

C07K 14/11 (2018.01) **A01H 5/00** (2018.01)
A61K 39/145 (2018.01) **A61P 31/16** (2018.01)
C12N 15/44 (2018.01) **C12N 15/82** (2018.01)
C12N 7/00 (2018.01) **C12N 7/01** (2018.01)
C07K 19/00 (2018.01)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2010.06.25**

(30) Prioridade(s): **2009.06.24 US 220161 P**

(43) Data de publicação do pedido: **2012.05.02**

(45) Data e BPI da concessão: **2018.03.28**
109/2018

(73) Titular(es):

MEDICAGO INC.

1020 ROUTE DE L' EGLISE SUITE 600 QUÉBEC,
QUÉBEC G1V 3V9 **CA**

(72) Inventor(es):

PIERRE-OLIVIER LAVOIE **CA**

MANON COUTURE **CA**

MICHÈLE DARGIS **CA**

MARC-ANDRÉ D'Aoust **CA**

LOUIS-PHILIPPE VEZINA **CA**

(74) Mandatário:

LUÍSA MARIA FERREIRA GUERREIRO
RUA RAUL PROENÇA, 3 2820-478 CHARNECA DA
CAPARICA **PT**

(54) Epígrafe: **PARTICULAS QUIMÉRICAS SEMELHANTES A VIRUS DA GRIPE**
COMPREENDENDO HEMAGLUTININA

(57) Resumo:

É PROPORCIONADO UM MÉTODO PARA SINTETIZAR PARTÍCULAS QUIMÉRICAS SEMELHANTES A VÍRUS DA GRIPE (VLPS) DENTRO DE UMA PLANTA OU UMA PORÇÃO DE UMA PLANTA. O MÉTODO ENVOLVE A EXPRESSÃO DE GRIPE HA QUIMÉRICO EM UMA PLANTA OU UMA PORÇÃO DE UMA PLANTA. A INVENÇÃO TAMBÉM É DIRIGIDA A UMA VLP COMPREENDENDO PROTEÍNA HA DE GRIPE QUIMÉRICA E LÍPIDOS DE PLANTAS. A INVENÇÃO É TAMBÉM DIRIGIDA A UM ÁCIDO NUCLEICO QUE CODIFICA HA DA GRIPE QUIMÉRICO, BEM COMO A VETORES. AS VLPS PODEM SER USADAS PARA FORMULAR VACINAS CONTRA A GRIPE, OU PODEM SER USADAS PARA ENRIQUECER AS VACINAS EXISTENTES.

RESUMO

PARTÍCULAS QUIMÉRICAS SEMELHANTES A VÍRUS DA GRIPE COMPREENDENDO HEMAGLUTININA

É proporcionado um método para sintetizar partículas quiméricas semelhantes a vírus da gripe (VLPs) dentro de uma planta ou uma porção de uma planta. O método envolve a expressão de gripe HA quimérico em uma planta ou uma porção de uma planta. A invenção também é dirigida a uma VLP compreendendo proteína HA de gripe quimérica e lípidos de plantas. A invenção é também dirigida a um ácido nucleico que codifica HA da gripe quimérico, bem como a vetores. As VLPs podem ser usadas para formular vacinas contra a gripe, ou podem ser usadas para enriquecer as vacinas existentes.

DESCRIÇÃO
PARTÍCULAS QUIMÉRICAS SEMELHANTES A VÍRUS DA GRIPE COMPREENDENDO
HEMAGLUTININA

CAMPO DE INVENÇÃO

A presente invenção refere-se a partículas semelhantes a vírus. Mais especificamente, a presente invenção refere-se a partículas semelhantes a vírus compreendendo hemaglutinina da gripe quimérica e a métodos de produção de partículas quiméricas semelhantes a vírus da gripe.

ANTECEDENTES DA INVENÇÃO

A gripe é a principal causa de morte em humanos devido a um vírus respiratório e, durante a "temporada de gripe", estima-se que 10 a 20% da população mundial possa estar infectada, levando a 250 a 500 mil mortes por ano.

O método atual de combate à gripe em humanos é por vacinação anual. A vacina é geralmente uma combinação de várias estirpes que se prevê serem as estirpes dominantes para a próxima temporada de gripe, no entanto, o número de doses de vacina produzidas anualmente não é suficiente para vacinar a população mundial. Por exemplo, Canadá e Estados Unidos obtêm doses de vacina suficientes para imunizar cerca de um terço de sua população, e na Europa, apenas cerca de 17% podem ser vacinados, dada a produção atual - em face de uma pandemia mundial de gripe, esta produção seria insuficiente. Mesmo que a produção anual necessária pudesse de alguma forma ser atendida em um determinado ano, as variedades dominantes mudam de ano para ano, portanto, o armazenamento a baixos tempos de necessidade no ano não é prático. A produção económica e em grande escala de uma vacina eficaz contra a gripe é de interesse significativo para o governo e para a indústria privada.

A glicoproteína de superfície da hemaglutinina (HA) da gripe é tanto uma proteína de ligação ao recetor como uma de fusão de membrana. É um trímero de subunidades idênticas, cada uma contendo dois polipéptidos ligados por dissulfeto, HA1 e HA2, que são derivados por clivagem proteolítica de um precursor, HA0, que tem uma sequência peptídica sinal no seu terminal N e uma sequência âncora da membrana no seu terminal C. A clivagem para formar HA1 e HA2 gera o terminal N do polipéptido menor, HA2, que tem a sequência âncora da membrana no seu terminal C. A clivagem é necessária para a atividade de fusão de membrana, mas não para imunogenicidade. A sequência N-terminal de HA2 é chamada de "peptídeo de fusão" porque a clivagem em sequências hidrofóbicas similares é também necessária para a atividade de outras proteínas de fusão de vírus, *in vitro*.

Geralmente, a superfície da "cabeça" globular compreende várias alças flexíveis com regiões antigénicas bem caracterizadas e variáveis designadas como locais A, B, C, D e E (revisto em Wiley et al., 1987. Annu. Rev Biochem 56: 365-394). Inserção ou substituição de sequências peptídicas curtas em alguns locais (por exemplo, B e E) para estudos de imunidade foram descritas (Garcia-Sastré et al. 1995. Biologicals 23: 171-178). Fator de crescimento epidérmico (EGF), anticorpo de cadeia única (scFV) e o domínio Fc de IgG, variando em tamanho de 53 a 246 aminoácidos, foram inseridos na extremidade N-terminal de um H7 e as quimeras foram expressas com sucesso (Hatzioannou et al., 1999. Human Gene Therapy 10: 1533-1544). Mais recentemente, 90 e 140 domínios de aminoácidos do antígeno protetor de *Bacillus anthracis* foram fundidos ao terminal amino de um H3 (Li et al., 2005. J. Virol 79: 10003-1002). Copeland (Copeland et al., 2005. J. Virol 79: 6459-6471) descreve a expressão da glicoproteína de superfície do HIV gp120 Env numa haste H3, em que o domínio gp120 substituiu toda a cabeça globular de HA.

Vários produtos recombinantes foram desenvolvidos como candidatos vacinais contra a gripe recombinantes. Estas abordagens centraram-se na expressão, produção e purificação de proteínas HA e NA de gripe

tipo A, incluindo a expressão destas proteínas utilizando células de inseto infetadas com baculovírus (Crawford et al, 1999 Vaccine 17: 2265-74; Johansson, 1999 Vaccine 17: 2073-80), vetores virais e construções de vacina de ADN (Olsen et al., Vaccine 15: 1149-56, 1997).

A produção de uma estirpe de vírus da gripe não infecciosa para fins de vacina é uma forma de evitar a infecção inadvertida. Em alternativa, partículas semelhantes a vírus (VLPs) como substitutos do vírus cultivado foram investigadas. As VLPs imitam a estrutura do capsídeo viral, mas não possuem um genoma e, portanto, não podem replicar ou fornecer um meio para uma infecção secundária. As tecnologias atuais de produção de VLP influenciam a coexpressão de múltiplas proteínas virais, e essa dependência representa uma desvantagem dessas tecnologias, já que no caso de uma epidemia pandêmica e anual, o tempo de resposta é crucial para a vacinação. Um sistema de produção de VLP mais simples, por exemplo, aquele que se baseia na expressão de apenas uma ou algumas proteínas virais sem requerer a expressão de proteínas virais não estruturais é desejável para acelerar o desenvolvimento de vacinas.

Vírus envelopados podem obter seu envelope lipídico quando 'brotando' para fora da célula infetada e obter a membrana da membrana plasmática, ou a de um organelo interno. Em sistemas celulares de mamíferos ou baculovírus, por exemplo, brotos de gripe da membrana plasmática (Quan et al 2007 J. Virol 81: 3514-3524). Apenas alguns vírus envelopados são conhecidos por infetar plantas (por exemplo, membros dos Tospovirus e Rhabdoviruses). Dos vírus envelopados em plantas conhecidos, são caracterizados por brotamento das membranas internas da célula hospedeira, e não da membrana plasmática. Embora tenha sido produzido um pequeno número de VLPs recombinantes em hospedeiros vegetais, nenhum foi derivado da membrana plasmática, levantando a questão de se as VLPs derivadas da membrana plasmática, incluindo as VLPs da gripe, podem ser produzidas nas plantas.

A formação de VLPs, em qualquer sistema, coloca demandas consideráveis na estrutura das proteínas - alterar trechos curtos de sequência que correspondem a alças de superfície selecionadas de uma estrutura globular pode não ter muito efeito na expressão do próprio polipeptídeo, no entanto estudos estruturais faltam demonstrar o efeito de tais alterações na formação de VLPs. A cooperação das várias regiões e estruturas de HA (por exemplo, as sequências de âncora de membrana, as regiões de caule ou tronco do trímero que separa a cabeça globular das membranas) evoluiu com o vírus e pode não ser passível de alterações similares sem perda de Integridade do trímero de HA e formação de VLP.

A produção de VLPs HA de gripe foi previamente descrita pelos inventores em WO 2009/009876.

SUMÁRIO DA INVENÇÃO

A presente invenção refere-se a partículas semelhantes a vírus. Mais especificamente, a presente invenção refere-se a partículas semelhantes a vírus compreendendo hemaglutinina de gripe quimérica e a métodos de produção de partículas semelhantes a vírus de hemaglutinina da gripe quiméricas.

É um objetivo da invenção proporcionar uma partícula semelhante a vírus da gripe quimérica (VLP) melhorada.

A presente invenção proporciona um polipéptido que pode ser expresso num hospedeiro vegetal, compreendendo um HA de gripe quimérico compreendendo um domínio do domínio da haste (SDC), um domínio do domínio principal (HDC) e um domínio do domínio transmembranar (TDC) em que: o SDC compreende um Subdomínio F'1, F'2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail; em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os subdomínios SDC e E1 e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é Gripe HI ou H5, e o segundo polipeptídeo de gripe HA é derivado de uma estirpe de gripe diferente

do primeiro polipeptídeo de gripe HA. Além disso, o polipéptido pode compreender um péptido sinal.

