGAA
GTCCTACAAGAAAATGCTGAAAGAAATGGGCGAGGTGG-
CCCCCGAGTACAGACACGACAG-
CCCCGACTGCGGCATGATCATCCTGTGTATCGCCGCCCTGGTGATC
ACAAAGCTGGCCGCTGGCGACAGATCTGGCCTGACAGCCGTGAT-
CAGACGGGCCAACAATG-
TGCTGAAGAACGAGATGAAGCGGTACAAGGGCCTGCTGCCCCAAGG
ACATTGCCAACAGCTTCTACGAGGTGTTCGAGAAGTACCCCCACTT-
CATCGACGTGTTCGTG-
CACTTCGGCATTGCCCAGAGCAGCACCAGAGGCGGCTCCAGAGTG
GAGGGCATCTTCGCCGGCCTGTTCATGAACGCCTACGGCGCTGG-
CCAGGTGATGCTGAGA-
TGGGGCGTGCTGGCCAAGAGCGTGAAGAACATCATGCTGGGCCAC
GCCAGCGTGCAGGCCGAGATGGAACAGGTGGTGGAGGTG-
TACGAGTACGCCCAGAAGCTGGG-
CGGAGAGGCCGGCTTCTACCACATCCTGAACAACCCTAAGGCCTCC
CTGCTGTCCCTGACCCAGTTCCCCCACTTCTCCAGCGTGGTGCTGG-
GAAATGCCGCCGGAC-
TGGGCATCATGGGCGAGTACCGGGGCACCCCCAGAAACCAGGACC

TGTACGACGCCGCCAAGGCCTACGCCGAGCAG-
CTGAAAGAAAACGGCGTGATCAACTACAG-
CGTGCTGGACCTGACCGCTGAGGAACTGGAAGCCATCAAGCACCA
GCTGAACCCCAAGGACAACGACGTGGAGCTGGGAGGCGGAGGA-
TCTGGCGGCGGAGG-
CATGAGCAGACGGAACCCCTGCAAGTTCGAGATCCGGGGCCACTG
CCTGAACGGCAAGCGGTGCCACTTCAGCCACA ACTACTTCGAG-
TGGCCCCCTCATGCTCTG-
CTGGTGCGGCAGAACTTCATGCTGAACCGGATCCTGAAGTCCATGG
ACAAGAGCATCGACACCCTGAGCGAGATCAGCGGAGCCGCCGAG-
CTGGACAGAACCGAGGAA-
TATGCCCTGGGCGTGGTGGGAGTGCTGGAAAGCTACATCGGCTCCA
TCAACAACATCACAAAGCAGAGCGCCTGCGTGGCCATGAGCAAG-
CTGCTGACAGAGCTGAA-
CAGCGACGACATCAAGAAGCTGAGGGACAACGAGGAACTGAACAG
CCCCAAGATCCGGGTGTACAACACCGTGATCAGCTACATTGAGAG-
CAACCGCAAGAACAA-
CAAGCAGACCATCCATCTGCTGAAGCGGCTGCCCCGCCGACGTGCTG
AAAAAGACCATCAAGAACACCCTGGACATCCACAAGTCCATCAC-
CATCAACAA-
TCCCAAAGAAAGCACCGTGTCTGACACCAACGATCACGCCAAGAA
CAACGACACCACCTGATGAGCGGCCGCGATCTGCTGTGCCTTC-
TAGTTGCCAGCCATCTG-
TTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACT
CCCCTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTG-
TCTGAGTAGGTGTCATTC-
TATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTG
GGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG-
CCGATCAGCGATCG-
CTGAGGTGGGTGAGTGGGCGTGGCCTGGGGTGGTCATGAAAATAT

ATAAGTTGGGGGTCTTAGGGTCTCTTTATTTGTGTTGCAGAGACCG-
CCGGAGCCATGAGCGG-
GAGCAGCAGCAGCAGCAGTAGCAGCAGCGCCTTGGATGGCAGCAT
CGTGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCCGGGG-
TGCGTCAGAATG-
TGATGGGCTCCAGCATCGACGGCCGACCCGTCCTGCCCCGCAAATTC
CGCCACGCTGACCTATGCGACCGTCGCGGGGACGCCGTTGGACG-
CCACCGCCGCCGCCGCCG-
CCACCGCAGCCGCCTCGGCCGTGCGCAGCCTGGCCACGGACTTTGC
ATTCCTGGGACCACTGGCGACAGGGGCTACTTCTCGGGCCGCTG-
CTGCCGCCGTTTCG-
CGATGACAAGCTGACCGCCCTGCTGGCGCAGTTGGATGCGCTTACT
CGGGAACCTGGGTGACCTTTCTCAGCAGGTCATGGCCCTGCGCCAG-
CAGGTCTCCTCCCTG-
CAAGCTGGCGGGAATGCTTCTCCACAAATGCCGTTTAAGATAAAT
AAAACCAGACTCTGTTTGGATTAAAGAAAAGTAGCAAGTGCATTG-
CTCTCTTTATTTCATAA-
TTTTCCGCGCGCGATAGGCCCTAGACCAGCGTTCTCGGTCGTTGAG
GGTGCGGTGTATCTTCTCCAGGACGTGGTAGAGGTGGCTCTGGA-
CGTTGAGATACATGGG-
CATGAGCCCGTCCCGGGGGTGGAGGTAGCACCCTGCAGAGCTTC
ATGCTCCGGGGTGGTGTTGTAGATGATCCAGTCGTAGCAGGAGCG-
CTGGGCATGGTGCC-
TAAAAATGTCCTTCAGCAGCAGGCCGATGGCCAGGGGGAGGCCCT
TGGTGTAAGTGTTTACAAAACGGTTAAGTTGGGAAGGGTGCATT-
CGGGGAGAGATGATGTG-
CATCTTGGACTGTATTTTTAGATTGGCGATGTTTCCGCCCAGATCCC
TTCTGGGATTCATGTTGTGCAGGACCACCAGTACAGTGTATCCGG-
TGCACTTGGGGAATTTG-
TCATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGGAGACGCCT

TTGTGGCCTCCCAGATTTTCCATGCATTTCGTCCATGATGATGGCAA-
TGGGCCCCGCGGGAGG-
CAGCTTGGGCAAAGATATTTCTGGGGTCGCTGACGTCGTAGTTGTG
TTCCAGGGTGAGGTCGTCATAGGCCATTTTTACAAAGCGCGGGCG-
GAGGGTGCCCCGAC-
TGGGGGATGATGGTCCCCTCTGGCCCTGGGGCGTAGTTGCCCTCGC
AGATCTGCATTTCCCAGGCCTTAATCTCGGAGGGGGGAATCATA-
TCCACCTGCGGGG-
CGATGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAACTGGGATG
AGAGCAGGTTTCTAAGCAGCTGTGATTTTCCACAACCGGTGGG-
CCCATAAATAACACCTA-
TAACCGGTTGCAGCTGGTAGTTTAGAGAGCTGCAGCTGCCGTCGTC
CCGGAGGAGGGGGGCCACCTCGTTGAGCATGTCCCTGACGCG-
CATGTTCTCCCCGACCAGA-
TCCGCCAGAAGGCGCTCGCCGCCAGGGACAGCAGCTCTTGCAAG
GAAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATG-
TTTTTCAGGGTCTGGCT-
CAGCAGCTCCAGGCGGTCCCAGAGCTCGGTGACGTGCTCTACGGCA
TCTCTATCCAGCATATCTCCTCGTTTCGCGGGTTGGGGCGACTTT-
CGCTGTAGGGCACCAAG-
CGGTGGTTCGTCCAGCGGGGCCAGAGTCATGTCCTTCCATGGGCGCA
GGGTCCTCGTCAGGGTGGTCTGGGTCACGGTGAAGGGGTGCG-
CTCCGGGCTGAGCGCTTG-
CCAAGGTGCGCTTGAGGCTGGTTCTGCTGGTGCTGAAGCGCTGCCG
GTCTTCGCCCTGCGCGTCGGCCAGGTAGCATTTGACCATGGTGTCA-
TAGTCCAGCCCCCTCCG-
CGGCGTGTCCCTTGGCGCGCAGCTTGCCCTTGGAGGTGGCGCCGCA
CGAGGGGCAGAGCAGGCTCTTGAGCGCGTAGAGCTTGGGGG-
CGAGGAAGACCGATTTCGGGG-
GAGTAGGCGTCCGCGCCGCAGACCCCGCACACGGTCTCGCACTCCA

CCAGCCAGGTGAGCTCGGGGCGCGCCGGGTCAAAAACCAGGTTT-
CCCCCATGCTTTTTTGATG-
CGTTTCTTACCTCGGGTCTCCATGAGGTGGTGTCCCCGCTCGGTGAC
GAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCTTTT-
CTCCAGGGGGG-
TCCCTCGGTCTTCCTCGTAGAGGAACCTCGGACCACTCTGAGACGAA
GGCCCGCGTCCAGGCCAGGACGAAGGAGGCTATGTGGGAGGGG-
TAGCGGTCGTTGTCCAC-
TAGGGGGTCCACCTTCTCCAAGGTGTGAAGACACATGTGCGCCTTCC
TCGGCGTCCAGGAAGGTGATTGGCTTGTAGGTGTAGGCCACGTGA-
CCGGGGGGTTCCTGA-
CGGGGGGGTATAAAAGGGGGGTGGGGGCGCGCTCGTCGTCACTCTC
TTCCGCATCGCTGTCTGCGAGGGCCAGCTGCTGGGGTGAGTATT-
CCCTCTCGAAGGCGGG-
CATGACCTCCGCGCTGAGGTTGTCAGTTTCCAAAAACGAGGAGGAT
TTGATGTTACCTGTCCCGAGGTGATACCTTTGAGGGTACCCGCG-
TCCATCTGGTCAGAAAA-
CACGATCTTTTTATTGTCCAGCTTGGTGGCGAACGACCCGTAGAGG
GCGTTGGAGAGCAGCTTGGCGATGGAGCGCAGGGTCTGGTTCTTG-
TCCCTGTCGGCGCG-
CTCCTTGGCCGCGATGTTGAGCTGCACGTACTCGCGCGCGACGCAG
CGCCACTCGGGGAAGACGGTGGTGGCTCGTCGGGCACCAGGCG-
CACGCGCCAGCCGCGG-
TTGTGCAGGGTGACCAGGTCCACGCTGGTGGCGACCTCGCCGCGCA
GGCGCTCGTTGGTCCAGCAGAGACGGCCGCCCTTGCGCGAGCA-
GAAGGGGGGCAGGGGG-
TCGAGCTGGGTCTCGTCCGGGGGGTCCGCGTCCACGGTGAAAACCC
CGGGGCGCAGGCGCGCGTCGAAGTAGTCTATCTTGCAACCTTG-
CATGTCCAGCGCCTGCTG-
CCAGTCGCGGGGCGGCGAGCGCGCGCTCGTAGGGGTTGAGCGGCGG

GCCCCAGGGCATGGGGTGGGTGAGTGCGGAGGCGTACATGCCGCA-
GATGTCATAGACG-
TAGAGGGGCTCCCGCAGGACCCCGATGTAGGTGGGGTAGCAGCGG
CCGCCGCGGATGCTGGCGCGCACGTAGTCATACAGCTCGTG-
CGAGGGGGGCGAGGAGG-
TCGGGGCCCAGGTTGGTGCGGGCGGGGCGCTCCGCGCGGAAGACG
ATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGGTGGGGCG-
CTGGAAGACGTTGAAG-
CTGGCGTCCTGCAGGCCGACGGCGTCGCGCACGAAGGAGGCGTAG
GAGTCGCGCAGCTTGTGTACCAGCTCGGCGGTGACCTGCACG-
TCGAGCGCGCAGTAG-
TCGAGGGTCTCGCGGATGATGTCATATTTAGCCTGCCCCTTCTTTTT
CCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTCCAGTAC-
TCTTGGATCGGGAAAC-
CGTCCGGTTCCGAACGGTAAGAGCCTAGCATGTAGAACTGGTTGAC
GGCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAGG-
CCTGCGCGGCCTTGCGGA-
GCGAGGTGTGGGTCAGGGCGAAGGTGTCCCTGACCATGACTTTGAG
GTACTGGTGCTTGAAGTCGGAGTCGTGCGAGCCGCCCCGCTCCCA-
GAGCGAGAAGTCGGTG-
CGCTTCTTGGAGCGGGGGTTGGGCAGAGCGAAGGTGACATCGTTG
AAGAGGATTTTGCCCGCGCGGGGCATGAAGTTGCGGGTGATGCG-
GAAGGGCCCCGGCACTT-
CAGAGCGGTTGTTGATGACCTGGGCGGCGAGCACGATCTCGTCGAA
GCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAGCGGGG-
CCGGCCCTTTACGG-
TGGGCAGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGC
GAGGCCGTGCTCGGCCAGGGCCCAGTCCGCGAGGTGCGGGTTG-
TCTCTGAGGAAGGACTT-
CCAGAGGTGCGGGGCCAGGAGGGTCTGCAGGCGGTCTCTGAAGGT

CCTGAACTGGCGGCCACGGCCATTTTTTCGGGGGTGATGCAG-
TAGAAGGTGAGGGGG-
TCTTGCTGCCAGCGGTCCCAGTCGAGCTGCAGGGCGAGGTCGCGCG
CGGCGGTGACCAGGCGCTCGTCGCCCCCGAATTTTCATGACCAG-
CATGAAGGGCACGAGCTG-
CTTTCCGAAGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTG
ACAAAGAGGCGCTCCGTGCGAGGATGCGAGCCGATCGGGA-
AGAACTGGATCTCCCGCCACCA-
GTTGGAGGAGTGGCTGTTGATGTGGTGGAAGTAGAAGTCCCGTCGC
CGGGCCGAACACTCGTGCTGGCTTTTGTAAGCGAGCGCAGTAC-
TGGCAGCGCTGCACGGG-
CTGTACCTCATGCACGAGATGCACCTTTCGCCCCGCGCACGAGGAAG
CCGAGGGGAAATCTGAGCCCCCGCCTGGCTCGCGGCATGGCTGG-
TTCTCTTCTACTTTGGA-
TGCGTGTCCGTCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAG
CGGACCACCACGCCGCGCGAGCCGCAGGTCCAGATATCGGCGCG-
CGGCGGTTCGGAG-
TTTGATGACGACATCGCGCAGCTGGGAGCTGTCCATGGTCTGGAGC
TCCCGCGGCGGCGGCAGGTTCAGCCGGGAGTTCTTGCAGGTTTAC-
CTCGCAGAGTCGGGCCA-
GGGCGCGGGGCAGGTCTAGGTGGTACCTGATCTCTAGGGGCGTGTT
GGTGGCGGCGTCGATGGCTTGCAGGAGCCCGCAGCCCCGGGGGG-
CGACGACGGTGCCCCG-
CGGGGTGGTGGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGC
GGGCGGGCCCCCGGAGGTAGGGGGGGCTCCGGTCCCGCGGGCA-
GGGGCGGCAGCGGCACG-
TCGGCGTGGAGCGCGGGCAGGAGTTGGTGCTGTGCCCCGAGGTTG
CTGGCGAAGGCGACGACGCGGCGGTTGATCTCCTGGATCTGGCG-
CCTCTGCGTGAAGACGA-
CGGGCCCCGGTGAGCTTGAACCTGAAAGAGAGTTCGACAGAATCAA

TCTCGGTGTCATTGACCGCGGCCTGGCGCAGGATCTCCTGCACG-
TCTCCCGAGTTGTCTTGG-
TAGGCGATCTCGGCCATGAACTGCTCGATCTCTTCCTCCTGGAGGT
CTCCGCGTCCGGCGCGTTCCACGGTGGCCGCCAGGTCGTTGGAGA-
TGCGCCCCATGAGCTG-
CGAGAAGGCGTTGAGTCCGCCCTCGTTCCAGACTCGGCTGTAGACC
ACGCCCCCCTGGTCATCGCGGGCGCGCATGACCACCTGCGCGAGG-
TTGAGCTCCACGTGCCG-
CGCGAAGACGGCGTAGTTGCGCAGACGCTGGAAGAGGTAGTTGAG
GGTGGTGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAG-
CGGCGCAACGTGGATTCTG-
TTGATGTCCCCCAAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGT
CCACGGCGAAGTTGAAAACTGGGAGTTGCGCGCCGACACGGT-
CAACTCCTCCTCCAGAAGA-
CGGATGAGCTCGGCGACGGTGTGCGGCACCTCGCGCTCGAAGGCTA
TGGGGATCTCTTCCTCCGCTAGCATCACACCTCCTCCTCTT-
CCTCCTCTTCTGGCACTT-
CCATGATGGCTTCCTCCTCTTCGGGGGGTGGCGGGCGGCGGCGGTGG
GGGAGGGGGGCGCTCTGCGCCGGCGGGCGGCGCACCGGGAGGCGG-
TCCACGAAGCGCGCGAT-
CATCTCCCCGCGGCGGGCGGCATGGTCTCGGTGACGGCGCGGCCG
TTCTCCCGGGGGCGCAGTTGGAAGACGCCGCCGGACATCTGGTG-
CTGGGGCGGGTGGCCG-
TGAGGCAGCGAGACGGCGCTGACGATGCATCTCAACAATTGCTGC
GTAGGTACGCCGCCGAGGGACCTGAGGGAGTCCATATCCACCGGA-
TCCGAAAACCTTTCGAG-
GAAGGCGTCTAACCAGTCGCAGTCGCAAGGTAGGCTGAGCACCGT
GGCGGGCGGCGGGGGGTGGGGGGAGTGTCTGGCGGAGGTGCTG-
CTGATGATGTAATTGAAG-
TAGGCGGACTTGACACGGCGGATGGTCGACAGGAGCACCATGTCC

TTGGGTCCGGCCTGCTGGATGCGGAGGCGGTTCGGCTATGCCCA-
GGCTTCGTTCTGG-
CATCGGCGCAGGTCCTTGTAGTAGTCTTGCATGAGCCTTTCCACCG
GCACCTCTTCTCCTTCCTCTTCTGCTTCTTCCATGTCTGCTTCGG-
CCCTGGGGCGGGCGCCG-
CGCCCCCTGCCCCCATGCGCGTGACCCCGAACCCCTGAGCGGT
TGGAGCAGGGCCAGGTCGGCGACGACGCGCTCGGCCAGGATGG-
CCTGCTGCACCTGCG-
TGAGGGTGGTTTGGAAAGTCATCCAAGTCCACGAAGCGGTGGTAGG
CGCCCGTGTTGATGGTGTAGGTGCAGTTGGCCATGACGGACCAG-
TTGACGGTCTGGTGG-
CCCGGTTGCGACATCTCGGTGTACCTGAGTCGCGAGTAGGCGCGGG
AGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAGGTACTGGTAG-
CCCACCAGGAAGTGCGG-
CGGCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCTCC
GGGGGCCAGGTCTTCCAGCATGAGGCGGTGGTAGGCGTAGATG-
TACCTGGACATCCAGGTGA-
TACCCGCGGCGGTGGTGGAGGCGCGCGGGAAGTCGCGCACCCGGT
TCCAGATGTTGCGCAGGGGCAGAAAGTGCTCCATGGTAGGCGTG-
CTCTGTCCAGTCAGACG-
CGCGCAGTCGTTGATACTCTAGACCAGGGAAAACGAAAGCCGGTC
AGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAGGGTAT-
CATGGCGGAGGGCCTCGG-
TTCGAGCCCCGGGTCCGGGCCGGACGGTCCGCCATGATCCACGCGG
TTACCGCCCCGCGTGTCGAACCCAGGTGTGCGACGTCAGACAACGG-
TGGAGTGTTCTTTTGG-
CGTTTTTCTGGCCGGGCGCCGGCGCCGCGTAAGAGACTAAGCCGCG
AAAGCGAAAGCAGTAAGTGGCTCGCTCCCCGTAGCCGGAGGGA-
TCCTTGCTAAGGGTTGCG-
TTGCGGCGAACCCCGGTTCTGAATCCCGTACTCGGGCCGGCCGGACC

CGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTATAAAGACCCCG-
CTTGCGGATTGAC-
TCCGGACACGGGGACGAGCCCCTTTTATTTTGTCTTTCCCCAGATGC
ATCCGGTGCTGCGGCAGATGCGCCCCCGCCCCAGCAGCAGCAA-
CAACACCAGCAAGAGCGG-
CAGCAACAGCAGCGGGAGTCATGCAGGGCCCCCTCACCCACCCTC
GGCGGGCCGGCCACCTCGGCGTCCGCGGCCGTGTCTGGCGCCTG-
CGGCGGCGGCGGGGGG-
CCGGCTGACGACCCCGAGGAGCCCCCGCGGCGCAGGGCCAGACAC
TACCTGGACCTGGAGGAGGGCGAGGGCCTGGCGCGGCTGGGGG-
CGCCGTCTCCCGAGCG-
CCACCCGCGGGGTGCAGCTGAAGCGCGACTCGCGCGAGGCGTACGT
GCCTCGGCAGAACCTGTTTCAGGGACCGCGCGGGGCGAGGAG-
CCCGAGGAGATGCGGGACAGGA-
GGTTCAGCGCAGGGCGGGAGCTGCGGCAGGGGCTGAACCGCGAGC
GGCTGCTGCGCGAGGAGGACTTTGAGCCCGACGCGCGGACGGGGA-
TCAGCCCCGCGCGCGCG-
CACGTGGCGGCCCGCCGACCTGGTGACGGCGTACGAGCAGACGGTG
AACCAGGAGATCAACTTCCAAAAGAGTTTCAACAACCACGTGCG-
CACGCTGGTGGCGCG-
CGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTTTGTAAGC
GCGCTGGTGCAGAACCCCAACAGCAAGCCTCTGACGGCGCAGCTG-
TTCCTGATAGTGCAGCA-
CAGCAGGGACAACGAGGCGTTTAGGGACGCGCTGCTGAACATCAC
CGAGCCCGAGGGTCGGTGGCTGCTGGACCTGATTAACATCCTGCA-
GAGCATAGTGGTGCAG-
GAGCGCAGCCTGAGCCTGGCCGACAAGGTGGCGGCCATCAACTAC
TCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAAGATCTACCA-
GACGCCGTACGTGCCCA-
TAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGCATGGCGCT

GAAGGTGCTCACCTGAGCGACGACCTGGGCGTGTACCG-
CAACGAGCGCATCCACAAGGCCG-
TGAGCGTGAGCCGGCGGCGGAGCTGAGCGACCGCGAGCTGATGC
ACAGCCTGCAGCGGGCGCTGGCGGGCGCCGGCAGCGGCGACAGG-
GAGGCGGAGTCCTACTT-
CGATGCGGGGGCGGACCTGCGCTGGGCGCCCAGCCGGCGGGCCCT
GGAGGCCGCGGGGGTCCGCGAGGACTATGACGAGGACGGCGAG-
GAGGATGAGGAGTACGAGC-
TAGAGGAGGGCGAGTACCTGGACTAAACCGCGGGTGGTGTTCCTG
GTAGATGCAAGACCCGAACGTGGTGGACCCGGCGCTGCGGGCGG-
CTCTGCAGAGCCAGCCG-
TCCGGCCTTAACTCCTCAGACGACTGGCGACAGGTCATGGACCGCA
TCATGTGCTGACGGCGCGTAACCCGGACGCGTTCCGGCAGCAG-
CCGCAGGCCAACAGG-
CTCTCCGCCATCCTGGAGGCGGTGGTGCCTGCGCGCTCGAACCCCA
CGCACGAGAAGGTGCTGGCCATAGTGAACGCGCTGGCCGAGAACA-
GGGCCATCCGCCCGGA-
CGAGGCCGGGCTGGTGTACGACGCGCTGCTGCAGCGCGTGGCCCCG
CTACAACAGCGGCAACGTGCAGACCAACCTGGACCGGCTGG-
TGGGGGACGTGCGCGAGGCGG-
TGGCGCAGCGCGAGCGCGCGGATCGGCAGGGCAACCTGGGCTCCA
TGGTGGCGCTGAATGCCTTCCTGAGCACGCAGCCGGCCAACGTG-
CCGCGGGGGCAGGAAGAC-
TACACCAACTTTGTGAGCGCGCTGCGGCTGATGGTGACCGAGACCC
CCCAGAGCGAGGTGTACCAGTCGGGCCCCGGACTACTTCTTCCAGA-
CCAGCAGACAGGGCCTG-
CAGACGGTGAACCTGAGCCAGGCTTTCAAGAACCTGCGGGGGGCTG
TGGGGCGTGAAGGCGCCCACCGGCGACCGGGCGACGGTGTCCAG-
CCTGCTGACGCCAAC-
TCGCGCCTGCTGCTGCTGCTGATCGCGCCGTTACGGACAGCGGCA

GCGTGTCCCGGGACACCTACCTGGGGCACCTGCTGACCCTGTAC-
CGCGAGGCCATCGGGCA-
GGCGCAGGTGGACGAGCACACCTTCCAGGAGATCACCAGCGTGAG
CCGCGCGCTGGGGCAGGAGGACACGAGCAGCCTGGAGGCGAC-
TCTGAACTACCTGCTGAC-
CAACCGGCGGCAGAAGATTCCCTCGCTGCACAGCCTGACCTCCGAG
GAGGAGCGCATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAAC-
CTGATGCGCGACGGGG-
TGACGCCCAGCGTGGCGCTGGACATGACCGCGCGCAACATGGAAC
CGGGCATGTACGCCGCGCACCGGCCTTACATCAACCGCCTGATG-
GACTACCTGCATCGCG-
CGGCGGCCGTGAACCCCGAGTACTTTACCAACGCCATCCTGAACCC
GCACTGGCTCCCGCCGCCCCGGGTTCTACAGCGGGGGCTTCGAGG-
TCCCGGAGAC-
CAACGATGGCTTCCTGTGGGACGACATGGACGACAGCGTGTTCTCC
CCGCGGCCGCGAGGCGCTGGCGGAAGCGTCCCTGCTGCG-
TCCCAAGAAGGAGGAGGAGGAGGA-
GGCGAGTCGCCGCCGCGGCAGCAGCGGCGTGGCTTCTCTGTCCGAG
CTGGGGGCGGCAGCCGCCGCGCGCCCCGGGTCCCTGGGCGGCAG-
CCCCTTTCCGAGCCTGG-
TGGGGTCTCTGCACAGCGAGCGCACCAACCCGCCCTCGGCTGCTGGG
CGAGGACGAGTACCTGAATAACTCCCTGCTGCAGCCGGTGCGGGA-
GAAAAACCTGCCTCCCG-
CCTTCCCCAACAACGGGATAGAGAGCCTGGTGGACAAGATGAGCA
GATGGAAGACCTATGCGCAGGAGCACAGGGACGCGCCTGCG-
CTCCGGCCGCCCACGCGGCG-
CCAGCGCCACGACCGGCAGCGGGGGGCTGGTGTGGGATGACGAGGA
CTCCGCGGACGATAGCAGCGTGCTGGACCTGGGAGGGAGCGG-
CAACCCGTTTCGCGCACCTG-
CGCCCCCGCCTGGGGAGGATGTTTTAAAAAAAAAAAAAAAAAAGCA

AGAAGCATGATGCAAAAATTAAATAAAACTCACCAAGGCCATGG-
CGACCGAGCGTTGGTTT-
CTTGTGTTCCCTTCAGTATGCGGCGCGCGGCGATGTACCAGGAGGG
ACCTCCTCCCTCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGG-
CGCCCTCTTCTCCCTTTG-
CGTCGCAGCTGCTGGAGCCGCGGTACGTGCCTCCGCGCTACCTGCG
GCCTACGGGGGGGAGAAACAGCATCCGTTACTCGGAGCTGGCG-
CCCCTGTTTCGACACCAC-
CCGGGTGTACCTGGTGGACAACAAGTCGGCGGACGTGGCCTCCCTG
AACTACCAGAACGACCACAGCAATTTTTTGACCACGGTCATCCA-
GAACAATGACTACAG-
CCCGAGCGAGGCCAGCACCCAGACCATCAATCTGGATGACCGGTC
GCACTGGGGCGGCGACCTGAAAACCATCCTGCACACCAACATG-
CCCAACGTGAACGAGTT-
CATGTTACCAATAAGTTCAAGGCGCGGGTGATGGTGTGCGGCTCG
CACACCAAGGAAGACCGGGTGGAGCTGAAGTACGAGTGGGTGGA-
GTTTCGAGCTGCCAGAGGG-
CAACTACTCCGAGACCATGACCATTGACCTGATGAACAACGCGATC
GTGGAGCACTATCTGAAAGTGGGCAGGCAGAACGGGGTCTCTGGA-
GAGCGACATCGGGGT-
CAAGTTCGACACCAGGAACTTCCGCCTGGGGCTGGACCCCGTGACC
GGGCTGGTTATGCCCCGGGGTGTACACCAACGAGGCCTT-
CCATCCCGACATCATCCTGCTG-
CCCGGCTGCGGGGTGGACTTCACTTACAGCCGCCTGAGCAACCTCC
TGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGGCTTCAGGAT-
CACCTACGAGGACCTGGA-
GGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCCTACCAGGAT
AGCTTGAAGGAAAATGAGGCGGGACAGGAGGATACCGCCCCCG-
CCGCCTCCGCCGCCG-
CCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCGCGGACGGGG

CAGAGGCCGACCCCGCTATGGTGGTGGAGGCTCCCGAGCAGGAG-
GAGGACATGAATGACAG-
TGCGGTGCGCGGAGACACCTTCGTCACCCGGGGGGAGGAAAAGCA
AGCGGAGGCCGAGGCCGCGGCCGAGGAAAAGCAACTGGCGGCAG-
CAGCGGCGGCGGGCGGGCG-
TTGGCCGCGGCGGAGGCTGAGTCTGAGGGGACCAAGCCCGCCAAG
GAGCCCGTGATTAAGCCCCTGACCGAAGATAGCAAGAAGCGCAGT-
TACAACCTGCTCAAGGA-
CAGCACCAACACCGCGTACCGCAGCTGGTACCTGGCCTACAACCTAC
GGCGACCCGTCGACGGGGGTGCGCTCCTGGACCCTGCTGTGCACG-
CCGGACGTGACCTGCGG-
CTCGGAGCAGGTGTACTGGTCGCTGCCCCGACATGATGCAAGACCCC
GTGACCTTCCGCTCCACGCGGCAGGTCAGCAACTTCCCGGTGG-
TGGGCGCCGAGCTGCTG-
CCCGTGCACTCCAAGAGCTTCTACAACGACCAGGCCGTCTACTCCC
AGCTCATCCGCCAGTTCACCTCTCTGACCCACGTGTTCAATCG-
CTTTCCTGAGAACCAGATT-
CTGGCGCGCCCGCCCGCCCCCACCATCACCACCGTCAGTGAAAACG
TTCCTGCTCTCACAGATCACGGGACGCTACCGCTGCGCAACAG-
CATCGGAGGAGTCCAG-
CGAGTGACCGTTACTGACGCCAGACGCCGCACCTGCCCCCTACGTTT
ACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCCTTTCCAGCCGCAC-
TTTTTGAGCAACAC-
CACCATCATGTCCATCCTGATCTCACCAGCAATAACTCCGGCTGG
GGACTGCTGCGCGCGCCCAGCAAGATGTTTCGGAGGGGGCGAGGA-
AGCGTTCCGAGCAGCAC-
CCCGTGCGCGTGCGCGGGGCACTTCCGCGCCCCCTGGGGAGCGCACA
AACGCGGCCGCGCGGGGGCGCACCAACCGTGGACGACGCCATCGAC-
TCGGTGGTGGAGCAGGCG-
CGCAACTACAGGCCCGCGGTCTCTACCGTGGACGCGGCCATCCAGA

CCGTGGTGCGGGGCGCGCGGCGGTACGCCAAGCTGAAGAGCCG-
CCGGAAGCGCGTGGCCCG-
CCGCCACCGCCGCGGACCCGGGGCCGCGCCAAACGCGCCGCCGC
GGCCCTGCTTCGCCGGGCCAAGCGCACGGGCCGCGCGCCG-
CCATGAGGGCCGCGCGCCG-
CTTGGCCGCCGGCATCACCGCCGCCACCATGGCCCCCGTACCCGA
AGACGCGCGGCCGCGCGCCGCGCGCCGCCATCAGTGACATGG-
CCAGCAGGCGCCGGGG-
CAACGTGTACTGGGTGCGCGACTCGGTGACCGGCACGCGCGTGCCC
GTGCGCTTCCGCCCCCGCGGACTTGAGATGATGTGAAAAACAA-
CACTGAGTCTCCTGCTG-
TTGTGTGTATCCCAGCGGGCGGCGGCGCGCAGCGTCATGTCCAAG
CGCAAATCAAAGAAGAGATGCTCCAGGTCGTCGCGCCGGAGATC-
TATGGG-
CCCCCGAAGAAGGAAGAGCAGGATTCGAAGCCCCGCAAGATAAAG
CGGGTCAAAAAGAAAAAGAAAGATGATGACGATGCCGATGGGGA-
GGTGGAGTTCCTGCGCG-
CCACGGCGCCCAGGCGCCCGGTGCAGTGGAAGGGCCGGCGCGTAA
AGCGCGTCCTGCGCCCCGGCACCGCGGTGGTCTTCACGCCCGG-
CGAGCGCTCCACCCGGAC-
TTTCAAGCGCGTCTATGACGAGGTGTACGGCGACGAAGACCTGCTG
GAGCAGGCCAACGAGCGCTTCGGAGAGTTTGCTTACGGGAAGCGT-
CAGCGGGCGCTGGGGA-
AGGAGGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCACCC
CCAGTCTGAAGCCCGTGACCCTGCAGCAGGTGCTGCCGAGCAGCG-
CACCTCCGAGGCGAAG-
CGGGGTCTGAAGCGCGAGGGCGGCGACCTGGCGCCCACCGTGACG
CTCATGGTGCCCAAGCGGCAGAGGCTGGAGGATGTGCTGGAGAAA-
ATGAAAGTAGACCCCGG-
TCTGCAGCCGGACATCAGGGTCCGCCCCATCAAGCAGGTGGCGCCG

GGCCTCGGCGTGCAGACCGTGGACGTGGTCATCCCCACCGGCAAC-
TCCCCCGCCGCCGCCAC-
CACTACCGCTGCCTCCACGGACATGGAGACACAGACCGATCCCGCC
GCAGCCGCAGCCGCAGCCGCCGCCGCGACCTCCTCGGCGGAGGTG-
CAGACGGACCCCTGG-
CTGCCGCCGGCGATGTCAGCTCCCCGCGCGCGTCGCGGGCGCAGGA
AGTACGGCGCCGCCAACGCGCTCCTGCCCGAGTACGCCTTG-
CATCCTTCCATCGCGCCAC-
CCCCGGCTACCGAGGCTATACCTACCGCCCGCGAAGAGCCAAGGG
TTCCACCCGCCGTCCCCGCCGACGCGCCGCCGCCACCACCCGCCG-
CCGCCGCCGCAGACG-
CCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCGCGCGAC
GGACACACCCTGGTGCTGCCCAGGGCGCGCTACCACCCAG-
CATCGTTTAAAAGCCTGTTG-
TGGTTCTTGCAGATATGGCCCTCACTTGCCGCCTCCGTTTCCCGGTG
CCGGGATACCGAGGAGGAAGATCGCGCCGCAGGAGGGGTCTGG-
CCGGCCGCGGCCTGAGCG-
GAGGCAGCCGCCGCGCGCACCGGCGGCGACGCGCCACCAGCCGAC
GCATGCGCGGCGGGGTGCTGCCCCTGTTAATCCCCCTGATCGCCG-
CGGCGATCGGCGCCGTG-
CCCGGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAGGCATTGAC
AGACTTGCAAACCTTGCAAATATGGAAAAAAAAAACCCAA-
TAAAAAAGTCTAGACTCTCACG-
CTCGCTTGGTCCTGTGACTATTTTGTAGAATGGAAGACATCAACTTT
GCGTCGCTGGCCCCGCGTCACGGCTCGCGCCCGTTCCTGGGACAC-
TGGAACGATATCGGCAC-
CAGCAACATGAGCGGTGGCGCCTTCAGTTGGGGCTCTCTGTGGAGC
GGCATTA AAAAGTATCGGGTCTGCCGT TAAAAATTACGG-
CTCCCGGGCCTGGAACAGCAG-
CACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGAACTTCCA

GCAGAAGGTGGTGGAGGGCCTGGCCTCCGGCATCAACGGGGTGG-
TGGACCTGGCCAACCAGG-
CCGTGCAGAATAAGATCAACAGCAGACTGGACCCCCGGCCGCCGG
TGGAGGAGGTGCCGCCGGCGCTGGAGACGGTGTCCCCCGATGGG-
CGTGGCGAGAAGCGCCCG-
CGGCCCCGATAGGGAAGAGACCACTCTGGTCACGCAGACCGATGAG
CCGCCCCCGTATGAGGAGGCCCTGAAGCAAGGTCTGCCACCACG-
CGGCCCATCGCG-
CCCATGGCCACCGGGGTGGTGGGCGCCACACCCCCGCCACGCTGG
ACTTGCTCCGCCCCGCCGATGTGCCGCAGCAGCAGAAGGCGGCA-
CAGCCGGGCCCCGCCCCG-
CGACCGCTCCCGTTCCTCCGCCGGTCCTCTGCGCCGCGCGGCCAG
CGGCCCCCGCGGGGGGGTTCGCGAGGCACGGCAACTGGCAGAG-
CACGCTGAACAGCATCG-
TGGGTCTGGGGGTGCGGTCCGTGAAGCGCCGCCGATGCTACTGAAT
AGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTATGTGCCCG-
CCAGAGGAGCTGCTGAG-
TCGCCGCCGTTTCGCGCGCCCACCACCACCGCCACTCCGCCCCCTCAA
GATGGCGACCCCATCGATGATGCCGCAGTGGTCGTACATGCACA-
TCTCGGGCCAGGACG-
CCTCGGAGTACCTGAGCCCCGGGCTGGTGCAGTTCGCCCCGCGCCAC
CGAGAGCTACTTCAGCCTGAGTAACAAGTTTAGGAACCCACGG-
TGGCGCCACGCACGATG-
TGACCACCGACCGGTCTCAGCGCCTGACGCTGCGGTTTATTCCCGT
GGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGGTTTAC-
CCTGGCCGTGGGCGACAAC-
CGCGTGCTGGACATGGCCTCCACCTACTTTGACATCCGCGGGGTGC
TGGACCGGGGTCCCACTTTCAAGCCCTACTCTGGCACCGCCTA-
CAACTCCCTGG-
CCCCCAAGGGCGCTCCCAACTCCTGCGAGTGGGAGCAAGAGGAAA

CTCAGGCAGTTGAAGAAGCAGCAGAAGAGGAAGAAGAAGATG-
CTGACGGTCAAGCTGAGGA-
AGAGCAAGCAGCTACCAAAAAGACTCATGTATATGCTCAGGCTCCC
CTTTCTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACG-
GACGCTACAGCTACA-
GAACAAAAACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCC
AAATCGGGGAGTCCCAGTGGAATGAGGCAGATGCTACAGTCG-
CCGGCGGTAGAGTGC-
TAAAGAAATCTACTCCCATGAAACCATGCTATGGTTCCTATGCAAG
ACCCACAAATGCTAATGGAGGTCAGGGTGTACTAACGGCAAATG-
CCCAGGGACAGCTAGAA-
TCTCAGGTTGAAATGCAATTCTTTTCAACTTCTGAAAACGCCCCGTA
ACGAGGCTAACAACATTCAGCCCAAATTGGTGCTGTATAGTGAG-
GATGTGCACATGGAGAC-
CCCGGATACGCACCTTTCTTACAAGCCCGCAAAAAGCGATGACAAT
TCAAAAATCATGCTGGGTCAGCAGTCCATGCCCAACAGACCTAAT-
TACATCGGCTTCAGAGA-
CAACTTTATCGGCCTCATGTATTACAATAGCACTGGCAACATGGGA
GTGCTTGCAGGTCAGGCCTCTCAGTTGAATGCAGTGGTGGACTTG-
CAAGACAGAAACACA-
GAACTGTCCTACCAGCTCTTGCTTGATTCCATGGGTGACAGAACCA
GATACTTTTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCA-
GATGTTAGAATTA-
TTGAAAATCATGGAAGTGAAGACGAGCTCCCCAACTATTGTTTCCC
TCTGGGTGGCATAGGGGTAACTGACACTTACCAGGCTGTAAAAC-
CAACAATGGCAA-
TAACGGGGGGCCAGGTGACTTGGACAAAAGATGAACTTTTGCAGA
TCGCAATGAAATAGGGGTGGGAAACAATTTTCGCTATGGAGAT-
CAACCTCAGTGCCAACCTG-
TGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCAGACA

AGCTTAAGTACAACCCCTCCAATGTGGACATCTCTGACAAC-
CCCAACACCTACGATTACA-
TGAACAAGCGAGTGGTGGCCCCGGGGCTGGTGGACTGCTACATCA
ACCTGGGCGCGCGCTGGTCGCTGGACTACATGGACAACGTCAAC-
CCCTTCAACCACCACCG-
CAATGCGGGCCTGCGCTACCGCTCCATGCTCCTGGGCAACGGGGCGC
TACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTTCTTTGCCAT-
CAAGAAC-
CTCCTCCTCCTGCCGGGCTCCTACACCTACGAGTGGAACCTTCAGGA
AGGATGTCAACATGGTCCTCCAGAGCTCTCTGGGTAACGATCTCA-
GGGTGGACGGGGCCAG-
CATCAAGTTCGAGAGCATCTGCCTCTACGCCACCTTCTTCCCCATGG
CCCACAACACGGCCTCCACGCTCGAGGCCATGCTCAGGAACGACA-
CCAACGACCAGTCCTT-
CAATGACTACCTCTCCGCCGCCAACATGCTCTACCCCATACCCGCC
AACGCCACCAACGTCCCCATCTCCATCCCCTCGCGCAACTGGG-
CGGCCTTCCGCGGCTGGG-
CCTTCACCCGCCTCAAGACCAAGGAGACCCCTCCCTGGGCTCGGG
ATTGACCCCTACTACACCTACTCGGGCTCCATTCCCTACCTGGA-
CGGCACCTTCTACCT-
CAACCACACTTTCAAGAAGGTCTCGGTCACCTTCGACTCCTCGGTC
AGCTGGCCGGGCAACGACCGTCTGCTCACCCCCAACGAGTTCGA-
GATCAAGCGCTCGGTCGA-
CGGGGAGGGCTACAACGTGGCCCAGTGCAACATGACCAAGGACTG
GTTCTTGGTCCAGATGCTGGCCAACTACAACATCGGCTACCAGGG-
CTTCTACATCCCA-
GAGAGCTACAAGGACAGGATGTACTCCTTCTTCAGGAACTTCCAGC
CCATGAGCCGGCAGGTGGTGGACCAGACCAAGTACAAGGACTAC-
CAGGAGGTGGGCAT-
CATCCACCAGCACAACAACCTCGGGCTTCGTGGGCTACCTCGCCCCC