A presente invenção também proporciona um ácido nucleico que codifica o polipéptido compreendendo uma HA de gripe quimérica compreendendo um agregado de domínio de haste (SDC), um agregado de domínio de cabeça (HDC) e um agregado de domínio transmembranar (TDC) em que: a SDC compreende um F'1, Subdomínio F'2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail; em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os subdomínios SDC e EI e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é Gripe HI ou H5, e o segundo polipeptídeo de gripe HA é derivado de uma estirpe de gripe diferente do primeiro polipeptídeo de gripe HA. O ácido nucleico pode também codificar o polipéptido que compreende um péptido sinal como definido na reivindicação 2 em adição à SDC, HDC e TDC.

Também proporcionado um modo para produzir partículas semelhantes a vírus da gripe quiméricas (VLPs) numa planta, compreendendo o modo:

a) introdução de um ácido nucleico que codifica um HA da gripe quimérico compreendendo um peptídeo de sinal, um cluster de domínio de tronco (SDC), um cluster de domínio de cabeça (HDC) e um cluster de domínio transmembrana (TDC) em que: o SDC compreende um F'1, F Subdomínio 2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail; em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os subdomínios SDC e EI e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é o gripe H1 ou H5, e o segundo polipéptido de HA de gripe é derivado de uma estirpe de gripe diferente do primeiro polipéptido de HA de gripe, na planta, ou porção deste, e

b) incubar a planta, ou sua porção, sob condições que permitam a expressão do ácido nucleico, produzindo assim as VLPs.

A presente invenção inclui o método descrito acima em que no passo de introdução (passo a), o ácido nucleico é introduzido na planta de um modo transitório. Alternativamente, no passo de introdução (passo a), o ácido nucleico é introduzido na planta e é integrado de forma estável. O método pode ainda compreender um passo de: c) colher o hospedeiro e purificar as VLPs.

A presente invenção proporciona uma planta, ou porção da mesma, compreendendo uma HA de gripe quimérica, ou uma sequência de nucleótidos que codifica a gripe HA quimérica, a HA de gripe quimérica compreendendo um domínio de domínio tronco (SDC), um domínio de domínio principal (HDC) e uma membrana transmembranar. cluster de domínio (TDC) em que: o SDC compreende um subdomínio F'1, F'2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail; em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os subdomínios SDC e EI e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é o gripe H1 ou H5, e o segundo polipéptido de HA de gripe é derivado de uma estirpe de gripe diferente do que o primeiro polipéptido de HA de gripe.

A planta, ou porção desta, pode ainda compreender um ácido nucleico compreendendo uma sequência nucleotídica codificando uma ou mais de uma proteína acompanhante operativamente ligada a uma região reguladora ativa numa planta. A uma ou mais de uma proteína de acompanhante pode ser selecionada do grupo compreendendo Hsp40 e Hsp70.

A presente invenção refere-se a uma partícula semelhante a vírus (VLP) compreendendo uma HA de gripe quimérica, a HA de gripe quimérica compreendendo um domínio de domínio tronco (SDC), um domínio de domínio principal (HDC) e um domínio de domínio transmembranar (TDC) em que: compreende um subdomínio F'1, F'2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail; em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os

subdomínios SDC e EI e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é Gripe HI ou H5, e o segundo polipeptídeo de gripe HA é derivado de uma estirpe de gripe diferente do primeiro polipeptídeo de gripe HA. A VLP pode ainda compreender N-glicanos específicos de plantas, ou N-glicanos modificados.

Uma composição compreendendo uma dose eficaz da VLP como acabado de descrever e um transportador farmacêuticamente aceitável também é proporcionada.

Num aspeto alternativo da presente invenção é proporcionado um método para induzir imunidade a uma infeção por vírus da gripe num indivíduo, compreendendo a administração da VLP ao sujeito. A VLP pode ser administrada a um indivíduo por via oral, intradérmica, intranasal, intramuscular, intraperitoneal, intravenosa ou subcutânea.

As regiões reguladoras que podem ser operativamente ligadas a uma sequência que codifica uma proteína HA quimérica incluem aquelas que são operativas numa célula vegetal, numa célula de inseto ou numa célula de levedura. Tais regiões reguladoras podem incluir uma região reguladora de plastocianina, uma região reguladora de Ribulose 1,5-bisfosfato carboxilase/oxigenase (RuBisCO), proteína de ligação clorofina a/b (CAB) ou ST-LS 1. Outras regiões reguladoras incluem um 5'UTR, 3'UTR ou sequências de terminação. A região reguladora da plastocianina pode ser uma região reguladora da plastocianina de alfafa; as sequências 5'UTR, 3'UTR ou terminator também podem ser sequências de alfafa.

A presente divulgação proporciona um polipeptídeo HA de gripe quimérico compreendendo uma primeira gripe e uma segunda gripe, a primeira gripe e a segunda gripe podem ser independentemente selecionadas do grupo compreendendo H1 e H5 ; com a condição de que a primeira gripe e a segunda gripe não sejam do mesmo tipo de gripe, subtipo ou da mesma origem.

De acordo com alguns aspetos da divulgação, o polipéptido HA de gripe quimérico compreende uma sequência peptídica de sinal, a sequência peptídica sinal pode ser selecionada do grupo compreendendo uma sequência peptídica sinal nativa, uma sequência peptídica sinal de PDI de alfafa, uma sequência peptídica sinal H5 de gripe e uma sequência peptídica de sinal HI da gripe

A presente divulgação proporciona um modo para produzir uma VLP contendo hemaglutinina de gripe quimérica (HA) num hospedeiro capaz de produzir uma VLP, incluindo uma planta, inseto ou levedura compreendendo a introdução de um ácido nucleico que codifica uma gripe HA quimérica compreendendo um domínio de domínio (SDC), um cluster de domínio de cabeça (HDC) e um aglomerado de domínios transmembrana (TDC) em que: o SDC compreende um subdomínio F'1, F'2 e F; o HDC compreende um subdomínio RB, E1 e E2; o TDC compreende um subdomínio TmD e Ctail;em que o subdomínio RB é de um primeiro polipéptido HA de gripe e os subdomínios SDC e EI e E2 são de um segundo polipéptido HA de gripe e, em que o primeiro polipéptido HA de gripe é de gripe HI ou H5 e o segundo polipéptido HA de gripe é o influenza HI ou H5, e o segundo polipéptido HA de gripe é derivado de uma estirpe de gripe diferente do primeiro polipéptido HA de gripe, no hospedeiro e incubação do hospedeiro sob condições que permitem a expressão do ácido nucleico, produzindo assim as VLPs.

A produção de VLPs em plantas apresenta várias vantagens sobre a produção dessas partículas na cultura de células de insetos. Os lípidos vegetais podem estimular células imunitárias específicas e aumentar a resposta imunitária induzida. As membranas vegetais são feitas de lípidios, fosfatidilcolina (PC) e fosfatidiletanolamina (PE), e também contêm glicosfingolipídeos que são únicos para as plantas e algumas bactérias e protozoários. Os esfingolípídeos são invulgares porque não são ésteres de glicerol como PC ou PE, mas consistem de um amino álcool de cadeia longa que forma uma ligação amida a uma cadeia de ácido gordo contendo mais de 18 carbonos. PC e PE, bem como glicosfingolipídeos podem se ligar a moléculas CD1

expressas por células imunes de mamíferos, tais como células apresentadoras de antígenos (APCs), como células dendríticas e macrófagos e outras células, incluindo linfócitos B e T no timo e fígado. Além disso, para além do potencial efeito adjuvante da presença de lípidos vegetais, a capacidade dos N-glicanos vegetais para facilitar a captura de antígenos de glicoproteína por células apresentadoras de antígeno, pode ser vantajosa da produção de VLPs quiméricas em plantas. Sem querer estar limitado pela teoria, prevê-se que as VLPs quiméricas produzam uma reação imune mais forte do que as VLPs quiméricas produzidas em outros sistemas de fabricação e que a reação imune induzida por essas VLPs quiméricas é mais forte quando comparada à imunológica reação induzida por vacinas vivas inteiras ou atenuadas.

Contrariamente às vacinas feitas de vírus completos, as VLP quiméricas proporcionam a vantagem, uma vez que não são infecciosas, pelo que a contenção biológica restritiva não é um problema tão significativo como seria trabalhar com um vírus infeccioso completo e não é necessário para produção. As VLPs quiméricas de plantas proporcionam uma vantagem adicional novamente permitindo que o sistema de expressão seja cultivado em uma estufa ou campo, sendo assim significativamente mais económico e adequado para o aumento de escala.

Adicionalmente, as plantas não compreendem as enzimas envolvidas na sintetização e adição de resíduos de ácido siálico às proteínas. As VLPs podem ser produzidas na ausência de neuraminidase (NA), e não há necessidade de co-expressar NA, ou tratar as células produtoras ou extrair com sialidase (neuraminidase), para garantir a produção de VLP nas plantas.

BREVE DESCRIÇÃO DOS DESENHOS

Estas e outras características da invenção tornar-se-ão mais evidentes a partir da seguinte descrição, na qual é feita referência aos desenhos anexos, em que:

A Figura 1A mostra um diagrama esquemático dos subdomínios de HA. SP: péptido sinal, F'1, F'2 e F: subdomínios de fusão; RB: subdomínio de ligação ao recetor, E1 e E2: subdomínios de esterase, DTM/CT: subdomínios da cauda transmembrana e citoplasmática. A Figura 1B mostra representações esquemáticas de cassetes de expressão à base de plastocianina (números de construção: 774, 540, 660, 690, 691, 696) para a expressão de hemaglutinina HI A/ Brisbane/59/2007 (H1/Bri), hemaglutinina HI A/Nova Caledónia/20/99 (H1/NC) e hemaglutinina H5 A/Indonésia/5/05 (H5/Indo) em formas nativas e quiméricas. Plasto pro: promotor de plastocianina de alga, Plasto ter: terminador de plastocianina de alga, SP: peptídeo sinal, RB: subdomínio de ligação a recetor, E1-RB-E2: subdomínios de esterase e de ligação a recetor, TMD/CT: subdomínios de cauda transmembrana e citoplasmática: dissulfeto isomerase de proteína de alga.

A Figura 2 mostra a sequência de aminoácidos dos subdomínios indicados de HA quimérico expresso com painel superior, construções 690, 734 (SEQ ID NO: 11), 696 (SEQ ID NO: 112) e painel inferior, 691 (SEQ ID NO: 113). Os aminoácidos 1-92 da SEQ ID NO: 111 é o domínio F'1 + E1 do H5/Indo; os aminoácidos 93-263 são um domínio da cabeça do RB de H1/ Brisbane e os aminoácidos 264-552 são um E2 + Domínio F'2 de H5/Indo. Aminoácidos 1-92 da SEQ ID NO: 112 é o domínio F'1 + E1 de H5/NC, aminoácidos 93-301 é um domínio de cabeça RB de H5/Indo e aminoácidos 302-586 são um domínio E2 + F'2 de H1/NC os aminoácidos 1-42 da SEQ ID NO: 113 são o domínio F'1 de H5/Indo; os aminoácidos 43-273 são um E1-RB- O domínio de cabeça E2 de H1/Brisbane e os aminoácidos 274-552 é um domínio F'2 de H5/Indo.