ACCATGCGCGAGGGACAGGCCTACCCCGCCAACTTCCCCTATCCG-
CTCATAGGCAAGACCG-
CGGTCGACAGCATCACCCAGAAAAAGTTCCTCTGCGACCGCACCCCT
CTGGCGCATCCCCTTCTCCAGCAACTTCATGTCCATGGGTGCG-
CTCTCGGACCTGGGCCA-
GAACTTGCTCTACGCCAACTCCGCCCACGCCCTCGACATGACCTTC
GAGGTCGACCCCATGGACGAGCCCACCCTTCTCTATGTTCTGTT-
CGAAGTCTTTGACGTGG-
TCCGGGTCCACCAGCCGCACCGCGGCGTCATCGAGACCGTGTACCT
GCGTACGCCCTTCTCGGCCGGCAACGCCACCACCTAAAGAAG-
CAAGCCGCAGTCATCGCCG-
CCTGCATGCCGTGCGGTTCACCGAGCAAGAGCTCAGGGGCCATCGT
CAGAGACCTGGGATGCGGGGCCCTATTTTTTGGGCACCTTCGA-
CAAGCGCTTCCCTGGCTTTG-
TCTCCCCACACAAGCTGGCCTGCGCCATCGTCAACACGGCCGGCCG
CGAGACCGGGGGCGTGCACTGGCTGGCCTTCGCCTGGAACCCGCG-
CTCCAAAACATGCTT-
CCTCTTTGACCCCTTCGGCTTTTCGGACCAGCGGCTCAAGCAAATCT
ACGAGTTCGAGTACGAGGGCTTGCTGCGTCGCAGCGCCATCG-
CCTCCTCGCCCGACCGCTG-
CGTCACCCTCGAAAAGTCCACCCAGACCGTGCAGGGGGCCCGACTCG
GCCGCCTGCGGTCTCTTCTGCTGCATGTTTCTGCACGCCTTTGTG-
CACTGGCCTCAGAG-
TCCCATGGACCGCAACCCCAACCATGAACTTGCTGACGGGGGTGCCC
AACTCCATGCTCCAGAGCCCCCAGGTCGAGCCCACCCTGCGCCG-
CAACCAGGAGCAGCTCTA-
CAGCTTCCTGGAGCGCCACTCGCCTTACTTCCGCCGCCACAGCGCA
CAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAGAGATG-
CAAGAAGGGTAATAACGATG-
TACACACTTTTTTTCTCAATAAATGGCATCTTTTTTATTTATAACAAGC

TCTCTGGGGTATTCATTTCCCACCACCACCCGCCGTTGTGCG-
CCATCTGGCTCTATTTAGAAA-
TCGAAAGGGTTCTGCCGGGAGTCGCCGTGCGCCACGGGCAGGGAC
ACGTTGCGATACTGGTAGCGGGTGCCCCACTTGA ACTCGGGCAC-
CACCAGGCGAGGCAG-
CTCGGGGAAGTTTTCGCTCCACAGGCTGCGGGTCAGCACCAGCGCG
TTCATCAGGTCGGGCGCCGAGATCTTGAAGTCGCAGTTGGGGCCG-
CCGCCCTGCGCGCG-
CGAGTTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCGC
CGGGTGCTTCACGCTGGCCAGCACGCTGCGGTTCGGAGATCAG-
CTCGGCGTCCAGGTCCTCCG-
CGTTGCTCAGCGCGAACGGGGTCATCTTGGGCACTTGCCGCCCCAG
GAAGGGCGCGTGCCCCGGTTTCGAGTTGCAGTCGCAGCGCAGCGG-
GATCAGCAGGTGCCCCG-
TGCCCGGACTCGGCGTTGGGGTACAGCGCGCGCATGAAGGCCTGC
ATCTGGCGGAAGGCCATCTGGGCCTTGGCGCCCTCCGAGAAGAA-
CATGCCGCAGGACTTG-
CCCGAGA ACTGGTTTGCGGGGCAGCTGGCGTCGTGCAGGCAGCAG
CGCGCGTCGGTGTTGGCGATCTGCACCACGTTGCGCCCCCACC GG-
TTCTTCACGATCTTGG-
CCTTGGACGATTGCTCCTTCAGCGCGCGCTGCCCGTTCTCGCTGGTC
ACATCCATCTCGATCACATGTTCTTGTTCACCATGCTGCTGCCG-
TGCAGACACTTCAG-
CTCGCCCTCCGTCTCGGTGCAGCGGTGCTGCCACAGCGCGCAGCCC
GTGGGCTCGAAAGACTTGTAGGTCACCTCCGCGAAGGACTGCAGG-
TACCCCTGCAAAAAG-
CGGCCCATCATGGTCACGAAGGTCTTGTGCTGCTGAAGGTCAGCT
GCAGCCCGCGGTGCTCCTCGTTCAGCCAGGTCTTGCACACGGCCG-
CCAGCGCCTCCACCTGG-
TCGGGCAGCATCTTGAAGTTCACCTTCAGCTCATTCTCCACGTGGTA

CTTGTCCATCAGCGTGCGCGCCGCCTCCATGCCCTTCTCCCAGG-
CCGACACCAGCGGCAGG-
CTCACGGGGTTCTTCACCATCACCGTGGCCGCCGCCTCCGCCGCGC
TTTCGCTTTCCGCCCCGCTGTTCTCTTCCTCTTCCTCCTCTTCCTCG-
CCGCCGCCCACTCG-
CAGCCCCCGCACACGGGGTCGTCTTCCTGCAGGCGCTGCACCTTG
CGCTTGCCGTTGCGCCCCCTGCTTGATGCGCACGGGCGGGTTG-
CTGAAGCCCACCATCACAG-
CGCGGCCTCTTCTTGCTCGTCCTCGCTGTCCAGAATGACCTCCGGGG
AGGGGGGGTTGGTCATCCTCAGTACCGAGGCACGCTTCTTTTTCTT-
CCTGGGGGCGTTCG-
CCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTCGAAGGCCGAGGGC
TGGGCGTGCGCGGCACCAGCGCGTCCTGCGAGCCGTCCTCG-
TCCTCCTCGGACTCGAGACG-
GAGGCGGGCCCGCTTCTTCGGGGGCGCGCGGGGCGGCGGAGGCGG
CGGCGGCGACGGAGACGGGGACGAGACATCGTCCAGGGTGGGTG-
GACGGCGGGCCGCGCCG-
CGTCCGCGCTCGGGGGTGGTCTCGCGCTGGTCCTCTTCCCGACTGG
CCATCTCCCACTGCTCCTTCTCCTATAGGCAGAAAGAGATCATGGA-
GTCTCTCATGCGAG-
TCGAGAAGGAGGAGGACAGCCTAACCGCCCCCTCTGAGCCCTCCA
CCACCGCCGCCACCACCGCCAATGCCGCCGCGGACGACGCG-
CCCACCGAGACCACCGCCAG-
TACCACCCTCCCCAGCGACGCACCCCCGCTCGAGAATGAAGTGCTG
ATCGAGCAGGACCCGGGTTTTGTGAGCGGAGAGGAGGATGAGGTG-
GATGAGAAGGAGAAGGA-
GGAGGTCGCCGCCTCAGTGCCAAAAGAGGATAAAAAGCAAGACCA
GGACGACGCAGATAAGGATGAGACAGCAGTCGGGCGGGGGAACG-
GAAGCCATGATG-
CTGATGACGGCTACCTAGACGTGGGAGACGACGTGCTGCTTAAGCA

CCTGCACCGCCAGTGCGTCATCGTCTGCGACGCGCTGCAGGAGCG-
CTGCGAAGTGCCCCTG-
GACGTGGCGGAGGTCAGCCGCGCCTACGAGCGGCACCTCTTCGCGC
CGCACGTGCCCCCAAGCGCCGGGAGAACGGCACCTGCGAG-
CCCAACCCGCGTCTCAACTTC-
TACCCGGTCTTCGCGGTACCCGAGGTGCTGGCCACCTACCACATCT
TTTTCCAAAAGTCAAGATCCCCCTCTCCTGCCGCGCCAACCGCAC-
CCGCGCCGACAAAAC-
CCTGACCCTGCGGCAGGGCGCCACATACCTGATATCGCCTCTCTG
GAGGAAGTGCCCAAGATCTTCGAGGGTCTCGGTCGCGA-
CGAGAAACGGGCGGCGAACG-
CTCTGCACGGAGACAGCGAAAACGAGAGTCACTCGGGGGTGCTGG
TGGAGCTCGAGGGCGACAACGCGCGCCTGGCCGTACTCAAGCGCA-
GCATAGAGGTCACCCAC-
TTTGCCTACCCGGCGCTCAACCTGCCCCCAAGGTCATGAGTGTGG
TCATGGGCGAGCTCATCATGCGCCGCGCCCAGCCCCTGGCCGCG-
GATGCAAACCTTGCAAGAG-
TCCTCCGAGGAAGGCCTGCCCCGCGGTCAGCGACGAGCAGCTGGCG
CGCTGGCTGGAGACCCGCGACCCCGCGCAGCTGGAGGAGCGGCG-
CAAGCTCATGATGGCCG-
CGGTGCTGGTCACCGTGGAGCTCGAGTGTCTGCAGCGCTTCTTCGC
GGACCCCGAGATGCAGCGCAAGCTCGAGGAGACCCTGCACTACAC-
CTTCCGCCAGGGCTACG-
TGCGCCAGGCCTGCAAGATCTCCAACGTGGAGCTCTGCAACCTGGT
CTCCTACCTGGGCATCCTGCACGAGAACCGCCTCGGGCAGAACG-
TCCTGCACTCCACCCT-
CAAAGGGGAGGCGCGCCGCGACTACATCCGCGACTGCGCCTACCT
CTTCCTCTGCTACACCTGGCAGACGGCCATGGGGGTCTGGCAGCA-
GTGCCTGGAGGAGCG-
CAACCTCAAGGAGCTGGAAAAGCTCCTCAAGCGCACCCCTCAGGGA

CCTCTGGACGGGCTTCAACGAGCGCTCGGTGGCCGCCGCGCTGG-
CGGACATCATCTTT-
CCCGAGCGCCTGCTCAAGACCCTGCAGCAGGGCCTGCCCCGACTTCA
CCAGCCAGAGCATGCTGCAGAACTTCAGGACTTTCATCCTGGAG-
CGCTCGGGCATCCTG-
CCGGCCACTTGCTGCGCGCTGCCCAGCGACTTCGTGCCCATCAAGT
ACAGGGAGTGCCCGCCGCGCTCTGGGGCCACTGCTACCTCTTCCA-
GCTGGCCAACTAC-
CTCGCCTACCACTCGGACCTCATGGAAGACGTGAGCGGCGAGGGC
CTGCTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCCCACCG-
CTCTCTAGTCTGCAAC-
CCGCAGCTGCTCAGCGAGAGTCAGATTATCGGTACCTTCGAGCTGC
AGGGTCCCTCGCCTGACGAGAAGTCCGCGGGCTCCAGGGCTGAAAC-
TCACTCCGGGGCTGTG-
GACTTCCGCCTACCTACGCAAATTTGTACCTGAGGACTACCACGCC
CACGAGATCAGGTTCTACGAAGACCAATCCCGCCCCGCCCAAGGCG-
GAGCTCACCGCCTGCGT-
CATCACCCAGGGGCACATCCTGGGGCAATTGCAAGCCATCAACAA
AGCCCGCCGAGAGTTCTTGCTGAAAAAGGGTCGGGGGGGTGTAC-
CTGGACCCCCAGTCCGG-
CGAGGAGCTAAACCCGCTACCCCCGCCGCCGCCAGCAGCGGGA
CCTTGCTTCCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGCCG-
CCGCCGCCGCAGCCATA-
CATGCTTCTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAGA
GGAGGTTTCGGACGAGGAGCAGGAGGAGATGATGGAAGACTGG-
GAGGAGGACAGCAGCCTA-
GACGAGGAAGCTTCAGAGGCCGAAGAGGTGGCAGACGCAACACCA
TCGCCCTCGGTTCGCAGCCCCCTCGCCGGGGCCCCTGAAA-
TCCTCCGAACCCAGCACCCAGCGC-
TATAACCTCCGCTCCTCCGGCGCCGGCGCCACCCGCCCGCAGACCC

AACCGTAGATGGGACACCACAGGAACCGGGGTCGGTAAG-
TCCAAGTGCCCCGCCGCCGCCAC-
CGCAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGCTCGTGGCGCG
GGCACAAGAACGCCATAGTCGCCTGCTTGCAAGACTGCGGGGG-
CAACATCTCTTTCGCCCCG-
CCGCTTCCTGCTATTCCACCACGGGGTCGCCTTTCCCCGCAATGTCC
TGCATTACTACCGTCATCTCTACAGCCCCTACTGCAGCGGGCGAC-
CCAGAGGCGGCAGCGGCA-
GCCACAGCGGCGACCACCACCTAGGAAGATATCCTCCGCGGGCAA
GACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCAGCGGGCGG-
GAGCGGTGGGCGCACTG-
CGCCTCTCGCCCAACGAACCCCTCTCGACCCGGGAGCTCAGACACA
GGATCTTCCCCACTTTGTATGCCATCTTCCAACAGAGCAGAGGCCA-
GGAGCAGGAGCTGAAA-
ATAAAAAACAGATCTCTGCGCTCCCTCACCCGCAGCTGTCTGTATC
ACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAGGACGCGGA-
GGCACTCTTCAGCAAATAC-
TGCGCGCTCACTCTTAAAGACTAGCTCCGCGCCCTTCTCGAATTTAG
GCGGGAGAAAACACTACGTCATCGCCGGCCGCCGCCAGCCCGCCCA-
GCCGAGATGAGCAAAGA-
GATTCCCACGCCATACATGTGGAGCTACCAGCCGCAGATGGGACTC
GCGGCGGGAGCGGCCCCAGGACTACTCCACCCGCATGAACTACA-
TGAGCGCGGGACCCCACA-
TGATCTCACAGGTCAACGGGATCCGCGCCCAGCGAAACCAAATACT
GCTGGAACAGGCGGCCATCACCGCCACGCCCCGCCATAATCT-
CAACCCCCGAAATTGGCCCCG-
CCGCCCTCGTGTACCAGGAAACCCCTCCGCCACCACCGTACTACT
TCCGCGTGACGCCAGGCCGAAGTCCAGATGACTAACTCAGGGG-
CGCAGCTCGCGGGCGG-
CTTTCGTCACGGGGCGCGGCCGCTCCGACCAGGTATAAGACACCTG

ATGATCAGAGGCCGAGGTATCCAGCTCAACGACGAGTCGGTGAG-
CTCTTCGCTCGGTCTCCG-
TCCGGACGGAAC TTTCCAGCTCGCCGGATCCGGCCGCTCTTCGTTC
ACGCCCCGCCAGGCGTACCTGACTCTGCAGACCTCGTCCTCGGAG-
CCCCGCTCCGGCGG-
CATCGGAACCCTCCAGTTCGTGGAGGAGTTCGTGCCCTCGGTCTAC
TTCAACCCCTTCTCGGGACCTCCCGGACGCTACCCCGACCAGTT-
CATTCCGAAC TTTGACG-
CGGTGAAGGACTCGGCGGACGGCTACGACTGAATGTCAGGTGTCG
AGGCAGAGCAGCTTCGCCTGAGACACCTCGAGCACTGCCGCCG-
CCACAAGTGCTTCGCCCCG-
CGGTTCTGGTGAGTTCTGCTACTTTTCAGCTACCCGAGGAGCATACC
GAGGGGCCGGCGCACGGCGTCCGCCTGACCACCCAGGGCGAGGT-
TACCTGTTCCCT-
CATCCGGGAGTTTACCCTCCGTCCCCTGCTAGTGGAGCGGGAGCGG
GGTCCCTGTGTCCTAACTATCGCCTGCAACTGCCCTAACCTGGAT-
TACATCAAGATCTTTG-
CTGTCATCTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAATCTA
CTGGGGCTCCTGTCGCCATCCTGTGAACGCCACCGTCTTCACCCAC-
CCCGACCAGGCCAGG-
CGAACCTCACCTGCGGTCTGCATCGGAGGGCCAAGAAGTACCTCAC
CTGGTACTTCAACGGCACCCCCTTTGTGGTTTACAACAGCTTCGA-
CGGGGACGGAG-
TCTCCCTGAAAGACCAGCTCTCCGGTCTCAGCTACTCCATCCACAA
GAACACCACCCTCCAAC TTTCCCTCCCTACCTGCCGGGAACC-
TACGAGTGCGTCACCGG-
CCGCTGCACCCACCTCACCCGCCTGATCGTAAACCAGAGCTTTCCG
GGAACAGATAACTCCCTCTTCCCCAGAACAGGAGGTGAGCTCAG-
GAAACTCCCCGGGGACCA-
GGGCGGAGACGTACCTTCGACCCTTGTGGGGTTAGGATTTTTTTATT

ACCGGGTTGCTGGCTCTTTTAATCAAAGTTTCCTTGAGATTTGTT-
CTTTCCTTCTACGTGTA-
TGAACACCTCAACCTCCAATAACTCTACCCTTTCTTCGGAATCAGGT
GACTTCTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTG-
TTGATTTTTTTCCTTATCATACT-
CAGCCTTCTGTGCCTCAGGCTCGCCGCCTGCTGCGCACACATCTAT
ATCTACTGCTGGTTGCTCAAGTGCAGGGGTCGCCACCCAAGA-
TGAACAGGTACATGGTCCTA-
TCGATCCTAGGCCTGCTGGCCCTGGCGGCCTGCAGCGCCGCCAAAA
AAGAGATTACCTTTGAGGAGCCCGCTTGCAATGTAACCTTTCAAG-
CCCGAGGGTGACCAATG-
CACCACCCTCGTCAAATGCGTTACCAATCATGAGAGGCTGCGCATC
GACTACAAAAACAAAACCTGGCCAGTTTGCGGTCTATAGTGTGTT-
TACGCCCCGGAGACCCCTC-
TAACTACTCTGTCACCGTCTTCCAGGGCGGACAGTCTAAGATATTC
AATTACACTTTCCCTTTTTATGAGTTATGCGATGCGGTCATGTACA-
TGTCAAAACAGTA-
CAACCTGTGGCCTCCCTCTCCCCAGGCGTGTGTGGAAAATACTGGG
TCTTACTGCTGTATGGCTTTCGCAATCACTACGCTCGCTCTAATCTG-
CACGGTGCTATACA-
TAAAATTCAGGCAGAGGCGAATCTTTATCGATGAAAAGAAAATGC
CTTGATCGCTAACACCGGCTTCTATCTGCAGAATGAATGCAAT-
CACCTCCCTACTAATCAC-
CACCACCCTCCTTGCGATTGCCCATGGGTTGACACGAATCGAAGTG
CCAGTGGGGTCCAATGTCACCATGGTGGGCCCCGCCGGCAATT-
CCACCCTCATGTGGGAAAA-
ATTTGTCCGCAATCAATGGGTTCAATTTCTGCTCTAACCGAATCAGTA
TCAAGCCCAGAGCCATCTGCGATGGGCAAAATCTAACTCTGATCA-
ATGTGCAAATGATGGA-
TGCTGGGTACTATTACGGGCAGCGGGGAGAAATCATTAATTACTGG

CGACCCCAAGGACTACATGCTGCATGTAGTCGAGGCACTT-
CCCCTACCACCCCACTAC-
CACCTCTCCCACCACCACCACCCTACTACTACTACTACTACTA
CTACTACTACCCTACCGCTGCCCCGCCATACCCGCAAAAGCAC-
CATGATTAGCACAAAG-
CCCCCTCGTGCTCACTCCCACGCCGGCGGGCCCATCGGTGCGACCT
CAGAAACCACCGAGCTTTGCTTCTGCCAATGCACTAACGCCAGCG-
CTCATGAACTGTTCGAC-
CTGGAGAATGAGGATGTCCAGCAGAGCTCCGCTTGCCTGACCCAGG
AGGCTGTGGAGCCCGTTGCCCTGAAGCAGATCGGTGATTCAATAA-
TTGACTCTTCTT-
CTTTTGCCACTCCCGAATACCCTCCCGATTCTACTTTCCACATCACG
GGTACCAAAGACCCTAACCTCTCTTTCTACCTGATGCTGCTG-
CTCTGTATCTCTGTGG-
TCTCTTCCGCGCTGATGTTACTGGGGATGTTCTGCTGCCTGATCTGC
CGCAGAAAGAGAAAAGCTCGCTCTCAGGGCCAACCACTGATG-
CCCTTCCCCTACCCCCCGGA-
TTTTGCAGATAACAAGATATGAGCTCGCTGCTGACACTAACCGCTT
TACTAGCCTGCGCTCTAACCCTTGTCGCTTGCGACTCGAGATTCCA-
CAATGTCACAGCTG-
TGGCAGGAGAAAATGTTACTTTCAACTCCACGGCCGATACCCAGTG
GTCGTGGAGTGGCTCAGGTAGCTACTTAACTATCTGCAATAG-
CTCCACTTCCCCCGGCATA-
TCCCCAACCAAGTACCAATGCAATGCCAGCCTGTTACCCTCATCA
ACGCTTCCACCCTGGACAATGGACTCTATGTAGGCTATGTAC-
CCTTTGGTGGGCAAGGAAA-
GACCCACGCTTACAACCTGGAAGTTCGCCAGCCCAGAACCACTACC
CAAGCTTCTCCCACCACCACCACCACCACCACCATCACCAGCAG-
CAGCAGCAGCAGCAGCCA-
CAGCAGCAGCAGCAGATTATTGACTTTGGTTTTGGCCAGCTCATCT

GCCGCTACCCAGGCCATCTACAGCTCTGTGCCCCGAAACCACTCA-
GATCCACCGCCCA-
GAAACGACCACCGCCACCACCCTACACACCTCCAGCGATCAGATGC
CGACCAACATCACCCCCTTGGCTCTTCAAATGGGACTTACAAG-
CCCCACTCCAAAACCAGTG-
GATGCGGCCGAGGTCTCCGCCCTCGTCAATGACTGGGCGGGGCTGG
GAATGTGGTGGTTCGCCATAGGCATGATGGCGCTCTGCCTGCTT-
CTGCTCTGGCTCATCTG-
CTGCCTCCACCGCAGGCGAGCCAGACCCCCCATCTATAGACCCATC
ATTGTCCTGAACCCCGATAATGATGGGATCCATAGATTGGATGG-
CCTGAAAAACCTAC-
TTTTTTCTTTTACAGTATGATAAATTGAGACATGCCTCGCATTTTCT
TGTACATGTTCTTCTCCACCTTTTCTGGGGTGTTCTACGCTGG-
CCGCTGTGTCTCACCTG-
GAGGTAGACTGCCTCTCACCTTCACTGTCTACCTGCTTTACGGATT
GGTCACCCTCACTCTCATCTGCAGCCTAATCACAGTAATCATCG-
CCTTCATCCAGTG-
CATTGATTACATCTGTGTGCGCCTCGCATACTTCAGACACCACCCG
CAGTACCGAGACAGGAACATTGCCCAACTTCTAAGACTGCTCTAA-
TCATGCATAAGACTG-
TGATCTGCCTTCTGATCCTCTGCATCCTGCCACCCTCACCTCCTGC
CAGTACACCACAAAATCTCCGCGCAAAAGACATGCCTCCTGCCG-
CTTCACCCAACTGTGGAA-
TATACCCAAATGCTACAACGAAAAGAGCGAGCTCTCCGAAGCTTG
GCTGTATGGGGTCATCTGTGTCTTAGTTTTCTGCAGCACTGTCTTTG-
CCCTCATAATCTAC-
CCCTACTTTGATTTGGGATGGAACGCGATCGATGCCATGAATTACC
CCACCTTTCCCGCACCCGAGATAATTCCACTGCGACAAGTTGTAC-
CCGTTGTCGTTAAT-
CAACGCCCCCATCCCCTACGCCCCTGAAATCAGCTACTTTAACC

TAACAGGCGGAGATGACTGACGCCCTAGATCTAGAAATGGACGG-
CATCAGTACCGAGCAGCG-
TCTCCTAGAGAGGCGCAGGCAGGCGGCTGAGCAAGAGCGCCTCAA
TCAGGAGCTCCGAGATCTCGTTAACCTGCACCAGTGCAAAAGAGG-
CATCTTTTGTCTGG-
TAAAGCAGGCCAAAGTCACCTACGAGAAGACCGGCAACAGCCACC
GCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCTGGTGCT-
CATGGTGGGTGAGAA-
TCCCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGTCTG
CACTCCCCCTGTCGGGGTCCAGAAGACCTCTGCACCCTGGTAAA-
GACCCTGTGCGGTCTCA-
GAGATTTAGTCCCCTTTAACTAATCAAACACTGGAATCAATAAAAA
GAATCACTTACTTAAAATCAGACAGCAGGTCTCTGTCCAGTTTATT-
CAGCAGCACCTCCTT-
CCCCTCCTCCCAACTCTGGTACTCCAAACGCCTTCTGGCGGCAAAC
TTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTG-
TCCCTCCGCACCCACTATCTT-
CATGTTGTTGCAGATGAAGCGCACCAAAACGTCTGACGAGAGCTTC
AACCCCGTGTACCCCTATGACACGGAAAGCGGCCCTCCCTCCG-
TCCCTTTCCTCAC-
CCCTCCCTTCGTGTCTCCCGATGGATTCCAAGAAAGTCCCCCGGG
GTCCTGTCTCTGAACCTGGCCGAGCCCCTGGTCACTTCCCACGG-
CATGCTCGCCCTGAAAA-
TGGGAAGTGGCCTCTCCCTGGACGACGCTGGCAACCTCACCTCTCA
AGATATCACCAACCGCTAGCCCTCCCCTCAAAAAAACCAAGAC-
CAACCTCAGCCTAGAAACCT-
CATCCCCCCTAACTGTGAGCACCTCAGGCGCCCTCACCGTAGCAGC
CGCCGCTCCCCTGGCGGTGGCCGGCACCTCCCTCACCATGCAATCA-
GAGGCCCCCCTGACAG-
TACAGGATGCAAAACTCACCTGGCCACCAAAGGCCCCCTGACCGT

GTCTGAAGGCAAACCTGGCCTTGCAAACATCGGCCCCGCTGACGG-
CCGCTGACAGCAGCAC-
CCTCACAGTCAGTGCCACACCACCCCTTAGCACAAGCAATGGCAGC
TTGGGTATTGACATGCAAGCCCCCATTTACACCACCAATGGA-
AAACTAGGACTTAACTTTGG-
CGCTCCCCTGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTT
ACTGGCCAAGGTCTTACGATAAACGGAACAGCCCTACAAAC-
TAGAGTCTCAGGTGCCCT-
CAACTATGACACATCAGGAAACCTAGAATTGAGAGCTGCAGGGGG
TATGCGAGTTGATGCAAATGGTCAACTTATCCTTGATGTAGCTTAC-
CCATTTGATGCA-
CAAAACAATCTCAGCCTTAGGCTTGGACAGGGACCCCTGTTTGTTA
ACTCTGCCCACAACCTTGGATGTAACTACAACAGAGGCCTCTAC-
CTGTTCACATCTGGAAA-
TACCAAAAAGCTAGAAGTTAATATCAAAACAGCCAAGGGTCTCATT
TATGATGACACTGCTATAGCAATCAATGCGGGTGATGGGCTACAG-
TTTGACTCAGGCTCAGA-
TACAAATCCATTAAAACTAACTTGGATTAGGACTGGATTATGAC
TCCAGCAGAGCCATAATTGCTAACTGGGAACTGGCCTAAG-
CTTTGACAACACAGGTGCCAT-
CACAGTAGGCAACAAAAATGATGACAAGCTTACCTTGTGGACCAC
ACCAGACCCATCCCCTAACTGTAGAATCTATTCAGAGAAAGATGC-
TAAATTCACACTTG-
TTTTGACTAAATGCGGCAGTCAGGTGTTGGCCAGCGTTTCTGTTTTA
TCTGTAAAAGGTAGCCTTGCGCCCATCAGTGGCACAGTAACTAG-
TGCTCAGATTGTCCTCA-
GATTTGATGAAAATGGAGTTCTACTAAGCAATTCTTCCCTTGACCCT
CAATACTGGAACCTACAGAAAAGGTGACCTTACAGAGGGCACTGCA-
TATACCAACGCAGTGG-
GATTTATGCCCAACCTCACAGCATACCCAAAAACACAGAGCCAAA

CTGCTAAAAGCAACATTGTAAGTCAGGTTTACTTGAATGGGGA-
CAAATCCAAACCCATGAC-
CCTCACCATTACCCTCAATGGAACTAATGAAACAGGAGATGCCACA
GTAAGCACTTACTCCATGTCATTCTCATGGAACTGGAATGGAAGTA-
ATTACATTAA-
TGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCGCCCAA
GAATAAAAAGCATGACGCTGTTGATTTGATTCAATGTGTTTCTG-
TTTTATTTTCAAGCACAA-
CAAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACACAGT
AGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAAC-
TAGTGGAGAAGTACTCG-
CCTACATGGGGGTAGAGTCATAATCGTGCATCAGGATAGGGCGGT
GGTGCTGCAGCAGCGCGCGAATAAACTGCTGCCGCCGCCGCTCCG-
TCCTGCAGGAATACAA-
CATGGCAGTGGTCTCCTCAGCGATGATTCGCACCGCCCGCAGCATA
AGGCGCCTTGTCCTCCGGGCACAGCAGCGCACCCCTGATCTCACT-
TAAATCAGCACAGTAAC-
TGCAGCACAGCACCACAATATTGTTCAAAATCCCACAGTGCAAGGC
GCTGTATCCAAAGCTCATGGCGGGGACCACAGAACCCACGTGG-
CCATCATACCACAAGCGCA-
GGTAGATTAAGTGGCGACCCCTCATAAACACGCTGGACATAAACAT
TACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGTACCATA-
TAAACCTCTGATTAAACA-
TGGCGCCATCCACCACCATCCTAAACCAGCTGGCCAAAACCTGCCC
GCCGGCTATACACTGCAGGGAACCGGGACTGGAACAATGACAGTG-
GAGAGCCCAGGACTCG-
TAACCATGGATCATCATGCTCGTCATGATATCAATGTTGGCACAAC
ACAGGCACACGTGCATACACTTCCTCAGGATTACAAGCTCCTCCCG-
CGTTAGAACCATA-
TCCCAGGGAACAACCCATTCTGAATCAGCGTAAATCCCACACTGC

AGGGAAGACCTCGCACGTAACCTCACGTTGTGCATTGTCAAAGTGT-
TACATTCGGGCAGCAG-
CGGATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAAAGGA
GGTAGACGATCCCTACTGTACGGAGTGCGCCGAGACAACCGAGA-
TCGTGTTGGTCGTAGTGT-
CATGCCAAATGGAACGCCGGACGTAGTCATATTTCTGAAGTCTTA
GATCTCTCAACGCAGCACCAGCACCAACACTTCGCAGTG-
TAAAAGGCCAAGTGCCGAGAGAG-
TATATATAGGAATAAAAAGTGACGTAAACGGGCAAAGTCCAAAAA
ACGCCCAGAAAAACCGCACGCGAACCTACGCCCCGAAACGAAAG-
CCAAAAAACACTAGACAC-
TCCCTTCCGGCGTCAACTTCCGCTTTCCCACGCTACGTCACTTGCCC
CAGTCAAACAACTACATATCCCGAACTTCCAAGTCGCCACG-
CCCAAAACACCGCCTACAC-
CTCCCCGCCCCGCCGGCCCCGCCCCCAAACCCGCCTCCCGCCCCGCGC
CCCGCCCCGCGCCGCCCATCTCATTATCATATTGGCTTCAA-
TCCAAAATAAGGTATATTATTGATGATG

SEQ ID NO: 5 Sequência de aminoácidos de RSV F0 Δ TM-N-M2-1

MELLILKANAITTILTAVTFCFASGQNITEEFYQSTCSAVSKGYLSAL-
RTGWYTSVITIELS-
NIKENKCNGTDAKVLIKQELDKYKNAVTELQLLMQSTPATNNRARR
ELPRFMNYTLNNAKKTNTLSKKRKRRFLGFLLGVSIAIAS-
GVAVSKVLHLEGEVN-
KIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYIDKQLLPIVNBKQSCSI
SNIETVIEFQQKNRLLLEITREFSVNAGVTTPVSTYMLTNSSELLSLNDM-
PITNDQKKLMSNN-
VQIVRQQSYSIMSIIKEEVLAYVVQLPLYGVIDTPCWKLHTSPLCTTNT
KEGSNICLTRDRGWYCDNAGSVSFFPQAETCKVQSNRVFCDTMNS-
LTLPSEVNLCNVDIFN-
PKYDCKIMTSKTDVSSSVITSLGAIVSCYGKTKCTASNKNRGIKTFNS

GCDYVSNKGVDTVSVGNTLYYVKNQEGKSLYVKGEPIINFYD-
 PLVFPSDEFDASIS-
 QVNEKINQSLAFIRKSDELLHNVNAGKSTTNRKRRAPVKQTLNFDLLK
 LAGDVESNPGPMALSKVKLNDTLNKDQLLSSSKYTIQRS-
 TGDSIDTPNYDVQKHINKLCG-
 MLLITEDANHKFTGLIGMLYAMSRLGREDTIKILRDAGYHVKANGVD
 VTTHRQDINGKEMKFEVLTASLTTEIQINIEIESRKSYYKKMLKEMGE-
 VAPEYRHDSPPDCG-
 MIILCIAALVITKLAAGDRSGLTAVIRRANNVLKNEMKRYKGLLPKDI
 ANS
 FYEVFEKYPHFIDVFVHFGIAQSSTRGGSRVEGIFA-
 GLFMNAYGAGQVMLR-
 WGVLAHSVKNIMLGHASVQAEMEQVVEVYEYAQKLGGEAGFYHILN
 NPKASLLSLTQFPHFSSVVLGNAAGLGIMGEYRGTPRNQDLYDAAKA-
 YAEQLKEN-
 GVINYSVLDLTAEELAIKHQLNPKDNDVELGGGGSGGGGMSRRNPC
 KFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDK-
 SIDTLSEISGAAELDR-
 TEEYALGVGVVLESYIGSINNITKQSACVAMSKLLTELNSDDIKKLRDN
 EELNSPKIRVYNTVISYIESNRKNNKQTIHLLKRLPADVLKKTIKNTL-
 DIHKSITINNPKESTVSDTNDHAKNNDTT

SEQ ID NO: 6 Sequência polinucleotídica que codifica o promo-
tor de hCMV aprimorado

MELLILKANAITTILTAVTFCFASGQNITEEFYQSTCSAVSKGYLSAL-
 RTGWYTSVITIELS-
 NIKENKCNGTDAKVKLIKQELDKYKNAVTELQLLMQSTPATNNRARR
 ELPRFMNYTLNNAKKTNTLSKKRKRRFLGFLGVSASIAS-
 GVAVSKVLHLEGEVN-
 KIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYIDKQLLPVKNQSCSI

SNIETVIEFQQKNRLLLEITREFSVNAGVTTPVSTYMLTNSELLSLNDM-
PITNDQKKLMSNN-
VQIVRQQSYSIMSIIKEEVLAYVVQLPLYGVIDTPCWKLHTSPLCTTNT
KEGSNICLTRTRDGRWYCDNAGSVSFFPQAETCKVQSNRVFCDTMNS-
LTLPSEVNLCNVDIFN-
PKYDCKIMTSKTDVSSSVITSLGAIVSCYGKTKCTASNKNRGIKTFNS
GCDYVSNKGVDTVSVGNTLYYVVKQEGKSLYVKGEPIINFYD-
PLVFPSDEFDASIS-
QVNEKINQSLAFIRKSDELLHNVNAGKSTTNRKRRAPVKQTLNFDLLK
LAGDVESNPGPMALSKVKLNDTLNKDQLLSSSKYTIQRS-
TGDSIDTPNYDVQKHINKLCG-
MLLITEDANHKFTGLIGMLYAMSRLGREDTIKILRDAGYHVKANGVD
VTTHRQDINGKEMKFEVLTLASLTTEIQINIEIESRKSYYKMLKEMGE-
VAPEYRHDSPPDCG-
MIILCIAALVITKLAAGDRSGLTAVIRRANNVLKNEMKRYKGLLPKDI
ANS
FYEVFEKYPHFIDVFVHFGIAQSSTRGGSRVEGIFA-
GLFMNAYGAGQVMLR-
WGVLAHSVKNIMLGHASVQAEMEQVVEVYEYAQKLGGEAGFYHILN
NPKASLLSLTQFPHFSSVVLGNAAGLGIMGEYRGTPRNQDLYDAAKA-
YAEQLKEN-
GVINYSVLDLTAELEAIKHQLNPKDNDVELGGGGSGGGGMSRRNPC
KFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDK-
SIDTLSEISGAAELDR-
TEEYALGVVGVLESYIGSINNITKQSACVAMSKLLTELNSDDIKKLRDN
EELNSPKIRVYNTVISYIESNRKNNKQTIHLLKRLPADVLKKTIKNTL-
DIHKSITINNPKESTVSDTNDHAKNNDTT

**SEQ ID NO: 7 Sequência polinucleotídica que codifica o cas-
sete hCMV NM2 bgbpolyA**

CCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT-
 CATGTCCAACAT-
 TACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTAC
 GGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACT-
 TACGGTAAATGG-
 CCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACG
 TATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG-
 GAGTATT-
 TACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTAC
 GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA-
 GTACATGACCT-
 TATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC
 CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGA-
 CTCACGGGGA-
 TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA
 ATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAA-
 TGGGCGGTAGG-
 CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAG
 ATCTCCCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAACCGT-
 CAGATCGCCTG-
 GAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCC
 AGCCTCCGCGGGCGGGAACGGTGCATTGGAACGCGGATTCCCCGTG-
 CCAAGAGTGA_{GATCTT}-
 CCGTTTATCTAGGTACCAGATATCGCCACCATGGCCCTGAGCAAAGT_{GAAAC}
 TGAACGATACACTGAACAAGGACCAGCTGCTGTCCAGCAGCAAGTACAC-
 CATCCAGCGGAG-
 CACCGGCGACAGCATCGATACCCCCAACTACGACGTGCAGAAGCACATCAAC
 AAGCTGTGCGGCATGCTGCTGATCACAGAGGACGCCAACCACAAGTTCAC-
 CGGCCTGATCGG-
 CATGCTGTACGCCATGAGCCGGCTGGGCGGGAGGACACCATCAAGATCCTG

CGGGACGCCGGCTACCACGTGAAGGCCAATGGCGTGGACGTGACCACACAC-
CGGCAGGACAT-
CAACGGCAAAGAAATGAAGTTCGAGGTGCTGACCCTGGCCAGCCTGACCACC
GAGATCCAGATCAATATCGAGATCGAGAGCCGGAAGTCCTACAAGAAAATG-
CTGAAAGAAA-
TGGGCGAGGTGGCCCCCGAGTACAGACACGACAGCCCCGACTGCGGCATGAT
CATCCTGTGTATCGCCGCCCTGGTGATCACAAAGCTGGCCGCTGGCGACA-
GATCTGGCCTGA-
CAGCCGTGATCAGACGGGCCAACAATGTGCTGAAGAACGAGATGAAGCGGTA
CAAGGGCCTGCTGCCCAAGGACATTGCCAACAGCTTCTACGAGGTGTT-
CGAGAAGTAC-
CCCCACTTCATCGACGTGTTCTGTCGACTTCGGCATTGCCCAGAGCAGCACCA
GAGGCGGCTCCAGAGTGGAGGGCATCTTCGCCGGCCTGTTTCATGAACGCC-
TACGGCGCTGG-
CCAGGTGATGCTGAGATGGGGCGTGCTGGCCAAGAGCGTGAAGAACATCATG
CTGGGCCACGCCAGCGTGACGGCCGAGATGGAACAGGTGGTGGAGGTG-
TACGAGTACGCCCA-
GAAGCTGGGCGGAGAGGGCCGGCTTCTACCACATCCTGAACAACCCTAAGGCC
TCCCTGCTGTCCCTGACCCAGTTCCCCCACTTCTCCAGCGTGGTGCTGGGA-
AATGCCGCCG-
GACTGGGCATCATGGGCGAGTACCGGGGCACCCCCAGAAACCAGGACCTGTA
CGACGCCGCCAAGGCCTACGCCGAGCAGCTGAAAGAAAACGGCGTGAT-
CAACTACAGCGTG-
CTGGACCTGACCGCTGAGGAACTGGAAGCCATCAAGCACCAGCTGAACCCCA
AGGACAACGACGTGGAGCTGGGAGGCGGAGGATCTGGCGGCGGAGG-
CATGAGCAGACGGAAC-
CCCTGCAAGTTCGAGATCCGGGGCCACTGCCTGAACGGCAAGCGGTGCCACT
TCAGCCACAACACTACTTCGAGTGGCCCCCTCATGCTCTGCTGGTGCGGCA-
GAACTTCATG-
CTGAACCGGATCCTGAAGTCCATGGACAAGAGCATCGACACCCTGAGCGAGA

TCAGCGGAGCCGCCGAGCTGGACAGAACCGAGGAATATGCCCTGGGCGTGG-
TGGGAGTGCTG-
GAAAGCTACATCGGCTCCATCAACAACATCACAAAGCAGAGCGCCTGCGTGG
CCATGAGCAAGCTGCTGACAGAGCTGAACAGCGACGACATCAAGAAG-
CTGAGGGACAACGAG-
GAACTGAACAGCCCCAAGATCCGGGTGTACAACACCGTGATCAGCTACATTG
AGAGCAACCGCAAGAACAACAAGCAGACCATCCATCTGCTGAAGCGGCTG-
CCCGCCGACGTG-
CTGAAAAAGACCATCAAGAACACCCTGGACATCCACAAGTCCATCACCATCA
ACAATCCCAAAGAAAGCACCGTGTCTGACACCAACGATCACGCCAAGAA-
CAACGACACCAC-
CTGATGAGCGGCCGCGATCTGCTGTGCCTTCTAGTTGCCAGCCATCTGTTGT
TTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTG-
TCCTTTCCTAA-
TAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGG
GGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCA-
GGCATGCTGGGGATGCGGTGGGCTCTATGG

Sequência do promotor CMV: negrito

Sequência do transgene NM2: itálico

bghpolyA *sinál PolyA: itálico + sublinhado*

SEQ ID NO: 8 Sequência de proteínas NM2

MALSKVKLNDTLNKDQLLSSSKYTIQIRSTGDSIDTPNYDVQKHINKLCG-
MLLITEDANHK-
FTGLIGMLYAMSRLGREDTIKILRDAGYHVKANGVDVTTHRQDINGKEMKFE
VLTLASLTTEIQINIEIESRKSYYKMLKEMGEVAPEYRHDSPDCGMIIIL-
CIAALVITKLAAG-
DRSGLTAVIRRANNVLKNEMKRYKGLLPKDIANSFYEVFEKYPHFIDVVFVHF
GIAQSSTRGGSRVEGIFAGLFMNAYGAGQVMLRWGVLAHSVKNIMLGHAS-
VQAE-
MEQVVEVYEYAQKLGGEAGFYHILNNPKASLLSLTQFPHFSSVVLGNAAGLG

IMGEYRGTPRNQDLYDAAKAYAEQLKENGVINYSVLDLTAELEEAIKHQLN-
PKDNDVEL-

GGGGSGGGGMSRRNPCKFEIRGHCLNGKRCHFESHNYFEWPPHALLVRQNFML
NRILKSMDKSIDTLSEISGAAELDRTEEYALGVVGVLESYIGSINNIT-
KQSACVAMSKLL-

TELNSDDIKKLRDNEELNSPKIRVYNTVISYIESNRKNNKQTIHLLKRLPAD
VLKKTIKNTLDIHKSITINNPKESTVSDTNDHAKNNDTT

**SEQ ID NO: 9 Sequência polinucleotídica que codifica o cas-
sete hCMV F0 WPRE bghpolyA**

CCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT-
CATGTCCAACAT-

TACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTAC
GGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACT-
TACGGTAAATGG-

CCCGCCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACG
TATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG-
GAGTATT-

TACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTAC
GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA-
GTACATGACCT-

TATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC
CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGA-
CTCACGGGGA-

TTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAA
ATCAACGGGACTTTCCAAATGTCGTAACAACTCCGCCCCATTGACGCAAA-
TGGGCGGTAGG-