A Figura 3 mostra a sequência de aminoácidos da região codificadora das construções 690 e 734 (SEQ ID NO: 80) compreendendo um subdomínio RB de H1/Bri, um péptido sinal H5/Indo e um domínio complexo de tronco (SDC) compreendendo subdomínios H5/Indo F'1, E1, E2, F'2 e F.

A Figura 4 mostra a sequência de aminoácidos da região de codificação da construção 691 (SEQ ID NO: 81) compreendendo o complexo do domínio

da cabeça H1/Bri (HDC) que contém E1, RB, E2, um péptido sinal H5/Indo e H5/Indo Complexo do domínio tronco (SDC) compreendendo subdomínios H5/Indo F'1, F'2 e F.

A Figura 5 mostra a sequência de aminoácidos da região de codificação da construção 696 (SEQ ID NO: 82) compreendendo um subdomínio RB de H5/Indo, um péptido de sinalização PDI e complexo de domínio de haste H1/NC compreendendo F'1, E1, E2 e F'2.

A Figura 6 mostra uma análise de imunotransferência da expressão de H1/Bri na forma nativa, construção 774 (compreendendo H1/Bri), construção 692 (compreendendo o complexo de domínio de cabeça (HDC) de H1/Bri) e construção 690 (compreendendo o subgrupo RB domínio de H1/Bri fundido com o complexo H5/Indo do domínio de tronco (SDC) em plantas. Para cada construção, foram analisados extratos proteicos totais de 3 plantas separadas. Vinte microgramas de proteína foram carregados para cada planta analisada. Anticorpos monoclonais anti-HA (anti-H1-Brisbane; FII 10-150) A construção 774 expressa H1/Bri com o péptido de sinal nativo de H1/Bri: as construções 690, 691 expressam o HA com o péptido de sinal nativo de H5/Indo.

A Figura 7 mostra uma análise de imunotransferência da expressão de H5/Indo na forma nativa, construção 660 (compreendendo H5/Indo, ou construção 696 (compreendendo subdomínio H1/Indo RB fundido com subdomínios H1/NC SDC, E1 e E2), foram analisados extratos proteicos totais de 3 plantas separadas, 20 microgramas de proteínas foram carregados para cada planta analisada e o Western blot foi revelado com anticorpos policlonais anti-H5 da Indonésia (ITC IT-003-005V). A construção 660 expressa H5/Indo como seu péptido sinal nativo, a construção 696 expressa o HA quimérico com um péptido sinal de PDI.

A Figura 8 mostra uma representação esquemática de cassetes de expressão baseadas em 35SCPMV/HT para a expressão de H1/ Bri nas formas nativa (construção 732) e quimérica (construções 733 e 734). Construção 733, compreendendo o péptido de sinal PDI e complexo HDC,

SDC e domínio transmembranar (TDC) de H1/Bri, e construção 734 compreendendo um péptido de sinal H5/Indo, F'1, E1, E2, F'2, F, e um RB de H1/Bri. Promotor 35S pro: CaMV 35S, NOSTer: terminador da nopalina sintase, SP: peptídeo sinal, RB: subdomínio de ligação ao recetor, E1-RB-E2: subdomínios de esterase e de ligação ao recetor, TMD/CT: subdomínios da cauda transmembrana e citoplasmática, PDI: dissulfeto isomerase de proteína de alfafe; CPMV-HT: elementos 5' e 3' do sistema de expressão do vírus hipercrítico do mosaico do caupi.

A Figura 9 mostra a análise de imunotransferência da expressão de H1/Bri na forma nativa, construção 732 (compreendendo H1/ Bri sob o controlo da cassette de expressão baseada em 35SCPMV/HT), construção 733 (compreendendo um péptido de sinal PDI fundido com H1/Bri; sob o controlo da cassette de expressão com base em 35SCPMV/HT), ou construção 734 compreendendo um subdomínio H1/Bri RB fundido com um subdomínio H5/Indo SDC, E1 e E2; sob o controlo da cassette de expressão com base em 35SCPMV/HT). Para cada construção, foram analisados extratos proteicos totais de 3 plantas separadas. Cinco microgramas de proteína foram carregados para cada planta analisada. O Western blot foi revelado com anticorpos monoclonais anti-HA (FII 10-150).

A Figura 10 mostra a representação esquemática de cassetes de expressão baseadas em 35SCPMV/HT para a expressão de hemaglutininas H3A/Brisbane / 10/2007 HA (H3/Bri) e B/ Flórida/4/2006 HA (B/Flo). A construção 736 compreende H3/ Bri fundido a um péptido sinal da PDI. A construção 737 compreende o H3/Bri fundido com um péptido sinal PDI e um H5/Indo TMD/CT. A construção 739 compreende B/Flo fundida a um péptido sinal PDI. A construção 745 compreende o B/Flo fundido com um péptido sinal PDI e um H5/Indo TMD/ CT. Promotor 35S pro: CaMV 35S, NOSTer: terminador da nopalina sintase, SP: peptídeo sinal, RB: subdomínio de ligação ao recetor, E1-RB-E2: subdomínios de esterase e de ligação ao recetor, TMD/CT: subdomínios da cauda transmembrana e citoplasmática, PDI: dissulfeto isomerase de proteína de alfafe; CPMV-HT: elementos 5 'e 3' do sistema de expressão do vírus hipercrítico do mosaico do caupi.

A Figura 11 mostra a borda de fusão nas construções número 745 e 737. A origem da sequência de HA é indicada por setas com ponta de bala. Os aminoácidos do domínio transmembrana são QILSIYSTVA e são precedidos por aminoácidos que fazem parte do ectodomínio.

A Figura 12 mostra a sequência de aminoácidos da hemaglutinina H5/H3 quimérica (SEQ ID NO: 83; estrutura 737) compreendendo um peptídeo sinal PDI, um ectodomínio de H3 A/ Brisbane/10/2007 e um TMD/CT de H5 A/Indonésia/5/2005.

A Figura 13 mostra a sequência de aminoácidos da hemaglutinina H5/B quimérica (SEQ ID NO: 84) compreendendo um ectodomínio de B/Florida/4/2006 e uma DTM/CT de H5 A/ Indonésia/5/2005 codificada por o quadro de leitura aberto no número de construção 745.

A Figura 14 mostra a análise de imunotransferência da expressão de B/Flo na forma nativa, a construção 739 (compreendendo PDI-B/Flo), ou com a construção 745 (compreendendo B/Flo HDC e SDC fundida com um H5/Indo TDC). Para cada construção, foram analisados extratos proteicos totais de 3 plantas separadas. Vinte microgramas de proteína foram carregados para cada planta analisada. O Western blot foi revelado com anticorpos policlonais anti-HA B/Florida (NIBSC 07/356).

A Figura 15 mostra a análise de imunotransferência da expressão de H3/Bri na forma nativa, construção 736 (compreendendo PDI sp-H3/Bri) ou com a construção 737 (H3/ Bri HDC e SCD fundidos com um H5/Indo TDC). Para cada construção, foram analisados extratos proteicos totais de 3 plantas separadas. Vinte microgramas de proteína foram carregados para cada planta analisada. O Western blot foi revelado com anticorpos policlonais anti-H3 Brisbane (NIBSC 08/124).

A Figura 16 mostra a cromatografia por exclusão de tamanho de extratos proteicos de folhas de plantas infiltradas com o número de construção 745. O teor relativo de proteína das frações de eluição é apresentado para cada fração. A imunodeteção (Western Blot) de hemaglutinina utilizando anticorpos policlonais anti-HAB/Florida (NIBSC 07/356) nas

frações 7 a 15 é apresentada sob o gráfico. O pico de eluição do Dextrano Azul 2000 é indicado pela seta (fração 8).

A Figura 17 mostra a sequência de ácido nucleico (SEQ ID NO: 52) do fragmento sintetizado compreendendo a região codificante H5 completa (A/Indonésia/5/05 (H5N1)) (incluindo o péptido sinal e o codão stop) flanqueada, em 5' por um sítio HindIII e, em 3', por um sítio SacI.

A Figura 18 mostra a sequência de ácido nucleico (SEQ ID NO: 53) da construção 660, uma cassette de expressão HA compreendendo um promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina de H5 forma A/ Indonésia/5/05 (H5N1), alfafa plastocianina 3'UTR e sequências terminadoras.

A Figura 19 mostra a sequência de ácido nucleico (SEQ ID NO: 54) da sequência codificadora de H1 do tipo selvagem (A/ New Caledonia/20/99 (H1N1) (GenBank AY289929) sem TmD e Ctail.

A Figura 20 mostra a sequência de ácido nucleico (SEQ ID NO: 55) de um fragmento sintetizado compreendendo H1 (A/Nova Caledónia/20/99 (H1N1) sequência codificante sem TmD e Ctail. Na região 5', os últimos nucleótidos são originários de PDI SP e incluem um sítio de restrição BglII e em 3', um local duplo SacI/StuI é encontrado imediatamente a jusante do codão de terminação.

A Figura 21 mostra a sequência de ácido nucleico (SEQ ID NO: 56) do fragmento sintetizado compreendendo a sequência de codificação C-ter H1 (A/Nova Caledónia/20/99 (H1N1) incluindo o TmD e Ctail do local KpnI até ao codão de terminação (flanqueado em 3'por um local duplo SacI/StuI).

A Figura 22 mostra a sequência de nucleótidos para ARNm de *Medicago sativa* para dissulfureto isomerase de proteína. GenBank Z11499 (SEQ ID NO: 57). Os nucleótidos 32-103 codificam o péptido sinal da PDI.

A Figura 23 mostra a sequência de nucleótidos para o plasmídeo PromPlasto-PDISP-Plasto 3'UTR. A Figura 23A mostra a sequência de nucleótidos para PromPlasto-PDISP (SEQ ID NO: 58). A Figura 23B mostra a sequência nucleotídica de Plasto 3'UTR (SEQ ID NO: 85). O péptido sinal da isomerase de dissulfureto de proteína (PDI) está sublinhado. Os locais de restrição BglIII (AGATCT) e SacI (GAGCTC) utilizados para clonagem são mostrados em negrito.

A Figura 24 mostra a sequência de ácido nucleico (SEQ ID NO: 59; construção 540) da cassete de expressão HA compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência codificadora do péptido sinal da PDI e da forma H1 A/New Caledonia/20/99 (H1N1), plastocianina de alfafa 3'UTR e sequências de terminação. H1 da sequência de codificação A/ Nova Caledónia/20/1999 está sublinhada.