CGTGTACGGTGGGAGGTCTATATAAGGCGAAGCGCTCCCTATCAGTGATAGA
GATCTCCCTATCAGTGATAGAGATCGTCGACGAGCTCGCGGCGGGCGGGAG-
TCGCTGCGCG-

CTGCCTTCGCCCCGTGCCCCGCTCCGCCGCCGCTCGCGCCGCCCGCCCCGG

**CTCTGACTGACCGCGTTACTAAAACAGGTAAGTCCGGCCTCCGCGCCGGG-
TTTTGGCG-
CCTCCCGCGGGCGCCCCCTCCTCACGGCGAGCGCTGCCACGTCAGACGAAG
GGCGCAGCGAGCGTCCTGATCCTTCCGCCCCGACGCTCAGGACAGCGG-
CCCGCTGCTCATAA-
GACTCGGCCTTAGAACCCCAGTATCAGCAGAAGGACATTTTAGGACGGGACT
TGGGTGACTCTAGGGCACTGGTTTTCTTTCCAGAGAGCGGAACAGGCGAG-
GAAAAGTAG-
TCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTGGGGCGGTGAACGCCGAT
GATGCCTCTACTAACCATGTTTCATGTTTTCTTTTTTTTTTCTACAGG-
TCCTGGGTGACGAACA-
GGATATCGCCACCATGGAACTGCTGATCCTGAAGGCCAACGCCATCACCACC
ATCCTGACCGCCGTGACCTTCTGCTTCGCCAGCGGCCAGAACATCACCGAG-
GAATTCTACCA-
GAGCACCTGTAGCGCCGTGAGCAAGGGCTACCTGAGCGCCCTGAGAACCGGC
TGGTACACCAGCGTGATCACCATCGAGCTGAGCAACATCAAAGAAAACAAG-
TGCAACGGCAC-
CGACGCCAAAGTGAAGCTGATCAAGCAGGAACTGGACAAGTACAAGAACGCC
GTGACCGAGCTGCAGCTGCTGATGCAGAGCACCCCCGCCACCAACAAC-
CGGGCCAGACGGGA-
GCTGCCCCGGTTCATGAACTACACCCTGAACAACGCCAAAAAGACCAACGTG
ACCCTGAGCAAGAAGCGGAAGCGGCGGTTCTCTGGGCTTTCTGCTGGGCG-
TGGGCAGCG-
CCATTGCCAGCGGCGTGGCCGTGTCTAAGGTGCTGCACCTGGAAGGCGAAGT
GAACAAGATCAAGAGCGCCCTGCTGAGCACCAACAAGGCCGTGGTG-
TCCCTGAGCAACGGCG-
TGAGCGTGCTGACCAGCAAGGTGCTGGATCTGAAGAACTACATCGACAAGCA
GCTGCTGCCCATCGTGAACAAGCAGAGCTGCAGCATCAGCAACATCGAGA-
CAGTGATCGAG-
TTCCAGCAGAAGAACAACCGGCTGCTGGAAATCACCCGGGAGTTCAGCGTGA**

ACGCCGGCGTGACCACCCCTGTGTCCACCTACATGCTGACCAACAGCGAG-
CTGCTGAG-
CCTGATCAACGACATGCCCATCACCAACGACCAGAAAAAGCTGATGAGCAAC
AACGTGCAGATCGTGCGGCAGCAGAGCTACTCCATCATGTCCATCAT-
CAAAGAAGAGGTG-
CTGGCCTACGTGGTGCAGCTGCCCCTGTACGGCGTGATCGACACCCCCTGCT
GGAAGCTGCACACCAGCCCCCTGTGCACCACCAACACCAAAGAGGGCAG-
CAACATCTG-
CCTGACCCGGACCGACAGAGGCTGGTACTGCGACAACGCCGGCAGCGTGTCA
TTCTTTCCACAGGCCGAGACATGCAAGGTGCAGAGCAACCGGGTGTCTG-
CGACACCATGAA-
CAGCCTGACCCTGCCCTCCGAAGTGAACCTGTGCAACGTGGACATCTTCAAC
CCCAAGTACGACTGCAAGATCATGACCTCCAAGACCGACGTGTCCAG-
CTCCGTGATCAC-
CTCCCTGGGCGCCATCGTGTCTGCTACGGCAAGACCAAGTGCACCGCCAGC
AACAAGAACCGGGGCATCATCAAGACCTTCAGCAACGGCTGCGACTACGTG-
TCCAA-
CAAGGGGGTGGACACCGTGTCCGTGGGCAACACCCTGTACTACGTGAACAAA
CAGGAAGGCAAGAGCCTGTACGTGAAGGGCGAGCCCATCATCAACTTC-
TACGACCCCCTGG-
TGTTCCCCAGCGACGAGTTCGACGCCAGCATCAGCCAGGTGAACGAGAAGAT
CAACCAGAGCCTGGCCTTCATCCGGAAGTCCGACGAGCTGCTGCACAATG-
TGAATGCCGG-
CAAGTCCACCACCAACTGATGAGCGGCCATCTAATCAACCTCTGGATTACAA
AATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGC-
TATGTGGATACG-
CTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTT
CTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGG-
CCCGTTGTCAGG-
CAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGG

GCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTT-
CCCCCTCCCTATTG-
CCACGGCGGAACTCATCGCCGCTGCCTTGCCCGCTGCTGGACAGGGGCTCG
GCTGTTGGGCACTGACAATTCCGTGGTGTTCGGGGAAATCATCG-
TCCTTTCCTTGGCTG-
CTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCC
CTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGG-
CTCTGCGGCCTCTT-
CCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCT
CCCCGCCTGCGGCCGCGATCTGCTGTGCCTTCTAGTTGCCAGCCATCTG-
TTGTTTG-
CCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTT
TCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTC-
TATTCTGGGGGG-
TGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT
GCTGGGGATGCGGTGGGCTCTATGG

Sequência aprimorada do promotor de CMV: negrito

Sequência do transgene F0: itálico

Sequência de WPRE: negrito sublinhado

bghpolyA sinal PolyA: itálico + sublinhado

SEQ ID NO: 10 Sequência de proteínas F0

MELLILKANAITTILTAVTFCFASGQNITEEFYQSTCSAVSKGYLSAL-
 RTGWYTSVITIELS-
 NIKENKCNGTDAKVKLKQELDKYKNAVTELQLLMQSTPATNNRARRELPRF
 MNYTLNNAKKTNVTLSKKRKRRLGFLGVSASIAAGVAVSKVLHLEGEVN-
 KIKSALLSTN-
 KAVVSLSNGVSVLTSKVLDLKNYIDKQLLPVNVKQSCSISNIETVIEFQQKN
 NRLLEITREFSVNAGVTTTPVSTYMLTNSELLSLINDMPITNDQKKLMSNN-
 VQIVR-
 QQSYSIMSIIKEEVLAYVVQLPLYGVIDTPCWKLHTSPLCTTNTKEGSNICL

TRTDRGWYCDNAGSVSFFPQAETCKVQSNRVFCDTMNSLTLPSEVNLCNV-
DIFNPKYDCKIM-

TSKTDVSSSVITSLGAIVSCYGKTKCTASNKNRGI IKTF SNGCDYVSNKGVD
TVSVGNTLYYV NKQEGKSLYVKGEPIINFYDPLVFPSDEFDASIS-
QVNEKINQSLAFIRKSDELLHNVNAGKSTTN

SEQ ID NO: 11 Sequência de aminoácidos de um ligante flexível

Gly-Gly-Gly-Ser-Gly-Gly-Gly

SEQ ID NO: 12 Sequência de aminoácidos de um ligante flexível

Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly

REIVINDICAÇÕES

1. Vetor adenoviral símio, caracterizado pelo de que compreende dois cassetes de expressão, em que cada cassete de expressão compreende um transgene e um promotor, e em que o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral símio e o segundo cassete de expressão é inserido em uma região do vetor adenoviral compatível com a replicação do vetor.

2. Vetor adenoviral símio, caracterizado pelo de que compreende dois cassetes de expressão, em que cada cassete de expressão compreende um transgene e um promotor, e em que o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral símio e o segundo cassete de expressão é inserido na região E3, entre os códons de parada de L5 e E4 (a região “HE1”) ou entre o final da ITR e o sítio de terminação do mRNA de E4 (a região “HE2”) do vetor adenoviral símio.

3. Vetor adenoviral símio, de acordo com a reivindicação 2, caracterizado pelo fato de que o segundo cassete de expressão é inserido na região E3 do vetor adenoviral símio.

4. Vetor adenoviral símio, de acordo com a reivindicação 2, caracterizado pelo fato de que o segundo cassete de expressão é inserido na região HE1 do vetor adenoviral símio.

5. Vetor adenoviral símio, de acordo com a reivindicação 2, caracterizado pelo fato de que o segundo cassete de expressão é inserido na região HE2 do vetor adenoviral símio.

6. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o vetor é um vetor adenoviral de chimpanzé.

7. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o vetor é um adenovírus.

8. Vetor adenoviral símio, de acordo com a reivindicação 7, caracterizado pelo fato de que o vetor é ChAd155.

9. Vetor adenoviral símio, de acordo com a reivindicação 7, caracterizado pelo fato de que o vetor é ChAd83.

10. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o primeiro cassete de expressão compreende um CMV humano ou um promotor de CMV humano aprimorado.

11. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o segundo cassete de expressão compreende um CMV humano ou um promotor de CMV humano aprimorado.

12. Vetor adenoviral símio, de acordo com a reivindicação 10 ou 11, caracterizado pelo fato de que o promotor de hCMV aprimorado tem uma sequência de ácidos nucleicos tendo pelo menos cerca de 90%, pelo menos cerca de 95%, pelo menos cerca de 96%, pelo menos cerca de 97%, pelo menos cerca de 98%, pelo menos cerca de 99% de identidade de sequência com a SEQ ID NO: 6.

13. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações 10 a 12, caracterizado pelo fato de que o promotor compreende ou consiste em uma sequência de ácidos nucleicos da SEQ ID NO: 6.

14. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o primeiro e o segundo cassete de expressão compreendem promotores diferentes.

15. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o vetor adenoviral é capaz de infectar uma célula de mamífero.

16. Vetor adenoviral símio, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que o primeiro

e/ou o segundo cassete de expressão compreendem ainda um elemento regulador pós-transcricional.

17. Vetor adenoviral símio, de acordo com a reivindicação 16, caracterizado pelo fato de que o elemento regulador pós-transcricional é um Elemento Regulador Pós-transcricional de Hepatite da Marmota.

18. Composição, caracterizada pelo fato de que compreende um vetor adenoviral símio como definido em qualquer uma das reivindicações precedentes e um excipiente farmacologicamente aceitável.

19. Vetor ou composição adenoviral símia, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que é para uso como medicamento.

20. Vetor ou composição adenoviral símia, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que é para uso como vacina.

21. Vetor ou composição adenoviral símia, de acordo com qualquer uma das reivindicações precedentes, caracterizado pelo fato de que é para terapia ou profilaxia de uma doença.

22. Método para induzir uma resposta imune em um indivíduo, caracterizado pelo de que compreende a administração do vetor ou composição adenoviral símia como definido em qualquer uma das reivindicações precedentes ao indivíduo.

23. Uso de um vetor, como definido em qualquer uma das reivindicações 1 a 17, ou uma composição, como definida na reivindicação 18, caracterizado pelo fato de que é para o preparo de um medicamento ou uma vacina para terapia ou profilaxia de uma doença.

24. Uso de um vetor, como definido em qualquer uma das reivindicações 1 a 17, ou uma composição, como definida na reivindicação 18, caracterizado pelo fato de que é para o preparo de um medicamento ou uma vacina para induzir uma resposta imune em um indivíduo.

25. Invenção, caracterizada em quaisquer formas de suas concretizações ou em qualquer categoria aplicável de reivindicação, por exemplo, produto ou processo ou uso, ou qualquer outro tipo de reivindicação englobada pela matéria inicialmente descrita, revelada ou ilustrada no pedido de patente.

FIG. 1

(i) RC1



(ii) RC3:



(iii) RC2:



FIG. 2A

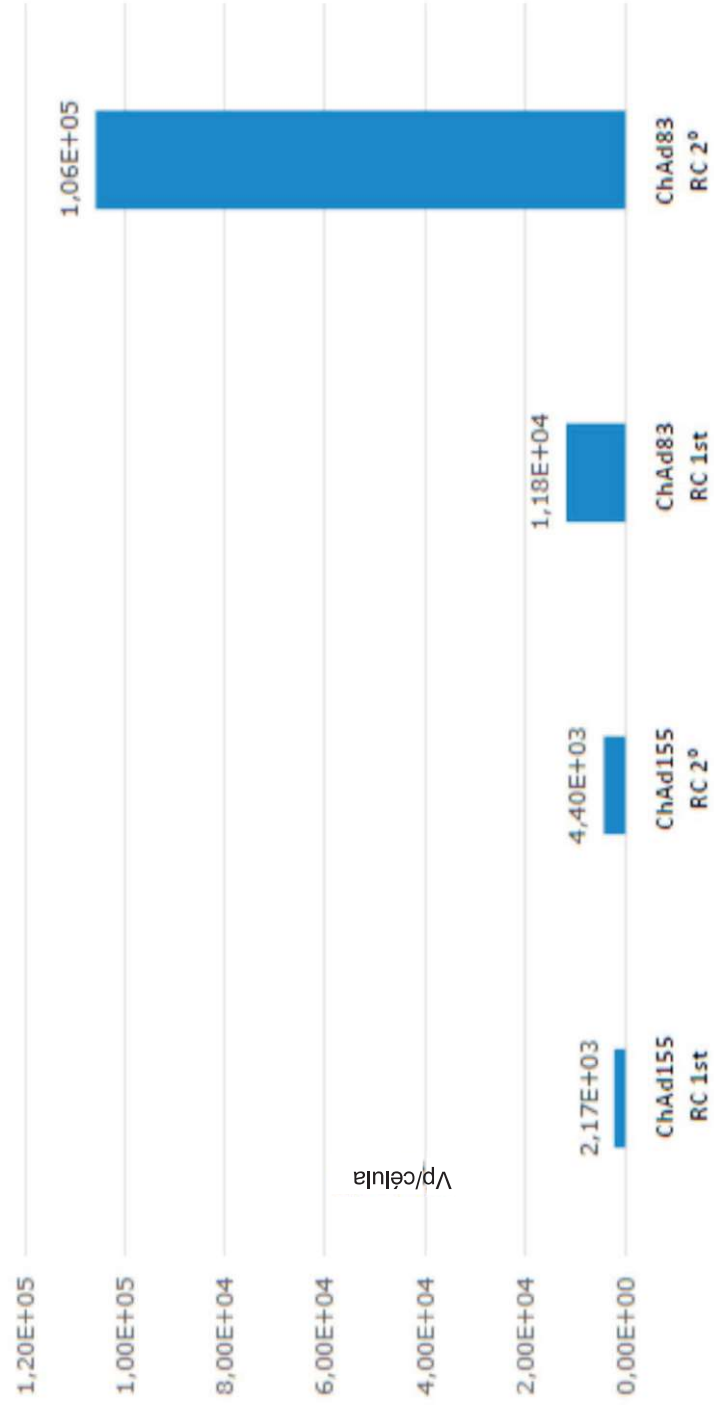


FIG. 2B

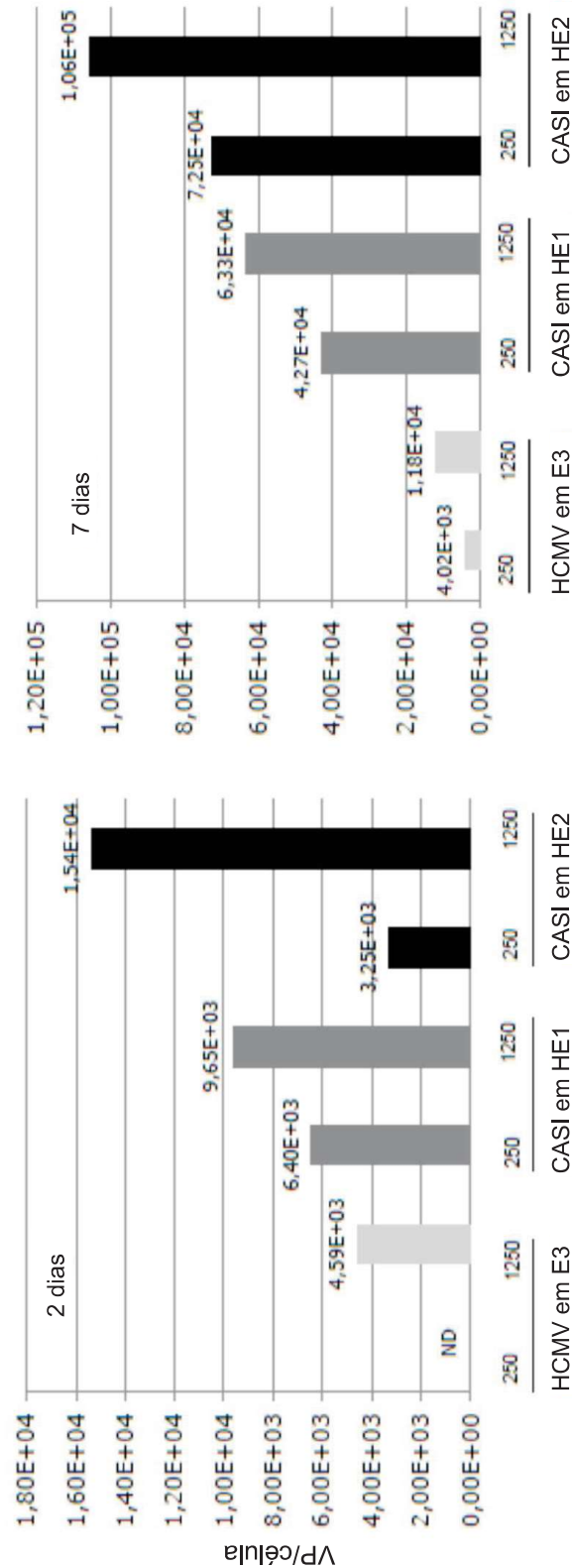


FIG. 3A

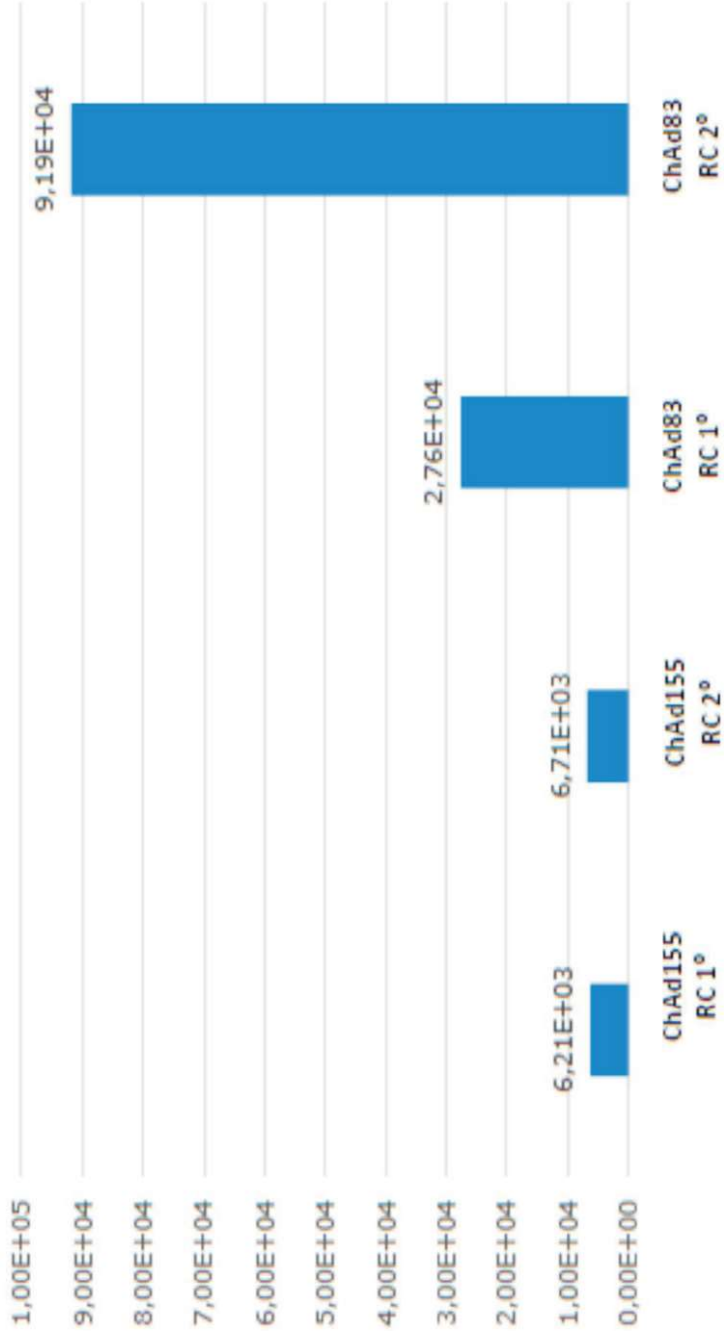


FIG. 3B

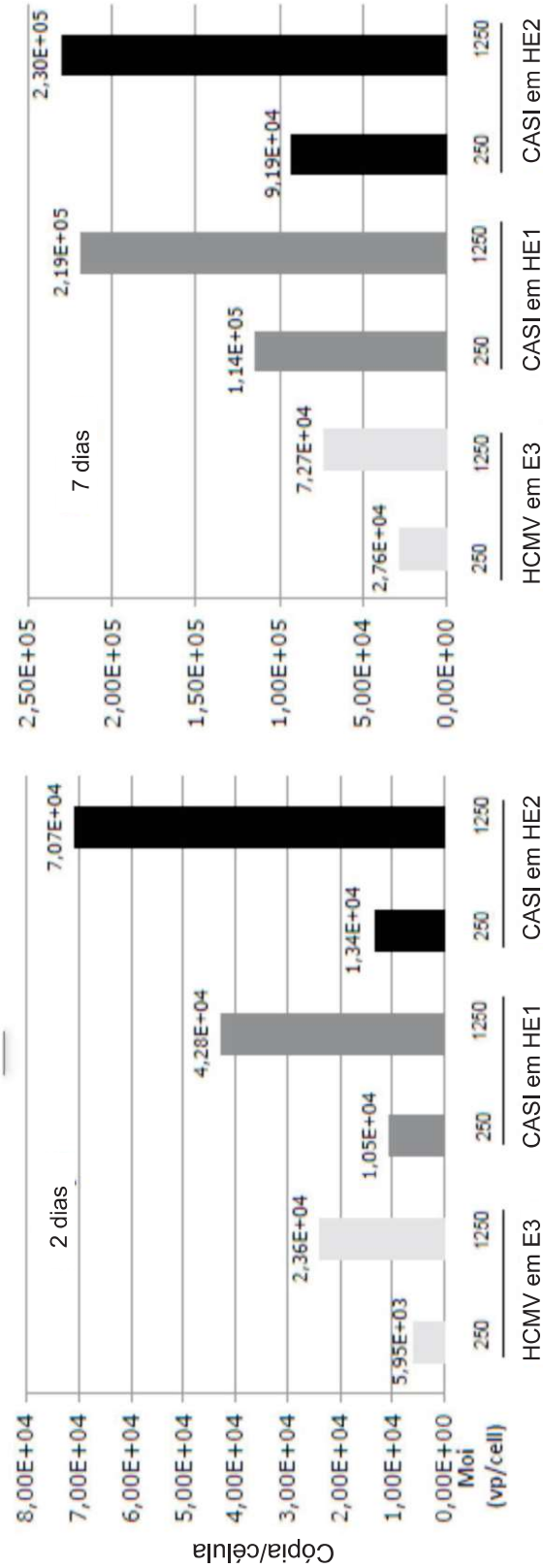
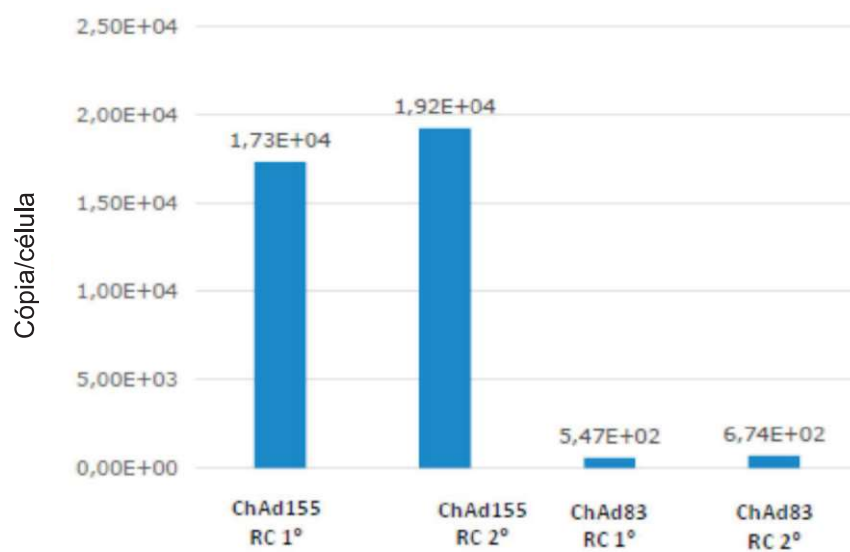
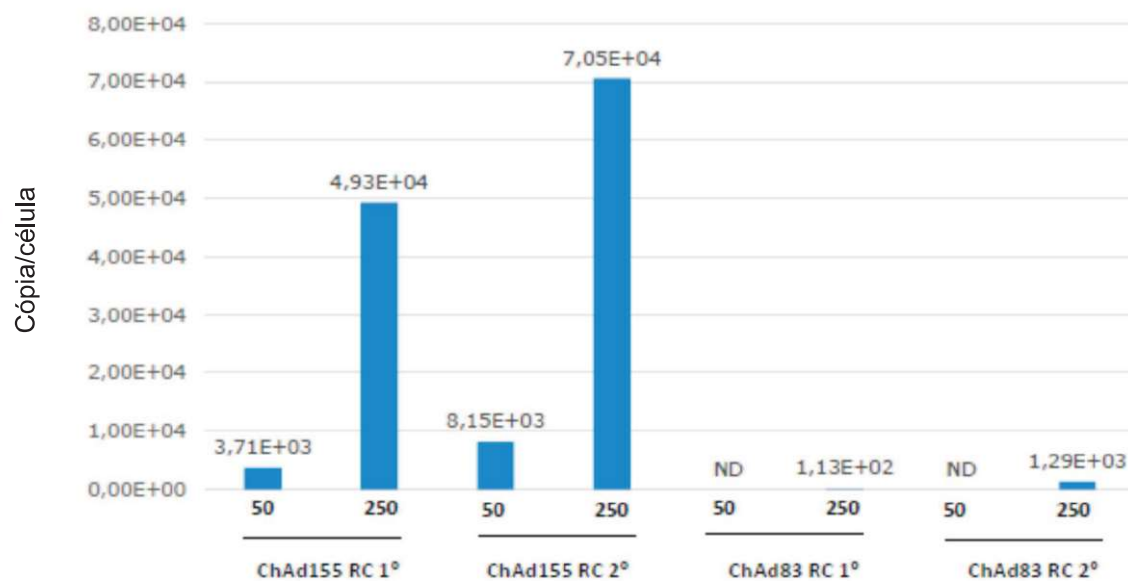


FIG. 4

(a)

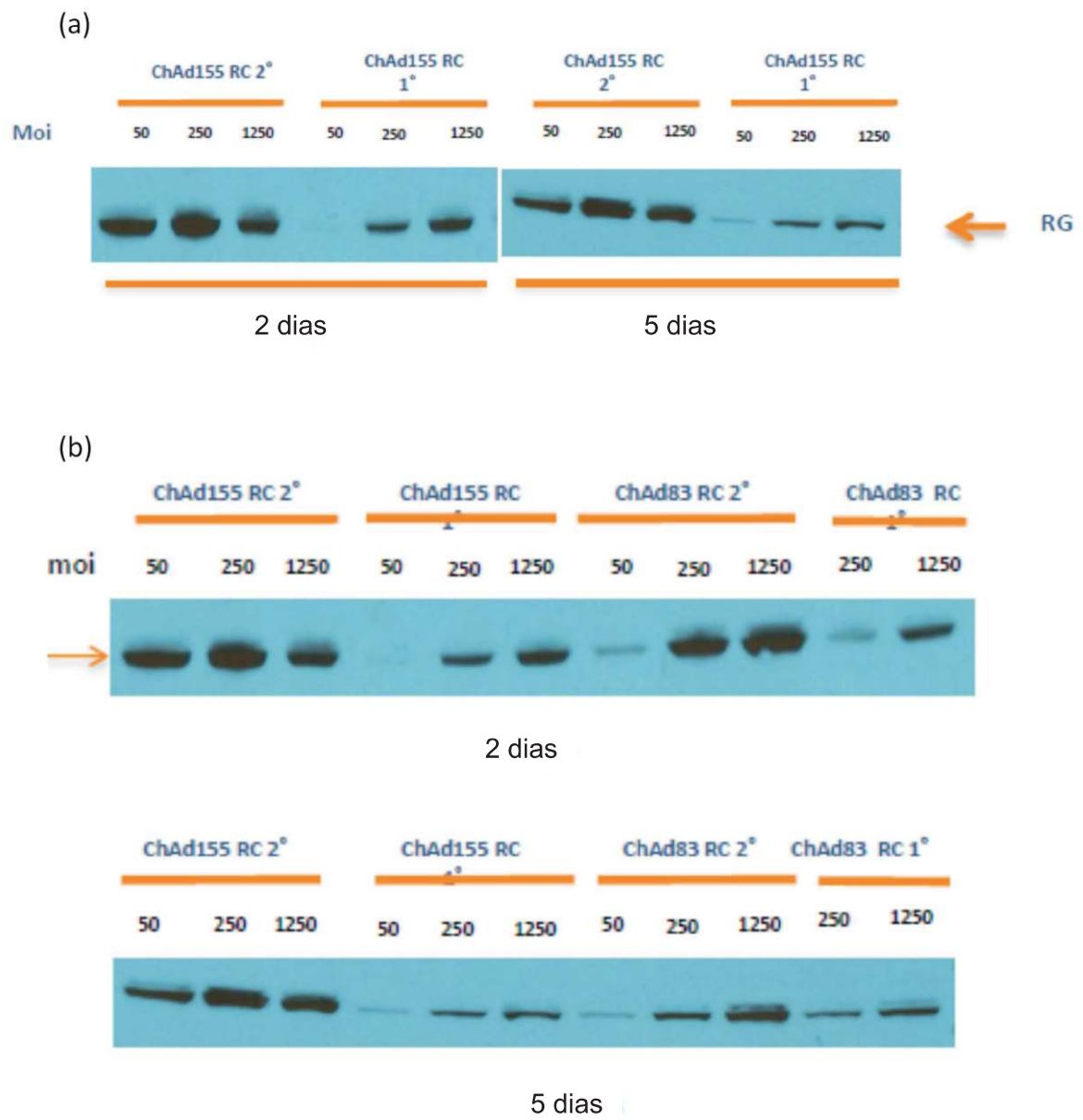


(b)



7/27

FIG. 5



8/27

FIG. 5

(c)

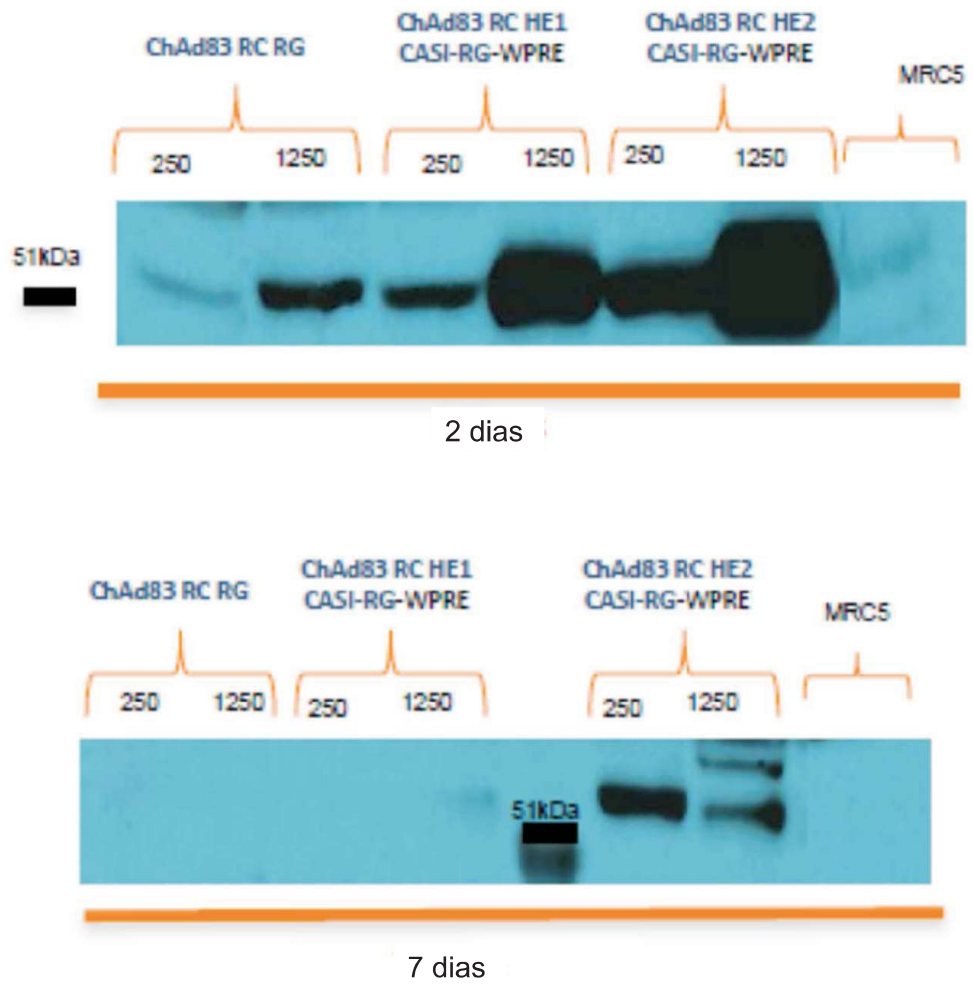


FIG. 6



FIG. 7

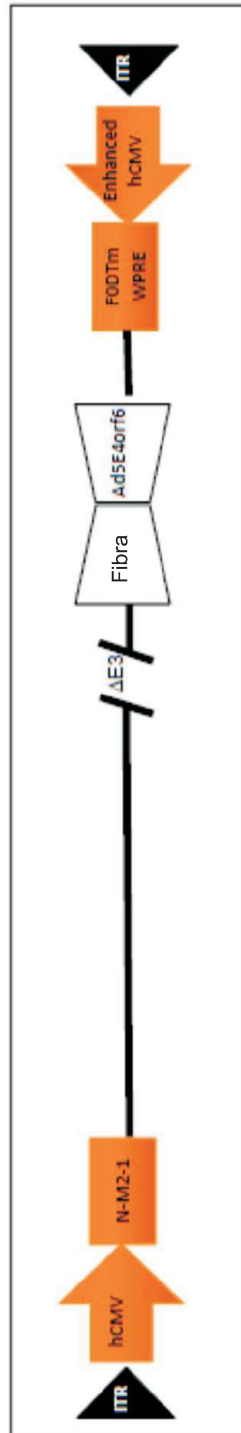


FIG. 8

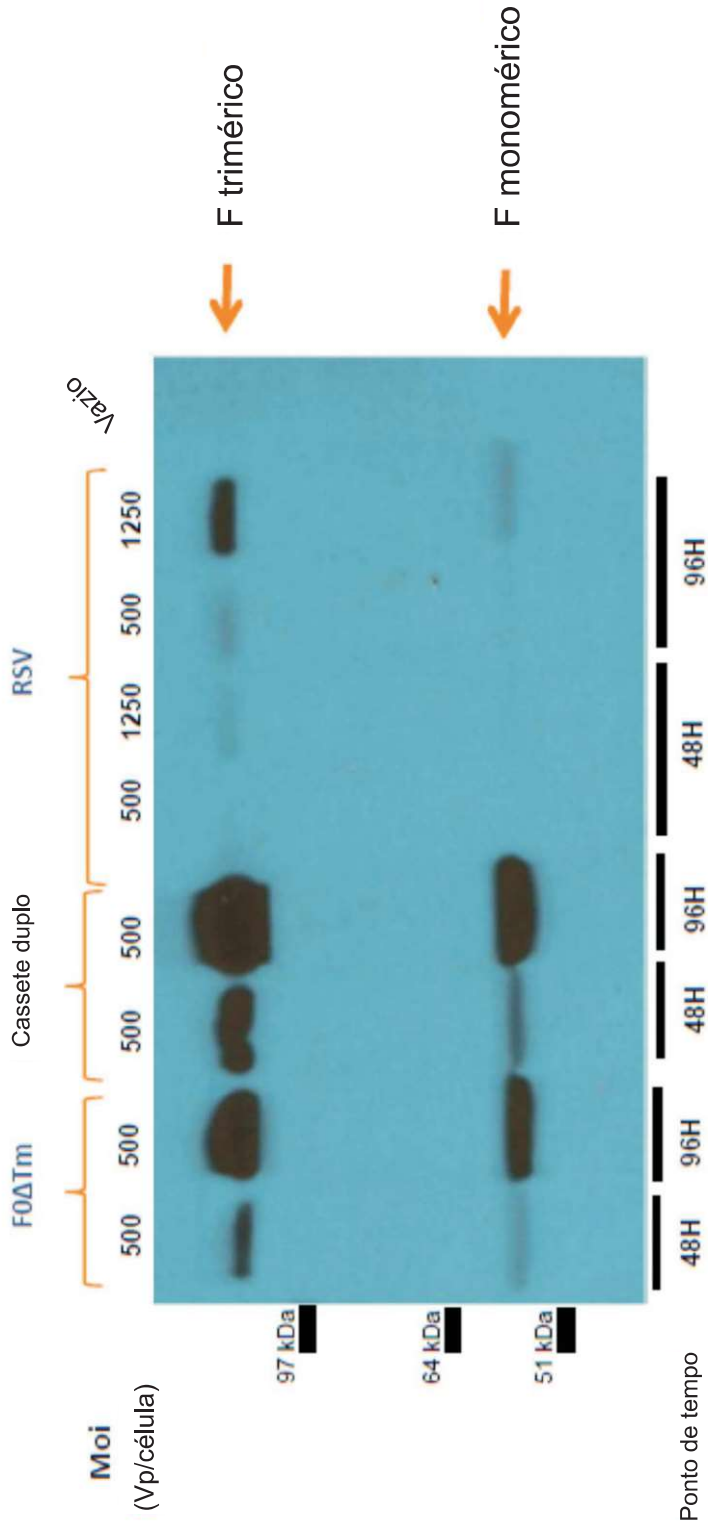


FIG. 9

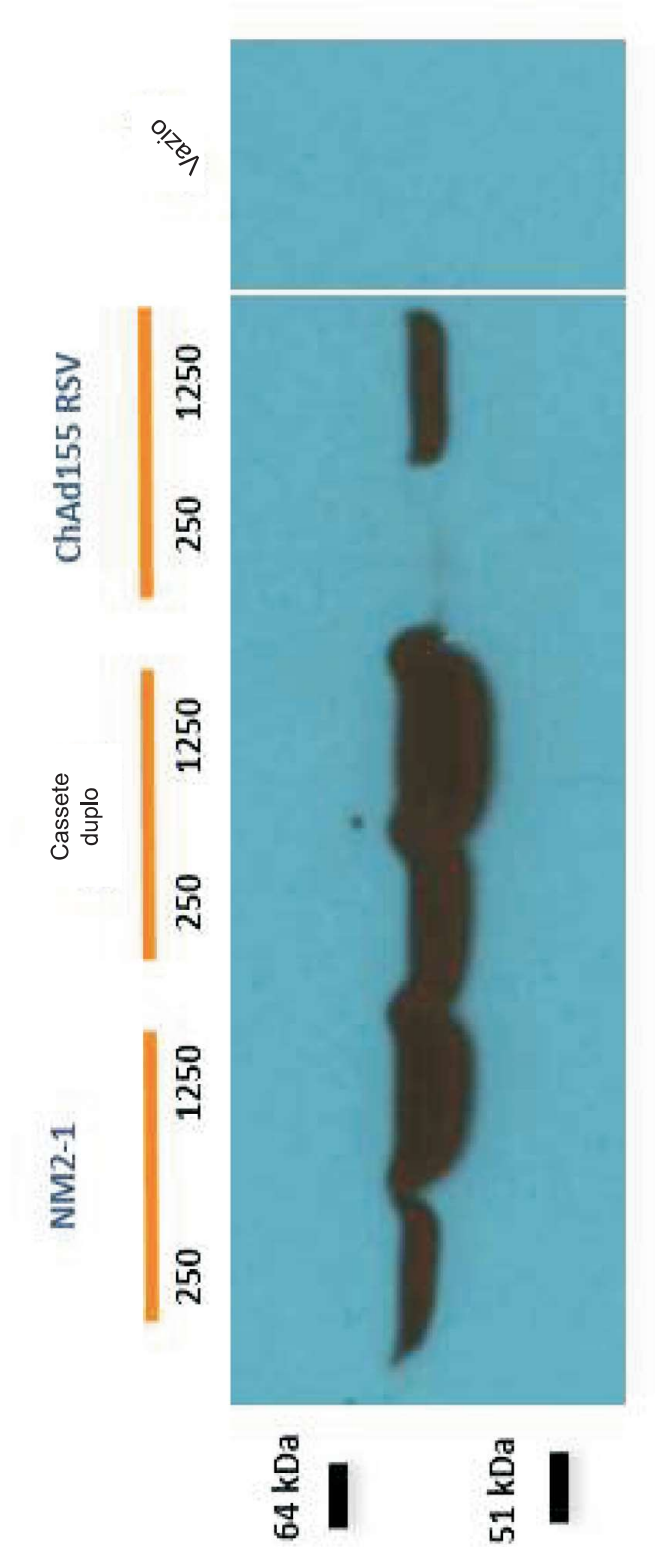


FIG. 10

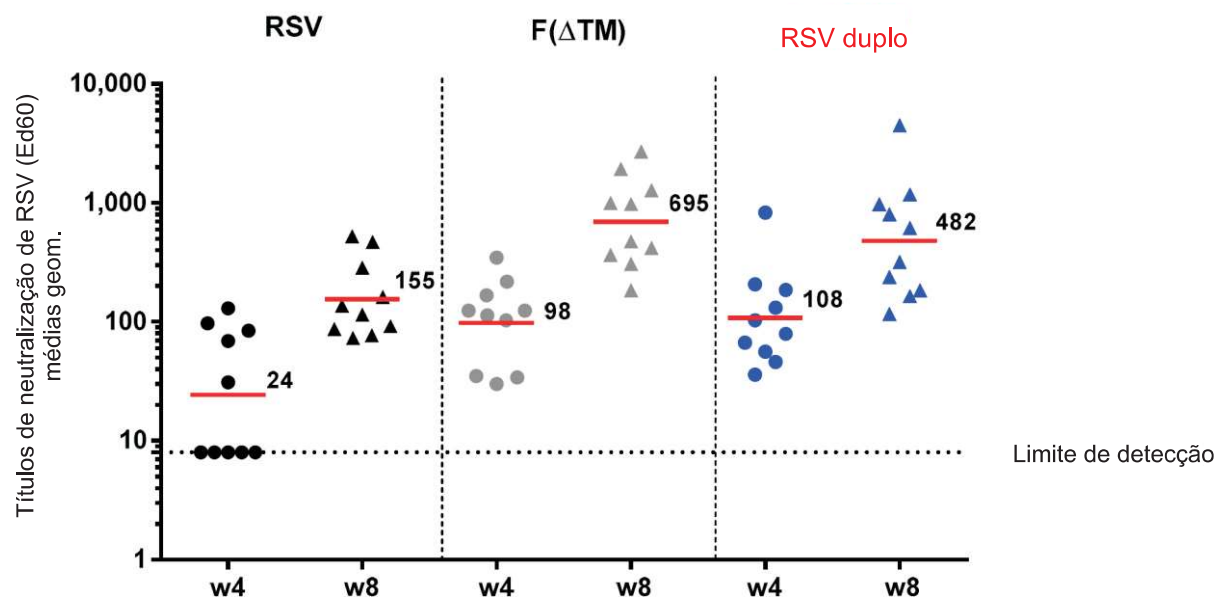
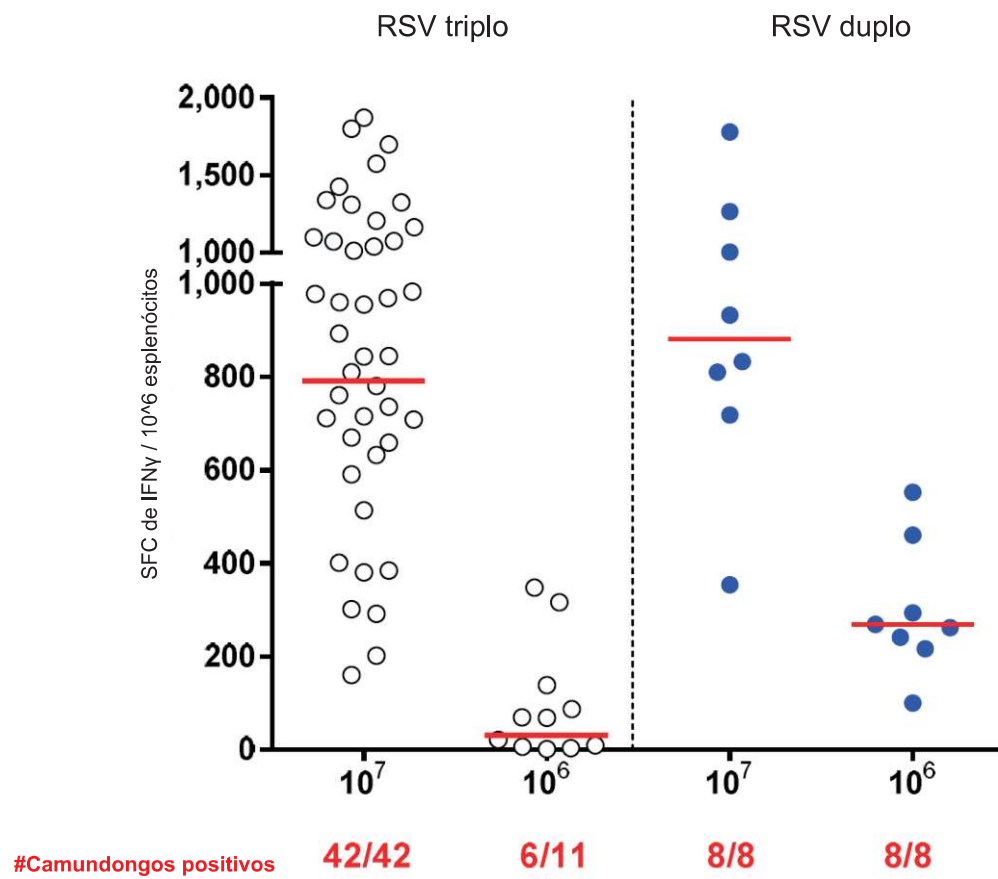