A Figura 25 mostra a sequência de ácido nucleico (SEQ ID NO: 60) do fragmento sintetizado compreendendo a região codificante HI completa (A/Brisbane/59/07 (H1N1)) (incluindo o péptido de sinal e o codão stop) flanqueada, em 5' por sequências de genes de plastocianina de alfafa correspondentes aos primeiros 84 nucleótidos a montante do ATG inicial, começando com um local DraIII e, em 3', por um local SacI.

A Figura 26 mostra a sequência de ácido nucleico (SEQ ID NO: 61; construção 774) da cassete de expressão HA compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina de H1 forma A/Brisbane/59/07 (H1N1), alfafa. plastocianina 3'UTR e sequências terminadoras.

A Figura 27 mostra a sequência de ácido nucleico da cassete de expressão número 828 (SEQ ID NO: 62), de PacI (a montante do promotor) a AscI (imediatamente a jusante do terminador NOS). **negrito**. Local de restrição **Apal** sublinhado e *itálico*.

A Figura 28 mostra a sequência de ácido nucleico (SEQ ID NO: 63; construção 690) de uma cassete de expressão quimérica H5/H1 compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência

de codificação de hemaglutinina quimérica, plastocianina de alfafa 3' UTR e sequências de terminação. A sequência de codificação de HA quimérica está sublinhada.

A Figura 29 mostra a sequência de ácido nucleico (SEQ ID NO: 64; construção 691) de uma cassette de expressão quimérica H5/H1 compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina quimérica, plastocianina de alfafa 3' UTR e sequências de terminação. A sequência de codificação de HA quimérica está sublinhada.

A Figura 30 mostra a sequência de ácido nucleico (SEQ ID NO: 65; construção 696) de uma cassette de expressão H1/H5 quimérica compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina quimérica, UTR de plastocianina de alfafa e sequências terminadoras. A sequência de codificação de HA quimérica está sublinhada.

A Figura 31 mostra a sequência de ácido nucleico (SEQ ID NO: 66; construção 732) da cassette de expressão HA compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência de codificação da hemaglutinina de HI forma A/Brisbane/59/07 (H1N1), CPMV-HT 3'UTR e sequências de terminação NOS. A sequência de codificação de H1/Bri está sublinhada.

A Figura 32 mostra a sequência de ácido nucleico (SEQ ID NO: 67) da sequência de codificação, de ATG para parar, do número de construção intermédio 787.

A Figura 33 mostra a sequência de ácido nucleico (SEQ ID NO: 68; construção 733) da cassette de expressão SpPDI H1/Bri compreendendo o promotor 35M de CaMV, CPMV-HT 5'UTR, sequência codificadora do péptido de sinal da PDI, sequência de codificação de hemaglutinina de sequências A/Brisbane/ 59/07 (H1N1), CPMV-HT 3'UTR e terminação NOS. A sequência de codificação do SpPDI H1/Bri está sublinhada.

A Figura 34 mostra a sequência de ácido nucleico (SEQ ID NO: 69; construção 734) de uma cassette de expressão quimérica H5/H1 compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência de codificação de hemaglutinina quimérica, CPMV-HT 3' UTR e sequências do terminador NOS. A sequência de codificação do HA quimérico está sublinhada.

A Figura 35 mostra a sequência de ácido nucleico (SEQ ID NO: 70) do fragmento sintetizado compreendendo a região codificante H3 completa (A/Brisbane/10/07 (H3N2)) (incluindo o péptido de sinal e o codão stop) flanqueada, em 5'por sequências de genes de plastocianina de alfafa correspondentes aos primeiros 84 nucleótidos a montante do ATG inicial, começando com um local DraIII e, em 3', por um local SacI.

A Figura 36 mostra a sequência de ácido nucleico (SEQ ID NO: 71; construção 736) da cassette de expressão HA compreendendo o promotor 35M de CaMV, CPMV-HT 5 'UTR, sequência codificadora do péptido sinal da PDI, sequência de codificação de hemaglutinina H3 forma A/Brisbane/10/07 (H2N3), CPMV-HT 3'UTR e sequências de terminação NOS. A sequência de codificação de Sp PDI H3/Bris está sublinhada.

A Figura 37 mostra a sequência de ácido nucleico (SEQ ID NO: 72; construção 737) de uma cassette de expressão quimérica H5/H3 compreendendo o promotor 35M de CaMV, CPMV-HT 5'UTR, sequência de codificação de hemaglutinina quimérica, CPMV-HT3'. Sequencias terminator UTR e NOS. A sequência de codificação do HA quimérico está sublinhada.

A Figura 38 mostra a sequência de ácido nucleico (SEQ ID NO: 73) do fragmento sintetizado compreendendo a região codificante completa de HA (B/Florida/4/06) (incluindo o péptido de sinal e o codão stop) flanqueada, em 5'por alfafa. sequências do gene da plastocianina correspondentes aos primeiros 84 nucleótidos a montante do ATG inicial, começando com um local DraIII e, em 3', por um sítio SacI.

A Figura 39 mostra a sequência de ácido nucleico (SEQ ID NO: 74; construção 739) da cassette de expressão HA compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência codificadora do péptido sinal da PDI, sequência de codificação da hemaglutinina de HA forma B/Florida/4/06, CPMV-HT 3'UTR e sequências terminadoras NOS. A sequência de codificação de Sp PDI B/Flo está sublinhada.

A Figura 40 mostra a sequência de ácido nucleico (SEQ ID NO: 75; construção 745) de uma cassette de expressão quimérica H5/B compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência de codificação de hemaglutinina quimérica, CPMV-HT 3' UTR e sequências do terminador NOS. A sequência de codificação do HA quimérico está sublinhada.

A Figura 41 mostra a sequência de ácido nucleico que codifica Msj1 (SEQ ID NO: 76).

A Figura 42 mostra a sequência de ácido nucleico (SEQ ID NO: 77) de uma porção do número de construção R850, de HindIII (no local de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). A sequência de codificação da HSP40 está sublinhada.

A Figura 43 mostra a sequência de ácido nucleico (SEQ ID NO: 78) de uma porção do número de construção R860, de HindIII (no local de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). A sequência de codificação da HSP70 está sublinhada.

A Figura 44 mostra a sequência de ácido nucleico (SEQ ID NO: 79) de uma porção do número de construção R870, de HindIII (no local de clonagem múltipla, 5 a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). A sequência de codificação da HSP40 está em sublinhado em itálico e a sequência de codificação da HSP70 está sublinhada. A) nucleótidos 1-4946; B) nucleótidos 4947-9493.

A figura 45 mostra uma representação esquemática do número de construção R472.

A Figura 46 mostra o padrão da ponte dissulfureto do tipo A da gripe. 1) Cys4HA1-Cys 137HA2, 2) Cys60HA1-Cys72HA1, 3) Cys94HA1-Cys143HA1, 4) Cys292HA1-Cys318HA1 5) Cys144HA2-Cys148HA2 e 6) Cys52HA1-Cys277HA1. As pontes dissulfureto que diferem entre os subtipos A e B (Figura 47) são indicadas por setas. A numeração da proteína H3 madura foi usada.

A Figura 47 mostra o padrão da ponte dissulfureto do tipo HA da gripe. Numeração de ponte: 1) Cys4HA1-Cys137HA2, 2) Cys60HA1-Cys72HA1, 3) Cys94HA1-Cys 143HA1, 4) Cys292HA1-Cys318HA1 5) Cys144HA2-Cys148HA2, 6) Cys52HA1-Cys277HA1, 7) Cys54HA1-Cys57HA1 e 8) Cys178HA1-Cys272HA1 As pontes dissulfureto que diferem entre os subtipos A (Figura 46) e B são indicadas por setas. A numeração da proteína H3 madura foi usada.

A Figura 48 mostra um diagrama esquemático de junções de fusão de troca de domínio. A Figura 48A mostra a fusão do subdomínio RB de H1/Bri, H3/Bri e B/Flo com H5/Indo SDC, e o subdomínio RB de H5/Indo com o domínio de tronco H1/NC. A Figura 48B mostra a fusão dos subdomínios E1-RB-E2 (HDC) de H1/Bri, H3/Bri ou B/Flo com H5/Indo SDC, e de H5/Indo HDC sem H1/NC SDC.

A Figura 49A mostra a sequência nucleotídica SEQ ID NO: 86) de H1 A/Califórnia/04/09. A sequência de codificação do péptido de sinal da isomerase de dissulfureto de proteína de alfafa é sublinhada e a sequência de codificação de H1 madura é destacada a negrito. A Figura 49B mostra a sequência de aminoácidos (SEQ ID NO: 87) do H1 A / Califórnia / 04/09. O péptido sinal isomerase de dissulfureto de proteína de alfafa está sublinhado.

A Figura 50 mostra uma análise de imunotransferência da expressão de hemaglutinina H5/B quimérica (número de construção 747; compreendendo B/Flo HDC e SDC fundido com um H5/Indo TDC) após infiltração de

AGL1/747 não diluído, co-infiltrado com AGL1/443 (vetor vazio) e co-infiltrado com AGL1/R870 (HSP40/HSP70). Para cada construção, foram analisados extratos proteicos totais de 3 plantas separadas. Vinte microgramas de proteínas foram carregados para cada planta analisada. O Western blot foi revelado com anticorpos policlonais anti-B Florida (NIBSC).

A Figura 51A mostra a sequência nucleotídica para a sequência promotora 2X35S (SEQ ID NO: 88). A Figura 51B mostra a sequência de nucleótidos para a construção 747 (SEQ ID NO: 93) de PacI (promotor 35S a montante) para AscI (terminador NOS imediatamente a jusante). A sequência de codificação do HA quimérico está sublinhada. A sequência do promotor 2X35S é indicada em *itálico*.

DESCRIÇÃO DETALHADA

A presente invenção refere-se a partículas semelhantes a vírus. Mais especificamente, a presente invenção refere-se a partículas semelhantes a vírus compreendendo hemaglutinina da gripe quimérica e a métodos de produção de partículas quiméricas semelhantes a vírus da gripe.

A descrição que se segue é de uma forma de realização preferida.

A presente invenção proporciona um ácido nucleico compreendendo uma sequência nucleotídica que codifica uma hemaglutinina da gripe quimérica (HA) operativamente ligada a uma região reguladora ativa numa planta.

Além disso, a presente invenção fornece um método de produção de partículas semelhantes a vírus (VLPs) em uma planta. O método envolve a introdução de um ácido nucleico que codifica um HA gripe quimérico operativamente ligado a uma região reguladora ativa na planta, na planta, ou porção da planta, e incubando a planta ou uma porção da planta sob condições que permitam a expressão de o ácido nucleico, produzindo assim as VLPs.

A presente invenção proporciona ainda uma VLP compreendendo uma gripe HA quimérica. A VLP pode ser produzida pelo método como proporcionado pela presente invenção.

Por "proteína quimérica" ou "polipéptido quimérico" entende-se uma proteína ou polipéptido que compreende sequências de aminoácidos de duas ou mais do que duas fontes, por exemplo, mas não limitadas a dois ou mais tipos ou subtipos de gripe, ou gripe de umas diferentes origens, que são fundidas como um único polipeptídeo. A proteína ou polipéptido quimérico pode incluir um péptido sinal que é o mesmo, ou heterólogo, com o restante do polipéptido ou proteína. A proteína quimérica ou polipéptido quimérico pode ser produzida como um transcrito de uma sequência nucleotídica quimérica, e a proteína quimérica ou polipéptido quimérico clivado após a síntese e, como requerido, associado para formar uma proteína multimérica. Assim sendo, uma proteína quimérica ou um polipéptido quimérico também inclui uma proteína ou polipéptido compreendendo subunidades que estão associadas através de pontes de dissulfureto (isto é, uma proteína multimérica). Por exemplo, um polipéptido quimérico compreendendo sequências de aminoácidos de duas ou mais do que duas fontes podem ser processadas em subunidades, e as subunidades associadas via pontes de dissulfureto para produzir uma proteína quimérica ou polipéptido quimérico (ver Figuras 46 e 47). O polipéptido pode ser hemaglutinina (HA), e cada uma das duas ou mais do que duas sequências de aminoácidos que formam o polipéptido podem ser obtidas a partir de HA diferentes para produzir um HA quimérico, ou HA gripe quimérico. Um HA quimérico também pode incluir uma sequência de aminoácidos compreendendo péptido de sinal heterólogo (uma proteína quimérica de HA) que clivado após ou durante a síntese proteica. Preferencialmente, o polipéptido quimérico, ou o HA de gripe quimérico não ocorre naturalmente. Um ácido nucleico que codifica um polipéptido quimérico pode ser descrito como um "ácido nucleico quimérico", ou uma "sequência nucleotídica quimérica". Uma partícula semelhante a vírus compreendida por HA quimérico pode ser descrita como uma "VLP quimérica".

A HA de gripe quimérica de acordo com várias formas de realização da presente invenção pode compreender um complexo de domínio de tronco (SDC) um complexo de domínio de cabeça (HDC) e um complexo de domínio transmembranar (TDC), onde um ou mais de um subdomínio de SDC, HDC ou TDC é de um primeiro tipo de HA de gripe, subtipo ou de uma origem, e um ou mais de um subdomínio de SDC, HDC ou TDC é de um segundo tipo de HA de gripe, subtipo ou de uma segunda origem ou diferente. Como aqui descrito, o "SDC" compreende um subdomínio F'1, F'2 e F, o "HDC" compreende um subdomínio RB, E1 e E2, o "TDC" compreende um subdomínio TmD e Ctail (TMD/CT; ver Figuras 1A, 46 e 47).

O termo "partícula semelhante a um vírus" (VLP) ou "partículas semelhantes a vírus" ou VLPs refere-se a estruturas que se automontam e compreendem proteínas estruturais, tais como proteína HA de gripe ou proteína HA de gripe quimérica. VLPs e VLPs quiméricas são geralmente morfológicas e antigenicamente semelhantes aos viriões produzidos em uma infecção, mas faltam informações genéticas suficientes para replicar e, portanto, não-infeciosas. As VLPs e as VLPs quiméricas podem ser produzidas em células hospedeiras adequadas, incluindo células hospedeiras vegetais. Após a extração da célula hospedeira e após isolamento e purificação adicional sob condições adequadas, as VLP e as VLPs quiméricas podem ser purificadas como estruturas intactas.

VLPs quiméricas, ou VLPs, produzidas a partir de proteínas derivadas de gripe, de acordo com a presente invenção, não compreendem proteína M1. Sabe-se que a proteína M1 se liga ao ARN (Wakefield e Brownlee, 1989) que é um contaminante da preparação de VLP. A presença de ARN indesejado quando se obtém a aprovação reguladora para o produto VLP quimérico, pelo que uma preparação de VLP quimérica sem ARN pode ser vantajosa.

As VLPs quiméricas da presente invenção podem ser produzidas numa célula hospedeira que é caracterizada por não possuir a capacidade de sialilar proteínas, por exemplo, uma célula vegetal. Ver, por exemplo, Gupta et al., 1999. *Nucleic Acids Research* 27:370-372; Toukach et al.,

2007. *Nucleic Acids Research* 35:D280-D286; Nakahara et al., 2008. *Nucleic Acids Research* 36:D368-D371. As VLPs quiméricas produzidas como aqui descrito não compreendem tipicamente neuraminidase (NA). No entanto, a NA pode ser co-expressa com HA se as VLPs compreendendo HA e NA forem desejadas.

A invenção também proporciona VLP compreendendo HA quimérico que obtém um envelope lipídico da membrana plasmática da célula na qual a HA quimérica é expressa. Por exemplo, se o HA quimérico é expresso num sistema à base de plantas, a VLP resultante pode obter um envelope lipídico a partir da membrana plasmática da célula vegetal.

Geralmente, o termo "lípidio" refere-se a moléculas naturalmente solúveis em gordura (lipofílicas). Uma VLP quimérica produzida numa planta de acordo com alguns aspetos da invenção pode ser complexada com lípidos derivados de plantas. Os lípidos derivados de plantas podem estar na forma de uma bicamada lipídica e podem ainda compreender um envelope envolvendo a VLP. Os lípidos derivados de plantas podem compreender componentes lipídicos da membrana plasmática da planta onde é produzida a VLP, incluindo fosfolípidos, tri-, di- e monoglicéridos, bem como esterois solúveis em gordura ou metabolitos compreendendo esteróis. Exemplos incluem fosfatidilcolina (PC), fosfatidiletanolamina (PE), fosfatidilinositol, fosfatidilserina, glicoesfingolipídeos, fitoesteróis ou uma combinação destes. Um lípidio derivado de plantas pode alternativamente ser referido como um "lípidio vegetal". Exemplos de fitoesteróis incluem campesterol, estigmasterol, ergosterol, brassicasterol, delta-7-estigmasterol, delta-7-avenasterol, daunosterol, sitosterol, 24-metilcolesterol, colesterol ou beta-sitosterol - ver, por exemplo, Mongrand et al., 2004. Como um perito na arte entenderia, a composição lipídica da membrana plasmática de uma célula pode variar com as condições de cultura ou crescimento da célula ou organismo, ou espécie, a partir das quais a célula é obtida. Geralmente, o beta-sitosterol é o fitoesterol mais abundante. a composição lipídica da membrana plasmática de uma célula pode variar com a cultura ou condições de

crescimento da célula ou organismo, ou espécie, a partir das quais a célula é obtida. Geralmente, o beta-sitosterol é o fitoesterol mais abundante. a composição lipídica da membrana plasmática de uma célula pode variar com a cultura ou condições de crescimento da célula ou organismo, ou espécie, a partir das quais a célula é obtida. Geralmente, o beta-sitosterol é o fitoesterol mais abundante.

As membranas celulares compreendem geralmente bicamadas lipídicas, bem como proteínas para várias funções. Concentrações localizadas de lipídios particulares podem ser encontradas na bicamada lipídica, conhecida como 'lipid rafts'. Estes microdomínios de jangada lipídica podem ser enriquecidos em esfingolipídeos e esteróis. Sem querer estar limitado pela teoria, jangadas lipídicas podem ter papéis significantes em endo e exocitose, entrada ou saída de vírus ou outros agentes infecciosos, transdução de sinal intercelular, interação com outros componentes estruturais da célula ou organismo, tais como intracelular e matrizes extracelulares.

A divulgação inclui VLPs compreendendo HA quimérico, dos quais os subdomínios podem ser obtidos de qualquer tipo, subtipo de vírus gripe que pode infectar humanos, incluindo, por exemplo, tipos ou subtipos H1 e H5. Exemplos não limitantes de tipos ou subtipos H1 ou H5 incluem o subtipo A/Nova Caledónia/20/99 (H1N1) ("H1/NC"; SEQ ID NO: 56), o subtipo H1 A/Califórnia 04/09 (H1N1). ("H1/Cal"; SEQ ID NO: 86), e o subtipo A/Indonésia/5/05 (H5N1) ("H5/Indo"), bem como A/Brisbane/59/2007 ("H1/Bri"). Além disso, o HA quimérico pode compreender um ou mais subdomínios de uma hemaglutinina que é isolada de um ou mais vírus da gripe emergentes ou recentemente identificados.

A presente divulgação também diz respeito a vírus gripe que infectam outros mamíferos ou animais hospedeiros, por exemplo humanos, primatas, cavalos, porcos, aves, aves aquosas aviárias, aves migratórias, codornizes, patos, gansos, aves, galinhas, camelos, caninos, outros felinos, gatos, tigres, leopardos, civetas, martas, martas de pedra, furões, animais domésticos, gado, ratinhos, ratos,

focas, baleias e similares. Alguns vírus da gripe podem infectar mais de um animal hospedeiro.

Com referência ao vírus da gripe, o termo "hemaglutinina" ou "HA", tal como aqui utilizado, refere-se a uma glicoproteína estrutural de partículas virais da gripe. A estrutura da hemaglutinina da gripe é bem estudada e demonstra um alto grau de conservação na estrutura secundária, terciária e quaternária. Esta conservação estrutural é observada mesmo que a sequência de aminoácidos possa variar (ver, por exemplo, Skehel e Wiley, 2000 *Ann Rev Biochem* 69: 531-69; Vaccaro et al 2005; que é aqui incorporado por referência). As sequências de nucleótidos que codificam para HA são bem conhecidas e estão disponíveis, por exemplo, na Base de Dados de BioDefesa e Saúde Pública (por exemplo, em [URL: biohealthbase.org/GSearch/home.do?decorator=Influenza](http://biohealthbase.org/GSearch/home.do?decorator=Influenza)) ou nas bases de dados mantidas pelo National Center for Biotechnology Information (NCBI; por exemplo, na [URL: ncbi.nlm.nih.gov/sites/entrez?db=nucleotide&cmd=search&term=influenza](http://ncbi.nlm.nih.gov/sites/entrez?db=nucleotide&cmd=search&term=influenza)), ambas aqui incorporadas por referência.

O monômero de HA pode ser subdividido em três domínios funcionais - um domínio tronco, ou domínio do domínio tronco (SDC), um domínio da cabeça globular ou domínio do domínio principal (HDC) e um domínio do domínio transmembranar (TDC). A SDC compreende quatro subdomínios, um péptido de fusão F, F'1 e F'2 (este subdomínio pode ser geralmente referido como uma 'espinha dorsal'). O CDT compreende dois subdomínios, a transmembrana (TmD) e uma cauda terminal C (CT). O HDC compreende três subdomínios, domínios esterase vestigial E1 'e E2 e um domínio de ligação ao recetor RB. A SDC e a HDC podem ser coletivamente referidas como 'ectodomínio'. Uma publicação de Ha et al. 2002 (*EMBO J.* 21: 865-875; que é aqui incorporado por referência) ilustra a orientação relativa dos vários subdomínios da SDC e HDC em vários subtipos de gripe, com base em estruturas cristalográficas de raios X. Um diagrama esquemático dos subdomínios relativos aos terminais N e C dos polipéptidos HA1 e HA2 é mostrado na Figura 1A. Um

alinhamento estrutural anotado de vários subtipos de gripe é fornecido na Figura 1C.