FIG. 11



% de IFNy específico para RSV + CD4 em BAL

FIG. 12A

BAL CD4

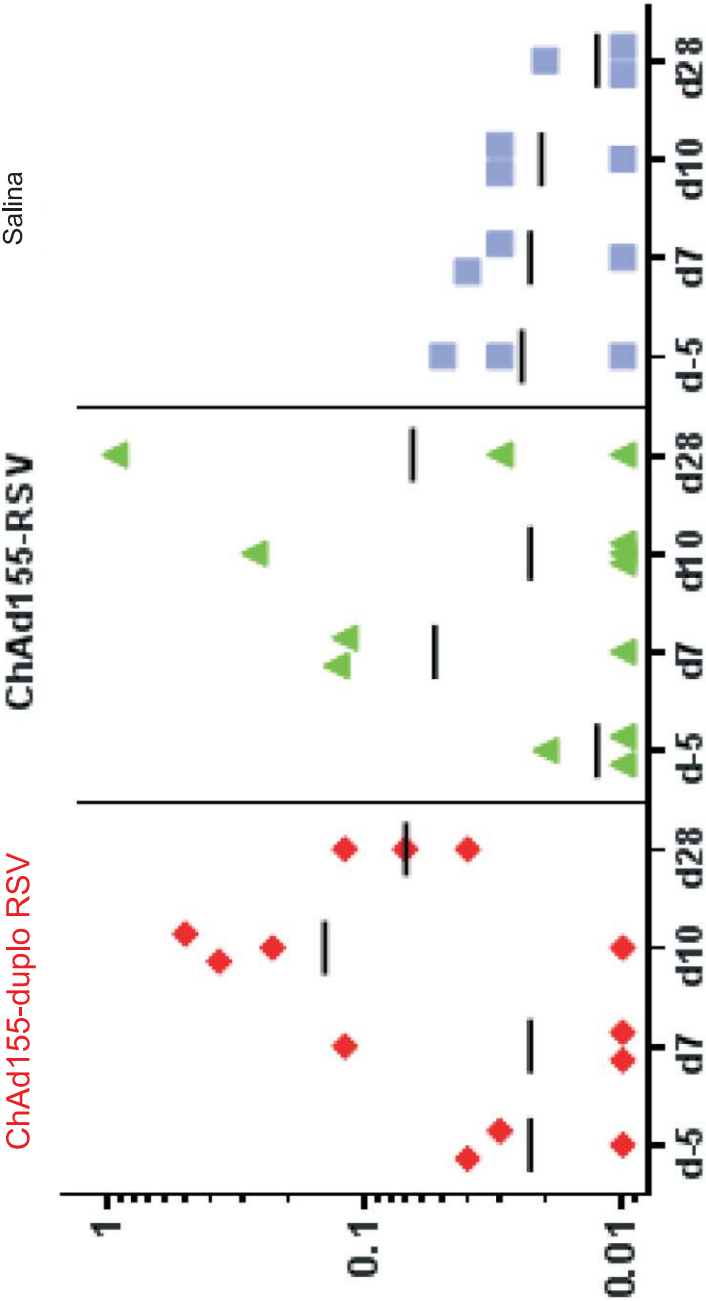


FIG. 12B

BAL CD8

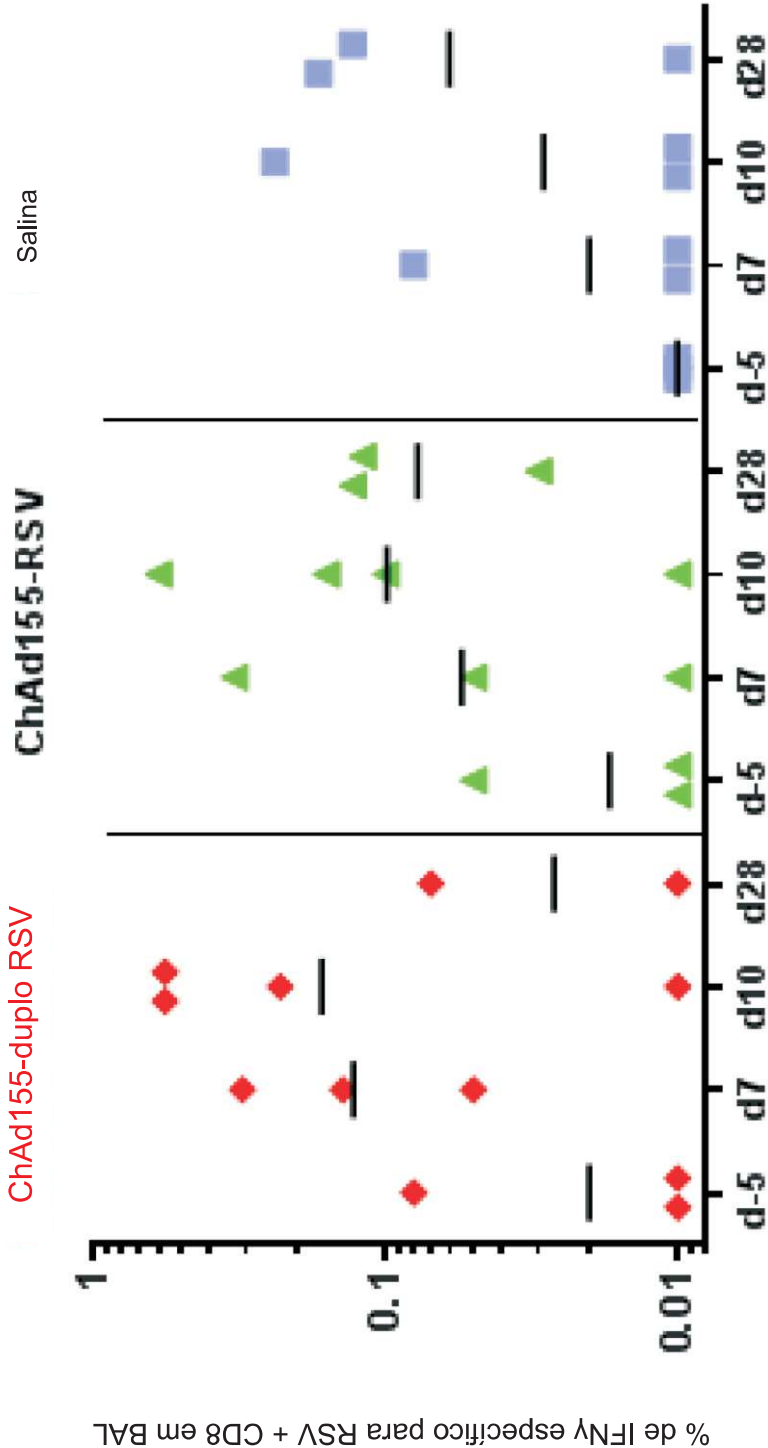


FIG. 13A

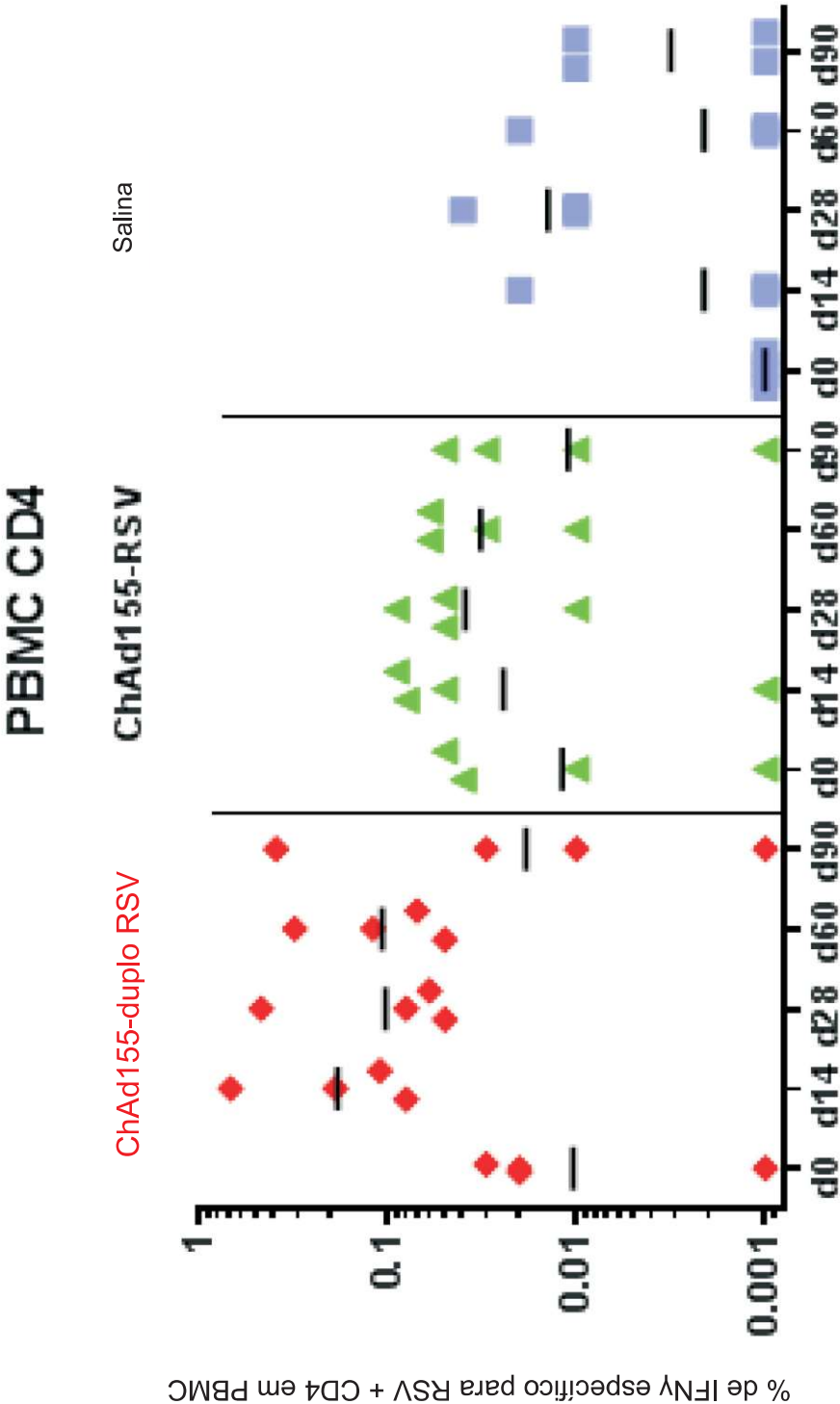


FIG. 13B

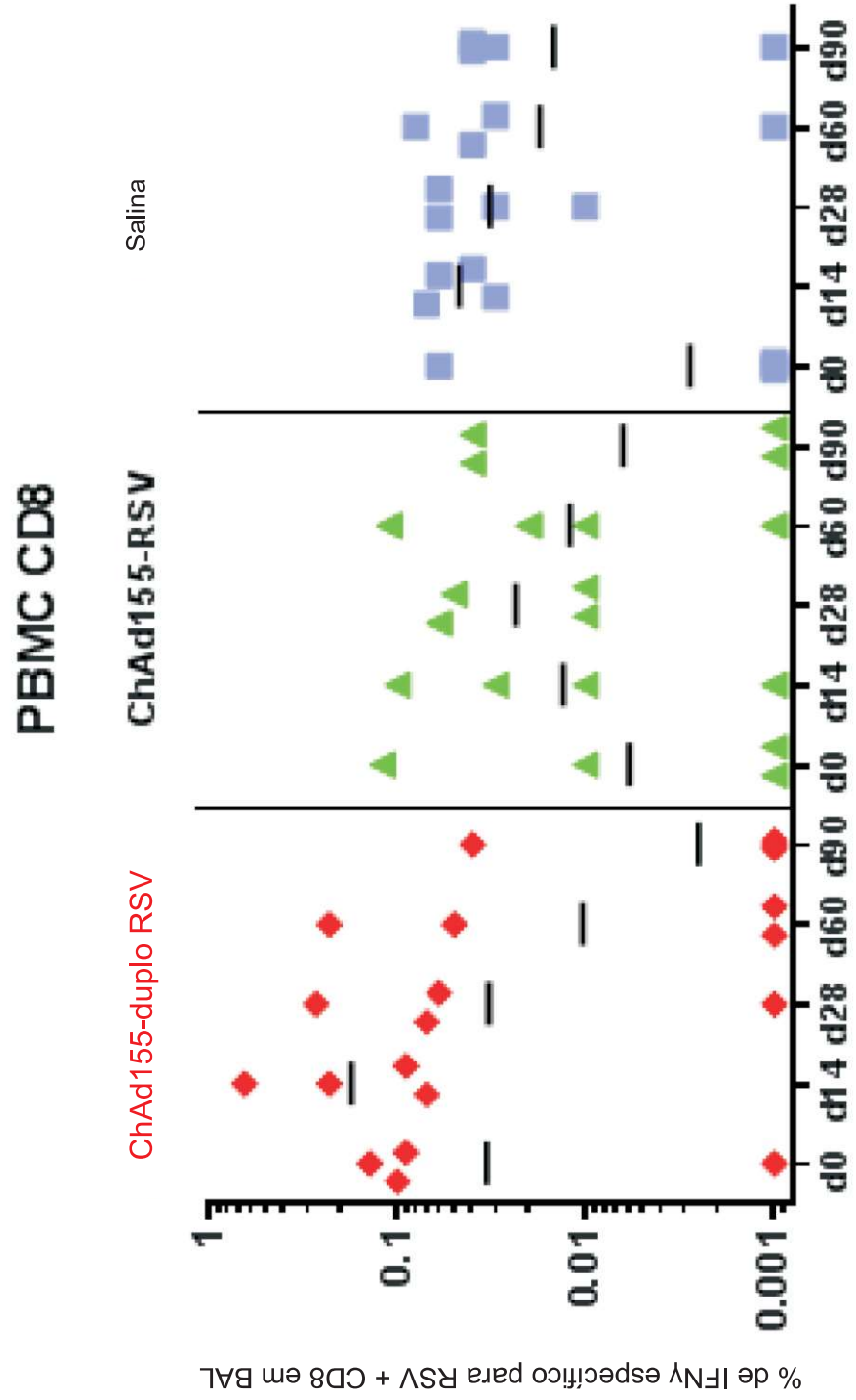


FIG. 14A

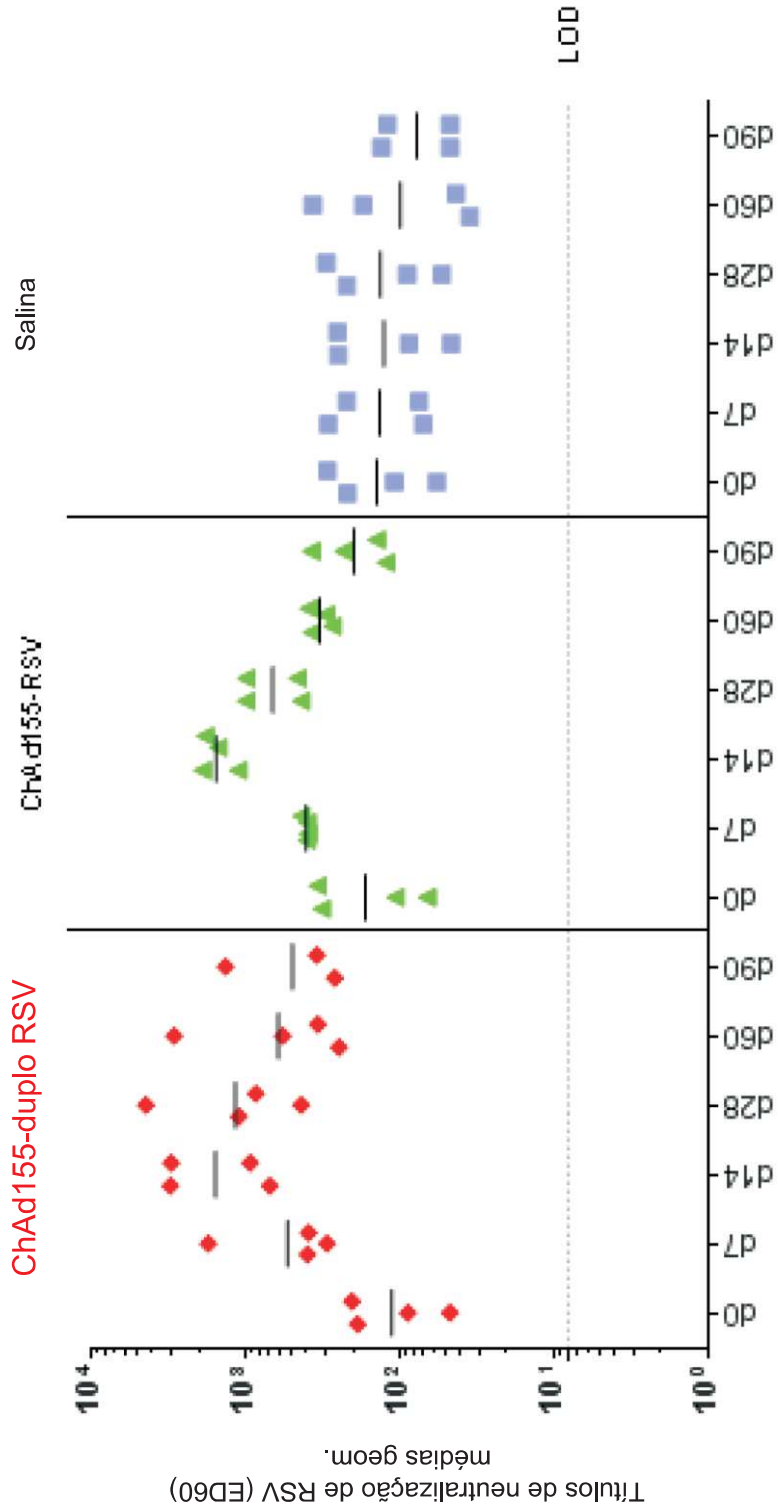


FIG. 14B

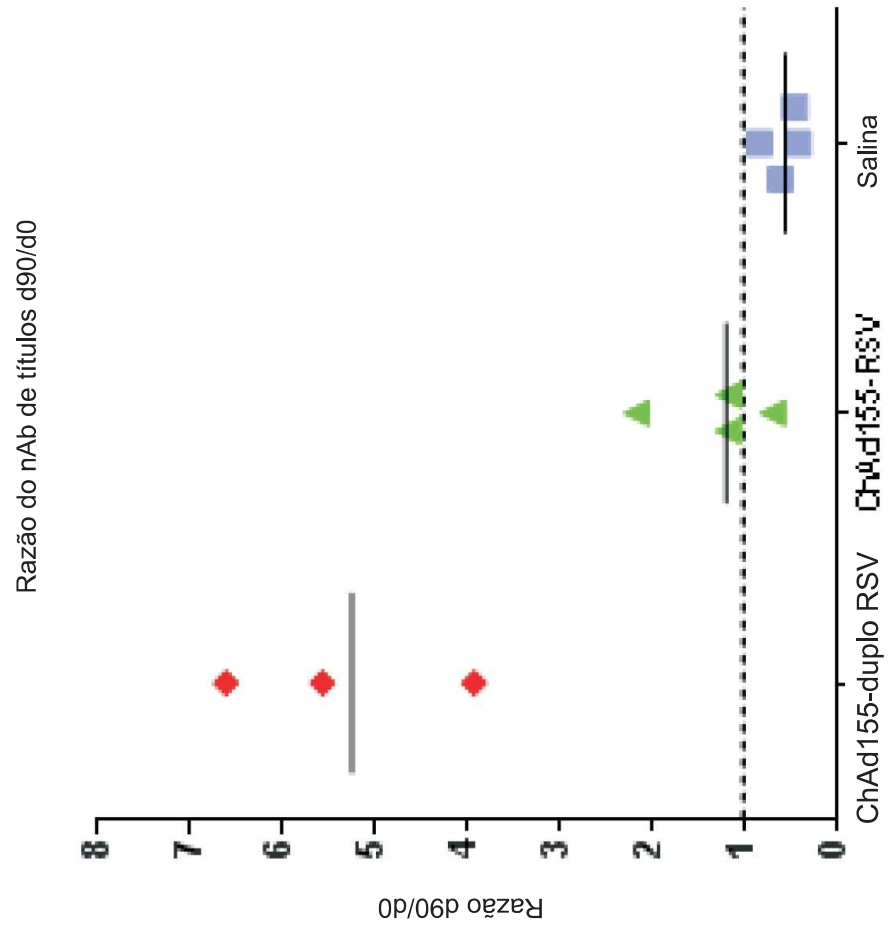


FIG. 15A

Resposta aos antígenos RSV M e N

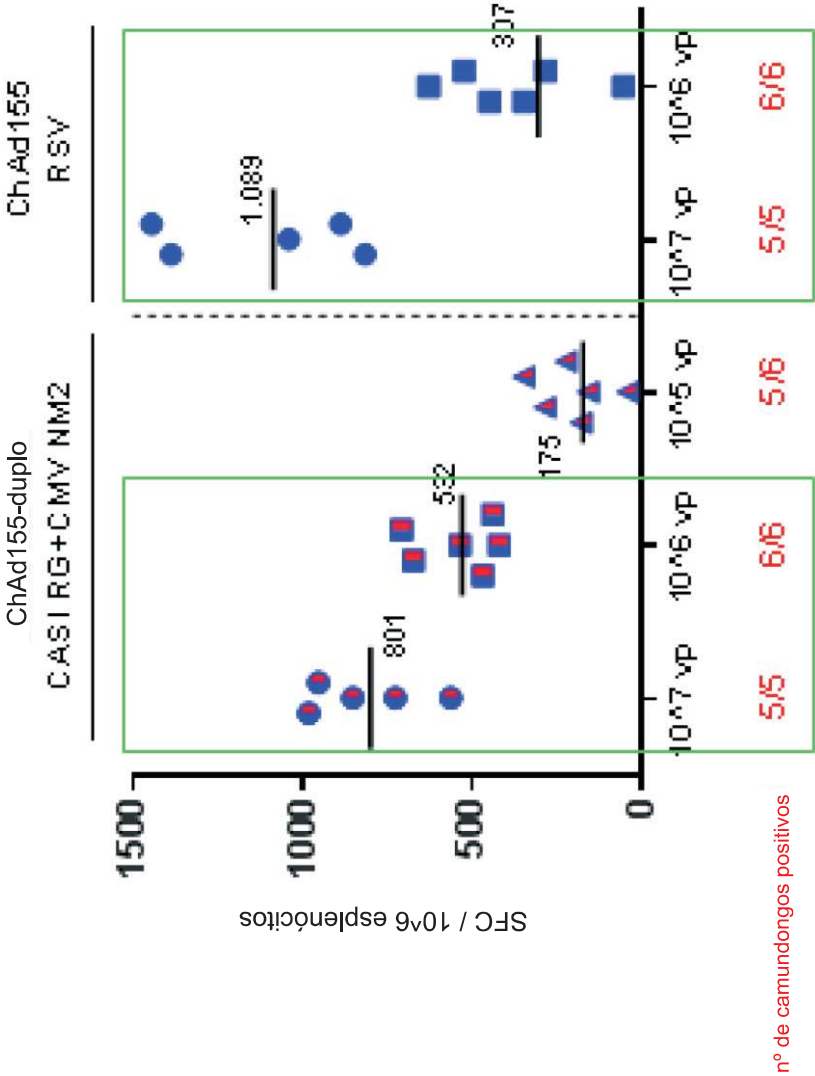
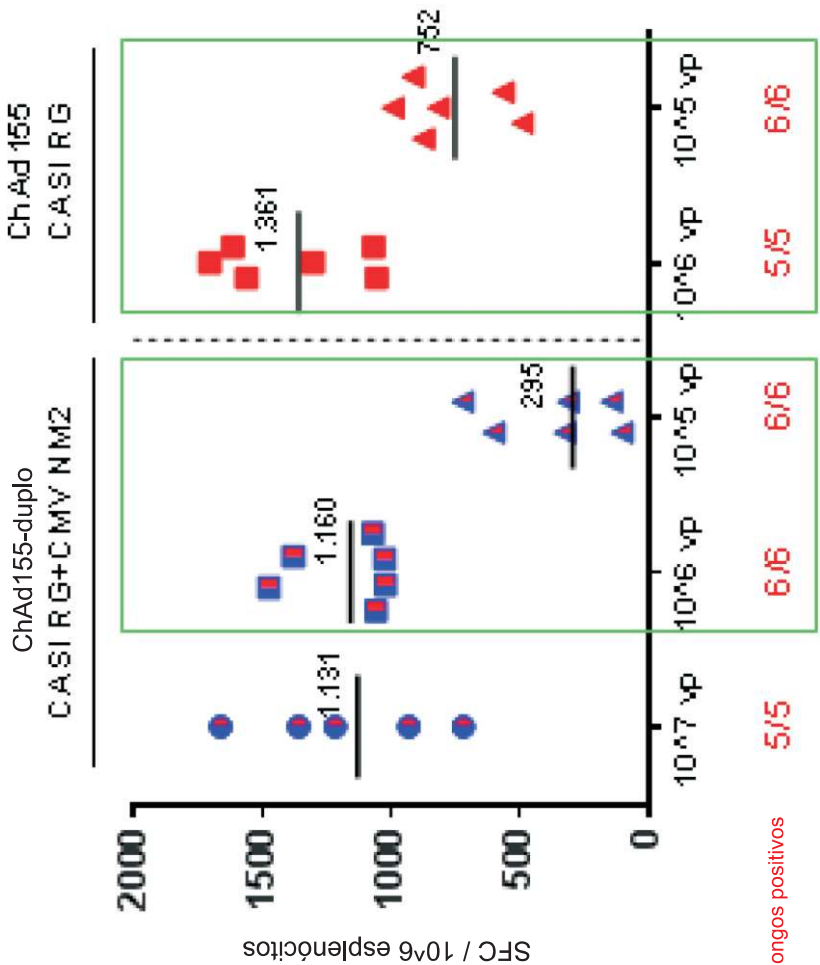


FIG. 15B

Resposta aos antígenos da raiva RG



n° de camundongos positivos

FIG. 15C

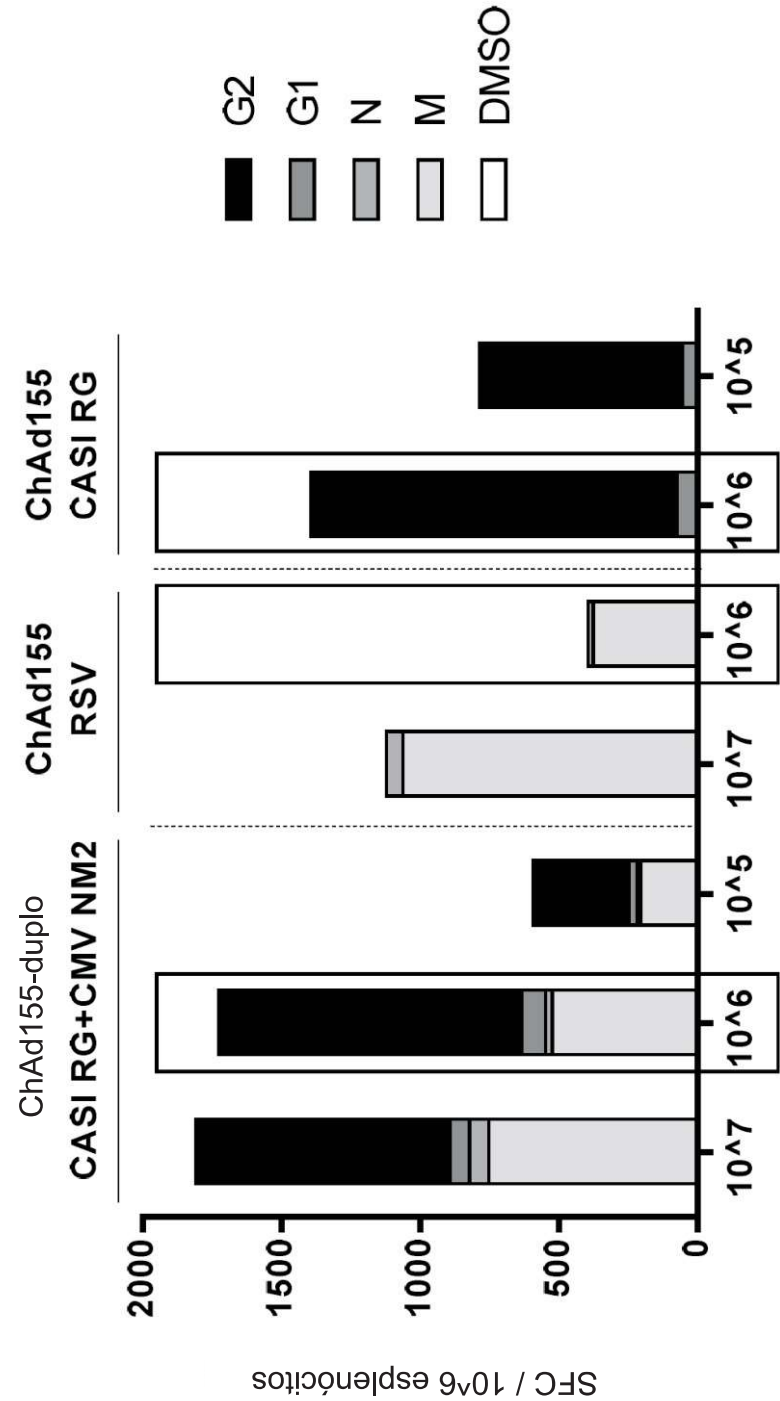


FIG. 16A

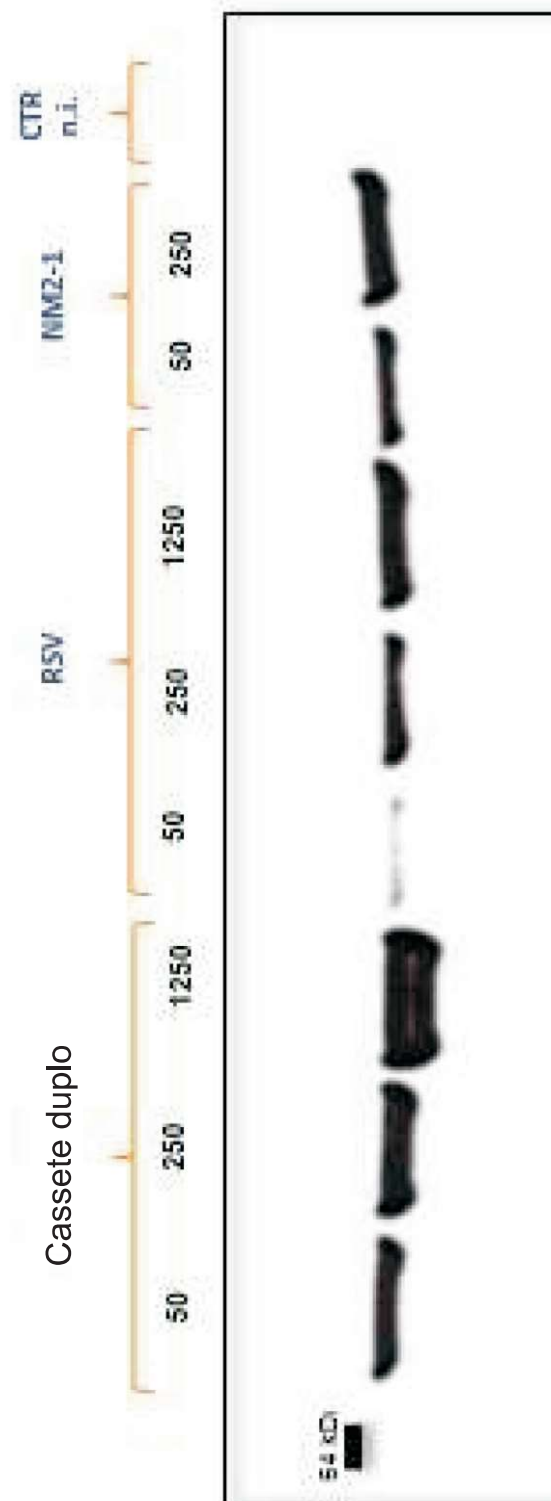


FIG. 16B

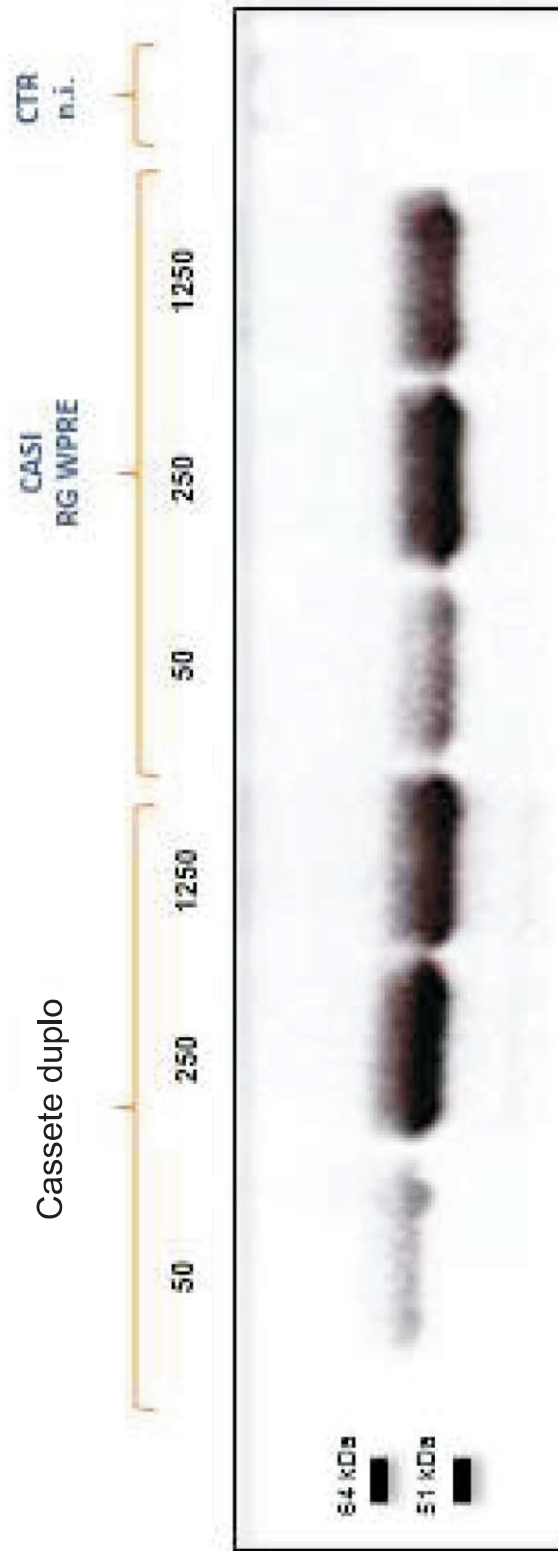


FIG. 17

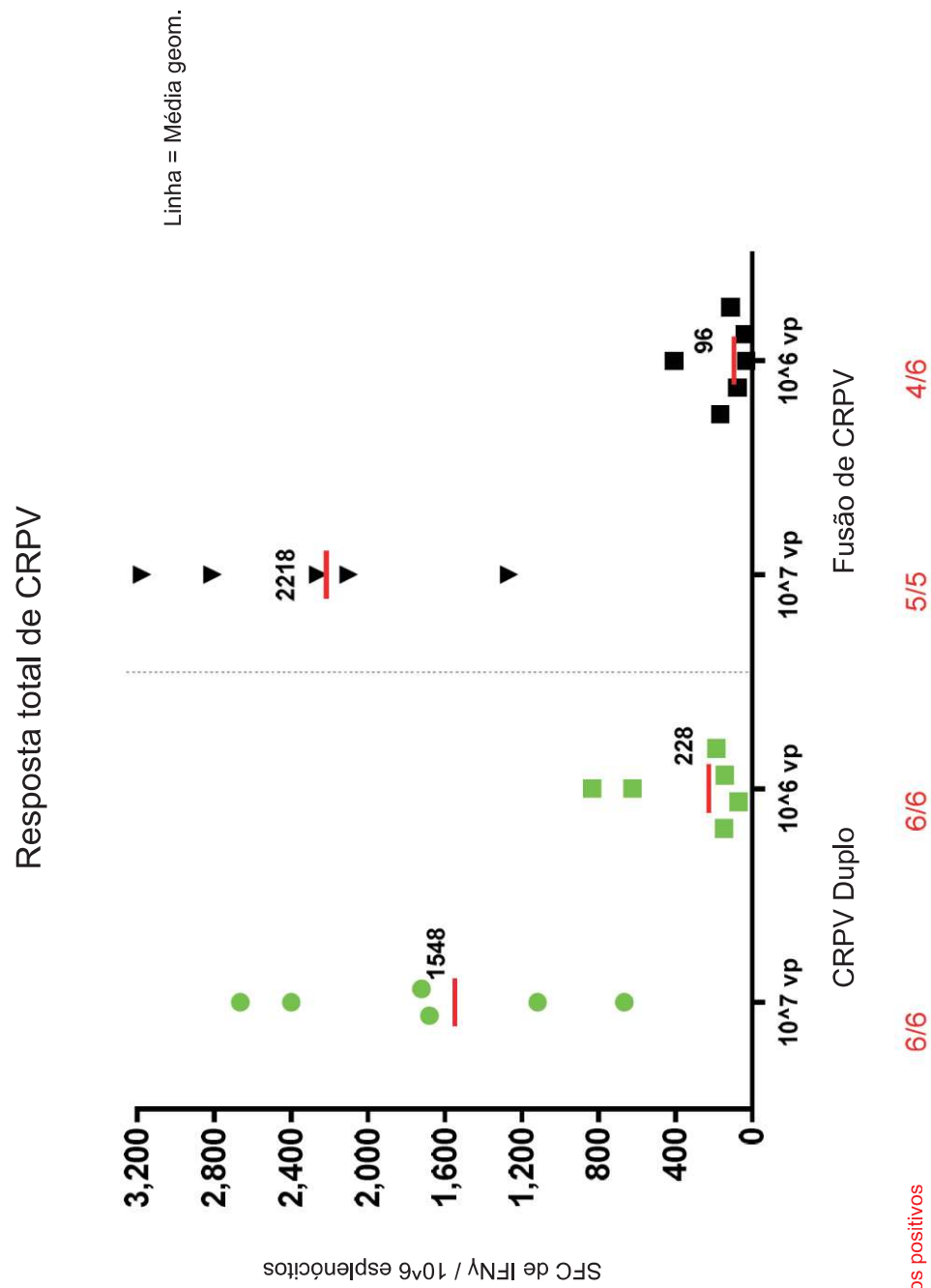
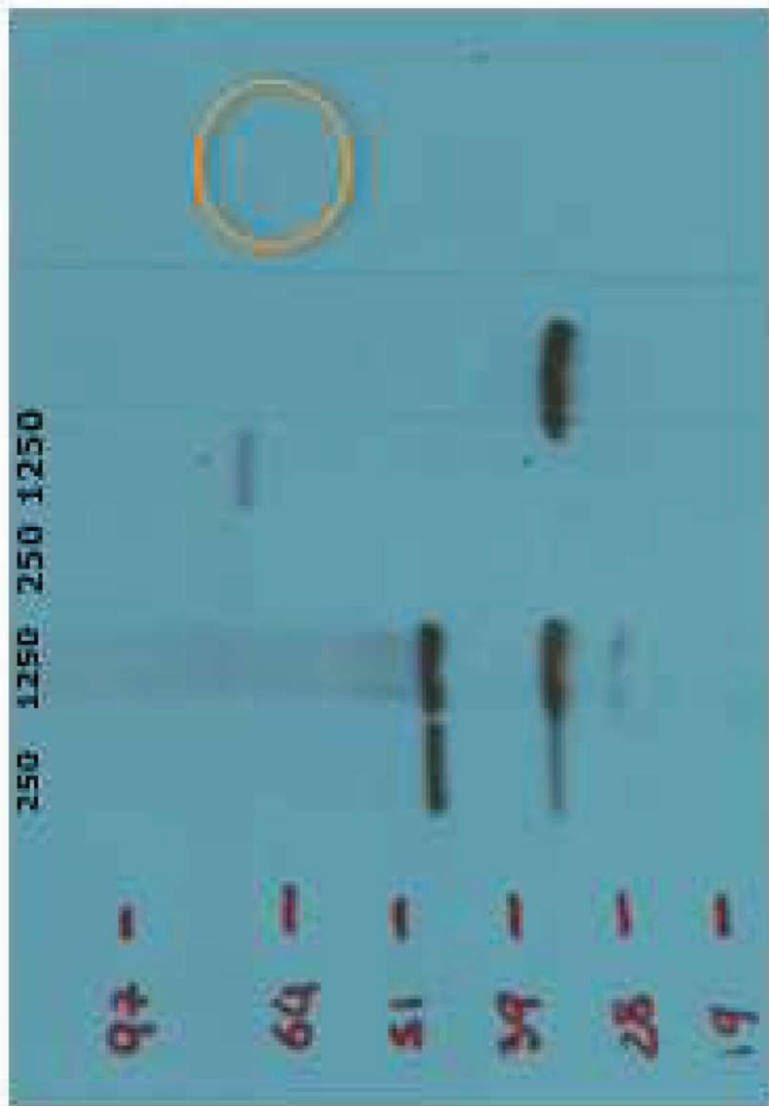


FIG.18

155 Duplo 155 Fusão Pvj DE2 Pvj Fusão



RESUMO

Patente de Invenção: “**VETORES ADENOVIRAIS COM DOIS CASSETES DE EXPRESSÃO**”.

A presente invenção refere-se a um vetor adenoviral símio que compreende dois cassetes de expressão, em que cada cassete de expressão compreende um transgene e um promotor, e em que o primeiro cassete de expressão é inserido na região E1 do vetor adenoviral símio e o segundo cassete de expressão é inserido em uma região do vetor adenoviral compatível com a replicação do vetor.

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

Código de Controle

Campo 1



Campo 2



Outras Informações:

- Nome do Arquivo: listagem de sequência p246456.TXT
- Data de Geração do Código: 14/04/2020
- Hora de Geração do Código: 16:38:33
- Código de Controle:
 - Campo 1: C4771B56B6AB6859
 - Campo 2: AEE3959BC42150AF