A variação de aminoácidos é tolerada em hemaglutininas do vírus gripe. Essa variação fornece novas estirpes continuamente identificadas. A infectividade entre as novas estirpes pode variar. Contudo, a formação de trómeros de hemaglutinina, que subsequentemente formam VLPs é mantida. A presente invenção, por conseguinte, proporciona uma sequência de aminoácidos de hemaglutinina compreendendo HA quimérico ou um ácido nucleico que codifica uma sequência de aminoácidos de hemaglutinina quimérica, que forma VLPs numa planta e inclui sequências conhecidas e sequências de HA variantes que podem desenvolver-se. A presente invenção também se refere à utilização de um polipéptido de HA quimérico compreendendo um TDC, SDC e HDC. Por exemplo, a proteína quimérica HA pode ser HA0, ou um HA quimérico clivado compreendendo subdomínios de HA1 e HA2 de dois ou mais tipos de gripe.

O HA0 pode ser expresso e dobrado para formar um trómero, que pode posteriormente ser montado em VLPs. A clivagem de HA0 produz polipéptidos HA1 e HA2 ligados por uma ponte dissulfureto (ver Figuras 1C, 46 e 47 para ilustração dos padrões da ponte dissulfureto). Para uma partícula de vírus infecciosa, é necessária a clivagem do precursor HA0 para desencadear a alteração conformacional de HA2 que liberta o péptido de fusão (no terminal N do polipéptido HA2) e disponibilizá-lo para a fusão das membranas celular e viral. No entanto, as VLPs não são infecciosas e a clivagem do HA em HA1 e HA2 não é necessária, por exemplo, para a produção de vacinas. O precursor HA0 não-clivado também se reúne em trómeros e brota a partir da membrana plasmática para formar nanopartículas de VLP.

O polipéptido HA0 compreende vários domínios. O subdomínio RB do HDC compreende vários laços em regiões antigénicas designadas como sítio AE. Anticorpos que podem neutralizar o vírus da gripe infecciosa são frequentemente direcionados para um ou mais desses locais. Os subdomínios da esterase vestigial (E1 e E2) podem ter um papel na

fusão e podem se ligar ao Ca^{++} . Os domínios F, F'1 e F'2 interagem e cooperam para formar uma haste, elevando a cabeça do trímero HA acima da membrana. Um TmD e CT podem estar envolvidos na ancoragem do HA dobrado a uma membrana. O TmD pode ter um papel na afinidade do HA para jangadas lipídicas, enquanto o CT pode ter um papel na secreção de HA, enquanto alguns dos resíduos de cisteína encontrados no subdomínio do CT podem ser palmitoilados. Um péptido sinal (SP) também pode ser encontrado no terminal N do polipéptido HA0. A Figura 2 e as Tabelas 4 e 5 fornecem exemplos das sequências de aminoácidos dos domínios SP, F'1, F'2, E1, RB, E2 e F de alguns subtipos do vírus da gripe.

O processamento de uma sequência de péptido sinal N-terminal (SP) durante a expressão e/ou secreção de hemaglutininas de gripe pode ter um papel no enrolamento do HA. O termo "péptido sinal" refere-se geralmente a uma sequência curta (cerca de 5-30 aminoácidos) de aminoácidos, encontrada geralmente no terminal N de um polipéptido de hemaglutinina que pode dirigir a translocação do polipéptido recentemente traduzido para um organelo particular, ou auxiliar no posicionamento de domínios específicos da cadeia polipeptídica em relação a outros. O péptido sinal das hemaglutininas visa a translocação da proteína para o retículo endoplasmático e foi proposto para ajudar no posicionamento do domínio proximal do terminal N em relação a um domínio âncora da membrana do polipéptido hemaglutinina nascente para ajudar na clivagem e dobragem do hemaglutinina madura.

A inserção de HA dentro da membrana do retículo endoplasmático (ER) da célula hospedeira, a clivagem do peptídeo sinal e a glicosilação da proteína são eventos co-translacionais. O correto enrolamento de HA requer glicosilação da proteína e formação de, pelo menos, 6 ligações dissulfureto intra-cadeia (ver Figuras 46 e 47). Na Figura 46, o HA do subtipo A mostra ter 6 pontes dissulfureto conservadas por monômero. Por comparação, o monômero de B HA (Figura 47) tem sete pontes dissulfeto, e cinco dessas pontes dissulfeto têm uma contraparte em A (revisado em Skehel e Wiley, 2000. Ann Rev

Biochem. 69: 531-569; exemplos de estruturas ilustrando pontes dissulfureto intra e intermoleculares e outros aminoácidos conservados e suas posições relativas são descritos, por exemplo, em Gamblin et al 2004, Science 303: 1838-1842; ambos os quais são aqui incorporados por referência). Como um perito na arte perceberia, é importante assegurar que um arranjo semelhante de pontes dissulfureto seja obtido quando se preparam HAs quiméricos.

Um péptido de sinal pode ser nativo para a hemaglutinina, ou um péptido de sinal pode ser heterólogo em relação de sequência primária da hemaglutinina a ser expressa. Um HA quimérico pode compreender um péptido de sinal de um primeiro tipo, subtipo ou estirpe de gripe com o equilíbrio do HA de um ou mais do que um tipo, subtipo ou estirpe de gripe diferente. Por exemplo, as SP nativas dos subtipos HA H1, H2, H3, H5, H6, H7, H9 ou gripe tipo B podem ser utilizadas para expressar o HA num sistema vegetal. Em algumas formas de realização da invenção, o SP pode ser de um tipo de gripe B, H1, H3 ou H5; ou do subtipo H1/Bri, H1/NC, H5/Indo, H3/ Bri ou B/Flo.

Um SP pode também ser não nativo, por exemplo, de uma proteína estrutural ou hemaglutinina de um vírus que não seja a gripe, ou de um polipéptido vegetal, animal ou bacteriano. Um exemplo não limitativo de um péptido de sinal que pode ser utilizado é o da isomerase de dissulfureto de proteína de alfafa (PDI SP; nucleótidos 32-103 do número de acesso Z11499; SEQ ID NO: 34; Figura 17) possuindo a sequência de aminoácidos:

MAKNVAIFGLLFSLLLLVPSSQIFAE (nucleótidos 32-103; SEQ ID NO: 34)

A presente invenção proporciona, por conseguinte, uma hemaglutinina de gripe quimérica compreendendo um péptido de sinal nativo, ou não nativo, e ácidos nucleicos que codificam tais hemaglutininas quiméricas.

O correto dobramento das hemaglutininas pode ser importante para a estabilidade da proteína, formação de multímeros, formação de VLPs e

função do AH (capacidade de hemaglutinato), entre outras características das hemaglutininas da gripe. O dobramento de uma proteína pode ser influenciado por um ou mais fatores, incluindo, mas não limitado a sequência da proteína, a abundância relativa da proteína, o grau de aglomeração intracelular, a disponibilidade de cofatores que podem se ligar ou ser transitoriamente associados com a proteína dobrada, parcialmente dobrada ou desdobrada, a presença de uma ou mais proteínas acompanhantes ou semelhantes.

As proteínas de choque térmico (Hsp) ou proteínas de stress são exemplos de proteínas acompanhantes, que podem participar em vários processos celulares, incluindo síntese de proteínas, tração intracelular, prevenção de dobragem incorreta, prevenção da agregação de proteínas, montagem e desmontagem de complexos proteicos, dobragem de proteínas e proteína de desagregação. Exemplos de tais proteínas chaperonas incluem, mas não se limitam a, Hsp60, Hsp65, Hsp70, Hsp90, Hsp100, Hsp20-30, Hsp10, Hsp100-200, Hsp100, Hsp90, Lon, TF55, FKBP, ciclofilinas, ClpP, GrpE, ubiquitina, calnexina e isomerases de dissulfureto de protea (ver, por exemplo, Macario, AJL, *Cold Spring Harbor Laboratory Res.* 25: 59-70, 1995; Parsell, DA e Lindquist, S. *Ann. Rev. Genet.* 27: 437-496 (1993); Patente US 5232833). Como aqui descrito, as proteínas acompanhantes, por exemplo, mas não limitadas a Hsp40 e Hsp70 podem ser utilizadas para assegurar a dobragem de um HA quimérico.

Exemplos de Hsp70 incluem Hsp72 e Hsc73 de células de mamífero, ADNK de bactérias, particularmente micobactérias tais como *Mycobacterium leprae*, *Mycobacterium tuberculosis* e *Mycobacterium bovis* (tais como Bacille-Calmette Guerin: aqui referidas como Hsp71). ADNK de *Escherichia coli*, levedura e outros procariontes, e BiP e Grp78 de eucariotas, como *A. thaliana* (Lin et al. 2001 (*Cell Stress and Chaperones* 6: 201-208)). Um exemplo particular de uma Hsp70 é a *A. thaliana* Hsp70 (codificada pelo Genbank ref: AY120747.1). A Hsp70 é capaz de ligar especificamente ATP, bem como polipéptidos e péptidos desdobrados, participando assim

no enrolamento e desdobramento de proteínas, bem como na montagem e desmontagem de complexos proteicos.

Exemplos de Hsp40 incluem ADNJ de procariontes como *E. coli* e micobactérias e HSJ1, HDJ1 e Hsp40 de eucariotas, como a alfafa (Frugis et al., 1999. *Plant Molecular Biology* 40: 397-408). Um exemplo particular de uma Hsp40 é *M. sativa* MsJ1 (AJ000995.1 ou SEQ ID NO: 76). A Hsp40 desempenha um papel como chaperona molecular no dobramento de proteínas, termotolerância e replicação de ADN, entre outras atividades celulares. A Figura 41 mostra a sequência de ácido nucleico que codifica Msj 1 (SEQ ID NO: 76).

Entre as Hsp, a Hsp70 e sua co-chaperona, Hsp40, estão envolvidas na estabilização de polipeptídeos tradutores e recém-sintetizados antes que a síntese esteja completa. Sem querer estar limitado pela teoria, a Hsp40 liga-se às manchas hidrofóbicas de polipéptidos desdobrados (nascentes ou recentemente transferidos), facilitando assim a interação do complexo Hsp70-ATP com o polipéptido. A hidrólise de ATP leva à formação de um complexo estável entre o polipeptídeo, Hsp70 e ADP, e a liberação de Hsp40. A associação do complexo Hsp70-ADP com os fragmentos hidrofóbicos do polipeptídeo impede sua interação com outros fragmentos hidrofóbicos, impedindo o enrolamento incorreto e a formação de agregados com outras proteínas (revisado em Hartl, FU. 1996. *Nature* 381: 571-579).

Proteínas chaperonas nativas podem ser capazes de facilitar o dobramento correto de baixos níveis de proteína recombinante, mas à medida que os níveis de expressão aumentam, a abundância de chaperonas nativas pode se tornar um fator limitante. Altos níveis de expressão de hemaglutinina nas folhas agroinfiltradas podem levar ao acúmulo de polipeptídeos de hemaglutinina no citosol, e a coexpressão de uma ou mais de uma proteína chaperona, como Hsp70, Hsp40 ou ambas Hsp70 e Hsp40, pode reduzir o nível de desdobramento ou polipéptidos de hemaglutinina agregados e aumentar o número de polipéptidos que exibem características estruturais terciárias e quaternárias que permitem a hemaglutinação e/ou formação de partículas semelhantes a vírus. SEQ

ID NO: 77 é uma sequência de ácido nucleico de uma porção do número de construção R850, de HindIII (no local de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS), codificando HSP40 (sublinhado). SEQ ID NO: 78 é uma sequência de ácido nucleico de uma porção do número de construção R860, de HindIII (no local de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS), codificando HSP70 (sublinhado). SEQ ID NO: 79 é uma sequência de ácido nucleico de uma porção do número de construção R870, de HindIII (no local de clonagem múltipla, 5 a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS) que codifica HSP40 (sublinhado *itálico*) e HSP70 (sublinhado). a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS), codificando HSP70 (sublinhado). SEQ ID NO: 79 é uma sequência de ácido nucleico de uma porção do número de construção R870, de HindIII (no local de clonagem múltipla, 5 a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS) que codifica HSP40 (sublinhado *itálico*) e HSP70 (sublinhado). a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS), codificando HSP70 (sublinhado). SEQ ID NO: 79 é uma sequência de ácido nucleico de uma porção do número de construção R870, de HindIII (no local de clonagem múltipla, 5 a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS) que codifica HSP40 (sublinhado *itálico*) e HSP70 (sublinhado).

Por conseguinte, a presente invenção também proporciona um método para produzir VLPs de gripe quiméricas numa planta, em que um primeiro ácido nucleico codificando um HA de gripe quimérico é co-expresso com um segundo ácido nucleico que codifica um acompanhante. O primeiro e segundo ácidos nucleicos podem ser introduzidos na planta no mesmo passo, ou podem ser introduzidos na planta sequencialmente.

As VLPs podem ser avaliadas quanto a estrutura e tamanho, por exemplo, por ensaio de hemaglutinação, microscopia eletrónica ou por cromatografia por exclusão de tamanho.

Para cromatografia por exclusão de tamanho, as proteínas solúveis totais podem ser extraídas do tecido vegetal por homogeneização (Polytron) da amostra de material vegetal triturado congelado em tampão de extração e o material insolúvel removido por centrifugação. Precipitação com PEG também pode ser benéfica. A proteína solúvel é quantificada e o extrato passa através de uma coluna Sephacryl™. O Blue Dextrano 2000 pode ser usado como padrão de calibração. Após cromatografia, as frações podem ser analisadas adicionalmente por imunotransferência para determinar o complemento de proteína da fração.

A presente invenção também proporciona uma planta compreendendo um ácido nucleico que codifica uma ou mais de uma hemaglutinina da gripe quimérica e um ácido nucleico que codifica uma ou mais de uma das chaperonas.

A presente divulgação inclui sequências nucleotídicas:

SEQ ID NO: 63 (construção 690; uma cassete de expressão H5/ H1 quimérica compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina quimérica, 3' UTR e sequências terminadoras de plastocianina de alfafa) e a porção sublinhada de SEQ ID NO: 63 codificando SP, F'1, E1 de H5/Indo-RB de H1/Bri-E2, F'2, F, TMD/CT de H5/Indo;

SEQ ID NO: 64 (construção 691; uma cassete de expressão quimérica H5/H1 compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina quimérica, UTR de plastocianina de alfafa 3' e sequências de terminador) e a porção sublinhada de SEQ ID NO: 64, codificando SP, F'1, de H5/Indo-E1, RB.E2 de H1/Bri-F'2, F, TMD/CT de H5/Indo;

SEQ ID NO: 65 (construção 696; uma cassete de expressão quimérica H1/H5 compreendendo promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina quimérica, sequências 3' UTR e terminadora de plastocianina de alfafa) e a porção sublinhada de

SEQ ID NO: 65 codificando PDI SP-F'1, E1 de H1/NC-RB de H5/Indo-E2, F'2, F, TMD/CT de H1/NC;

SEQ ID NO: 68 (construção 733; a cassette de expressão SpPDI H1/Bri compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência codificadora do péptido de sinal da PDI, sequência codificadora da hemaglutinina da forma H1 A/Brisbane/59/07 (H1N1), CPMV-HT 3'UTR e sequências de terminação NOS), e a porção sublinhada de SEQ ID NO: 68, codificando PDI SP-F'1, E1, RB, E2, F'2, F, TMD/CT de H1/BRI;

SEQ ID NO: 69 (construção 734; uma cassette de expressão quimérica H5/H1 compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência codificadora de hemaglutinina quimérica, CPMV-HT 3' UTR e sequências terminadoras NOS). A sequência de codificação de HA quimérico está sublinhada, codificando o mesmo HA quimérico como SEQ ID NO: 63;

SEQ ID NO: 71 (construção 736; uma cassette de expressão de HA compreendendo o promotor 35M de CaMV, CPMV-HT 5 'UTR, sequência codificadora do péptido de sinal de PDI, sequência codificadora de hemaglutinina de H3 forma A/Brisbane/10/07 (H2N3), as sequências terminadoras UTR e NOS CPMV-HT), e a porção sublinhada de SEQ ID NO: 71 codificando PDI SP-F'1, E1, RB, E2, F'2, F, TMD/CT de H3/Bri;

SEQ ID NO: 72 (construção 737; uma cassette de expressão quimérica H5/H3 compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência de codificação da hemaglutinina quimérica, CPMV-HT 3' UTR e sequências de terminador NOS) e o poro sublinhada da SEQ ID NO: 72 que codifica PDI SP-F'1, E1, RB, E2, F'2, F, TMD/CT de H5/Indo;

SEQ ID NO: 74 (construção 739; uma cassette de expressão HA compreendendo o promotor CaMV 35S, CPMV-HT 5 'UTR, sequência codificadora do péptido de sinal da PDI, sequência codificadora da hemaglutinina da forma HA B/Florida/4/06, CPMV-HT sequências terminadoras 3'UTR e NOS) e a porção sublinhada da SEQ ID NO: 74 codificando PDI SP-F'1, E1, RB, E2, F'2, F, TMD/CT de B/Flo;

SEQ ID NO: 75 (construção 734; uma cassete de expressão quimérica H5/B compreendendo o promotor CaMV 35S, CPMV-HT 5'UTR, sequência de codificação hemaglutinina quimérica, CPMV-HT 3' UTR e sequências terminadoras NOS) e o poro sublinhada de SEQ ID NO: 75 que codifica PDI SP-F'1, E1, RB, E2, F'2, F de B/Flo-TND/CT de H5/Indo.

A presente divulgação também inclui uma sequência nucleotídica que hibrida sob condições de hibridação rigorosas com as pores sublinhadas de qualquer uma das SEQ ID NOs: 63-65, 68, 69 e 71-75. A presente divulgação também inclui uma sequência de nucleótidos que hibrida sob condições de hibridação rigorosas com um complemento das pores sublinhadas de qualquer uma das SEQ ID Nos: 63-65, 68, 69 e 71-75. Estas sequências nucleotídicas que hibridam com as pores sublinhadas de SEQ ID NOs: 63-65, 68, 69 e 71-75, ou um complemento das pores sublinhadas de SEQ ID NOs: 63-65, 68, 69 e 71-75, codificam uma proteína hemaglutinina quimérica que, quando expressa forma uma VLP quimérica, e a VLP quimérica induz a produção de um anticorpo quando administrada a um indivíduo. Por exemplo, a expressão da sequência nucleotídica dentro de uma célula vegetal forma uma VLP quimérica, e a VLP quimérica pode ser utilizada para produzir um anticorpo que capaz de se ligar a HA, incluindo HA maduro, HA0, HA1 ou HA2 de um ou mais tipos ou subtipos de gripe. A VLP quimérica, quando administrada a um sujeito, induz uma resposta imunitária.

A hibridação sob condições de hibridação rigorosas conhecida na técnica (ver por exemplo Current Protocols in Molecular Biology, Ausubel et al., Eds. 1995 e suplementos; Maniatis et al., Em Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982; Sambrook e Russell, em Molecular Cloning: A Laboratory Manual, 3a edição, 2001; cada um dos quais é aqui incorporado por referência). Um exemplo de uma dessas condições de hibridação rigorosas pode ser cerca de 16-20 horas de hibridação em 4 X SSC a 65, seguido por lavagem em 0,1 X SSC a 65 durante uma hora, ou 2 lavagens em 0,1 X SSC a 65? C por 20 ou 30 minutos. Em alternativa, uma condição exemplar de hibridação rigorosa pode ser durante a noite

(16-20 horas) em formamida a 50%, 4X SSC a 42°C, seguida por lavagem em 0,1 X SSC a 65°C durante uma hora, ou 2 lavagens em 0,1 X SSC a 65°C cada tampão fosfato aquoso para 20 ou 30 minutos, ou durante a noite (16-20 horas), ou hibridação in Church (7% de SDS; 0,5 M NaPO₄ tampão pH 7,2; EDTA a 10 mM) a 65° C, com 2 lavagens a 50 em 0,1 X SSC, 0,1% de SDS durante 20 ou 30 minutos cada, ou 2 lavagens a 65 em 2 X SSC, 0,1% de SDS durante 20 ou 30 minutos cada.

Adicionalmente, a presente divulgação inclui sequências nucleotídicas que são caracterizadas como tendo cerca de 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% ou qualquer quantidade entre elas, identidade de sequência, ou semelhança de sequência, com a sequência nucleotídica que codifica HA quimérica de acordo com as porções sublinhadas de qualquer uma de SEQ ID N°: 63 SEQ ID N°: 64, SEQ ID N°: 65, SEQ ID N°: 68, SEQ ID N°: 69, SEQ ID NO: 71, em que a sequência nucleotídica codifica uma proteína de hemaglutinina que quando expressa forma uma VLP quimérica e que a VLP quimérica induz a produção de um anticorpo. Por exemplo, a expressão da sequência nucleotídica numa célula vegetal forma uma VLP quimérica e a VLP quimérica pode ser utilizada para produzir um anticorpo que capaz de se ligar a HA, incluindo HA madura, HA0, HA1 ou HA2. A VLP, quando administrada a um sujeito, induz uma resposta imune.

Uma "resposta imune" geralmente se refere a uma resposta do sistema imune adaptativo. O sistema imune adaptativo geralmente compreende uma resposta humoral e uma resposta mediada por células. A resposta humoral é o aspeto da imunidade que é mediada por anticorpos segregados, produzidos nas células da linhagem de linfócitos B (célula B). Anticorpos segregados se ligam a antígenos na superfície de micróbios invasores (como vírus ou bactérias), que os sinalizam para serem destruídos. A imunidade humoral é geralmente utilizada para referir a produção de anticorpos e os processos que a acompanham, bem como as funções efectoras de anticorpos, incluindo a ativação de células Th2 e produção de citocinas, geração de células de memória, promoção de fagocitose de opsonina, eliminação de patógenos e

semelhantes. Os termos "modular" ou "modulação" ou semelhantes referem-se a um aumento ou diminuição de uma resposta ou parâmetro particular, conforme determinado por qualquer um dos vários ensaios geralmente conhecidos ou utilizados, alguns dos quais são aqui exemplificados.

Uma resposta mediada por células é uma resposta imune que não envolve anticorpos, mas envolve a ativação de macrófagos, células natural killer (NK), linfócitos T citotóxicos específicos de antígeno e a liberação de várias citocinas em resposta a um antígeno. A imunidade mediada por células é geralmente utilizada para referir alguma ativação de células Th, ativação de células Tc e respostas mediadas por células T. A imunidade mediada por células é de particular importância na resposta a infecções virais.

Por exemplo, a indução de linfócitos T CD8 positivos específicos de antígeno pode ser medida utilizando um ensaio ELISPOT; A estimulação de linfócitos T CD4+ positivos pode ser medida utilizando um ensaio de proliferação. Os títulos de anticorpos anti-gripe podem ser quantificados utilizando um ensaio ELISA; Os isotipos de anticorpos específicos de antígeno ou reativos cruzados também podem ser medidos usando anticorpos anti-isotipo (por exemplo, anti-IgG, IgA, IgE ou IgM). Os métodos e técnicas para realizar tais ensaios são bem conhecidos na técnica.

Um ensaio de inibição de hemaglutinação (HI ou HAI) também pode ser utilizado para demonstrar a eficácia de anticorpos induzidos por uma vacina, ou composição de vacina compreendendo HA quimérico ou VLP quimérica pode inibir a aglutinação de glóbulos vermelhos (RBC) por HA recombinante. Os títulos de anticorpos inibidores de hemaglutinação de amostras de soro podem ser avaliados por HAI de microtitulação (Aymard et al. 1973). Os eritrócitos de qualquer uma das várias espécies pode ser usada - por exemplo, cavalo, peru, galinha ou semelhantes. Este ensaio fornece informação indireta sobre a montagem do trímero HA na superfície da VLP, confirmando a apresentação adequada de sítios antigênicos em HAs.

Os títulos de reação cruzada de HAI também podem ser usados para demonstrar a eficácia de uma resposta imune a outras estirpes de vírus relacionadas ao subtipo de vacina. Por exemplo, soro de um sujeito imunizado com uma composição de vacina compreendendo uma hemaglutinina química compreendendo um HDC de um primeiro tipo ou subtipo de gripe pode ser utilizado num ensaio de HAI com uma segunda estirpe de vírus inteiro ou partículas de vírus e o título de HAI determinado.

Sem querer estar limitado pela teoria, a capacidade de HA de se ligar a RBC de diferentes animais é impulsionada pela afinidade de HA por ácidos siálicos ligados a ligações $\alpha 2,3$ ou $\alpha 2,6$ e a presença destes ácidos siálicos na superfície de RBC. EH de equinos e aves dos vírus da gripe aglutinam os eritrócitos de todas as várias espécies, incluindo perus, galinhas, patos, porquinhos-da-índia, humanos, ovelhas, cavalos e vacas; considerando que os AH humanos se ligam aos eritrócitos de perus, galinhas, patos, cobaias, seres humanos e ovelhas (Ito T. et al, 1997, Virology, 227: 493-499; Medeiros R et al, 2001. Virology 289: 74-85).

A presença ou níveis de citocinas também podem ser quantificados. Por exemplo, uma resposta de células T auxiliares (Th1/Th2) será caracterizada pela medição de células secretoras de IFN- γ e IL-4 utilizando por ELISA (por exemplo, kits BD Biosciences OptEIA). Podem ser cultivadas células mononucleares do sangue periférico (PBMC) ou esplenócitos obtidos a partir de um sujeito, e o sobrenadante analisado. Os linfócitos T podem também ser quantificados por separação celular ativada por fluorescência (FACS), utilizando marcadores fluorescentes específicos para marcadores e métodos como são conhecidos na técnica.

Um ensaio de microneutralização pode também ser conduzido para caracterizar uma resposta imunitária num indivíduo, ver por exemplo os métodos de Rowe et al., 1973. Os títulos de neutralização de vírus podem ser obtidos de várias maneiras, incluindo: 1) enumeração de placas de lise (ensaio de placa) após fixação de cristal violeta / coloração de células; 2) observação microscópica da lise celular em

cultura; 3) ELISA e detecção espectrofotométrica da proteína do vírus NP (correlacionada com a infecção viral das células do hospedeiro)

A identidade de sequência ou similaridade de sequência pode ser determinada utilizando um programa de comparação de sequências, tal como o fornecido no ADNSIS (por exemplo, utilizando, mas não limitado aos seguintes parâmetros: penalidade GAP 5, #de diagonais superiores 5, penalidade fixa GAP 10, k-tuple 2, gap flutuante 10 e tamanho da janela 5). No entanto, outros métodos de alinhamento de sequências para comparação são bem conhecidos na arte, por exemplo, os algoritmos de Smith & Waterman (1981, Adv. Appl. Math. 2: 482) Needleman & Wunsch (J. Mol. Biol. 48: 443, 1970) Pearson e Lipman (1988, Proc. Natl. Acad. Sci. USA 85: 2444), e por implementações computadorizadas desses algoritmos (por exemplo, GAP, BESTFIT, FASTA e BLAST (Altschul et al., 1990. J. Mol Biol 215: 403-410), ou por alinhamento manual e inspeção visual. As sequências de ácidos nucleicos ou de aminoácidos podem ser comparadas ou alinhadas e as sequências de consenso podem ser determinadas utilizando qualquer um dos vários pacotes de software conhecidos na técnica, por exemplo MULTALIN (Corpet F., 1988, Nucl. Acids Res., 16 (22), 10881-10890), BLAST, CLUSTAL ou similares; em alternativa as sequências podem ser alinhadas manualmente e semelhanças e diferenças entre as sequências determinadas.

Um fragmento ou porção de uma proteína, proteína de fusão ou polipéptido inclui um péptido ou polipéptido compreendendo um subconjunto do complemento de aminoácidos de uma proteína ou polipéptido particular, desde que o fragmento possa formar uma VLP quimérica quando expressa. O fragmento pode, por exemplo, compreender uma região antigénica, uma região indutora da resposta ao stress ou uma região compreendendo um domínio funcional da proteína ou polipéptido. O fragmento pode também compreender uma região ou domínio comum a proteínas da mesma família geral, ou o fragmento pode incluir sequência de aminoácidos suficiente para identificar especificamente a proteína de comprimento total a partir da qual é derivada.

Por exemplo, um fragmento ou porção pode compreender desde cerca de 60% até cerca de 100%, do comprimento do comprimento total da proteína, ou qualquer quantidade entre eles, desde que o fragmento possa formar uma VLP quimérica quando expressa. Por exemplo, de cerca de 60% a cerca de 100%, de cerca de 70% a cerca de 100%, de cerca de 80% a cerca de 100%, de cerca de 90% a cerca de 100%, de cerca de 95% a cerca de 100% comprimento do comprimento total da proteína, ou qualquer quantidade entre elas. Alternativamente, um fragmento ou porção pode ser de cerca de 150 a cerca de 500 aminoácidos, ou qualquer quantidade entre eles, dependendo do HA quimérico, e desde que o fragmento possa formar uma VLP quimérica quando expressa. Por exemplo, um fragmento pode ser de 150 a cerca de 500 aminoácidos, ou qualquer quantidade entre eles, de cerca de 200 a cerca de 500 aminoácidos, ou qualquer quantidade entre eles, de cerca de 250 a cerca de 500 aminoácidos, ou qualquer quantidade entre, de cerca de 300 a cerca de 500 aminoácidos, ou qualquer quantidade entre eles, de cerca de 350 a cerca de 500 aminoácidos, ou qualquer quantidade entre eles, de cerca de 400 a cerca de 500 ou qualquer quantidade entre eles, de cerca de 450 a cerca de 500 ou qualquer quantidade entre elas, dependendo do HA quimérico, e desde que o fragmento possa formar uma VLP quimérica quando expressa. Por exemplo, cerca de 5, 10, 20, 30, 40 ou 50 aminoácidos, ou qualquer quantidade entre os mesmos, podem ser removidos do terminal C, do terminal N ou de ambos os terminais N e C de uma proteína HA quimérica, desde que o fragmento pode formar uma VLP quimérica quando expressa. de cerca de 400 a cerca de 500 ou qualquer quantidade entre elas, de cerca de 450 a cerca de 500 ou qualquer quantidade entre elas, dependendo do HA quimérico, e desde que o fragmento possa formar uma VLP quimérica quando expresso. Por exemplo, cerca de 5, 10, 20, 30, 40 ou 50 aminoácidos, ou qualquer quantidade entre os mesmos, podem ser removidos do terminal C, do terminal N ou de ambos os terminais N e C de uma proteína HA quimérica, desde que o fragmento pode formar uma VLP quimérica quando expressa. de cerca de 400 a cerca de 500 ou qualquer quantidade entre elas, de cerca de 450 a cerca de 500 ou qualquer quantidade entre

elas, dependendo do HA quimérico, e desde que o fragmento possa formar uma VLP quimérica quando expresso. Por exemplo, cerca de 5, 10, 20, 30, 40 ou 50 aminoácidos, ou qualquer quantidade entre os mesmos, podem ser removidos do terminal C, do terminal N ou de ambos os terminais N e C de uma proteína HA quimérica, desde que o fragmento pode formar uma VLP quimérica quando expressa.

A numeração de aminoácidos em qualquer sequência dada é relativa à sequência particular, no entanto, um especialista pode facilmente determinar a equivalência de um aminoácido particular numa sequência baseada na estrutura e/ou sequência. Por exemplo, se 6 aminoácidos do terminal N fossem removidos, isto mudaria a identidade numérica específica do aminoácido (por exemplo, em relação ao comprimento total da proteína), mas não alteraria a posição relativa do aminoácido na estrutura.

A presente divulgação descreve, mas não está limitada a expressão de um ácido nucleico que codifica um HA quimérico num poro de planta de uma planta, ou uma célula vegetal, e a produção de VLPs de gripe quiméricas da planta, adequadas para produção de vacinas. Exemplos de tais ácidos nucleicos incluem, por exemplo, mas não estão limitados a SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69.

A presente divulgação proporciona ainda a expressão de um ácido nucleico que codifica um HA quimérico, por exemplo mas não limitado a SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69 , em uma planta, uma porção de uma planta, ou uma célula vegetal, e produção de candidatos ou reagentes da vacina contra gripe composta de proteínas estruturais recombinantes da gripe que se auto-agrupam em estruturas proteicas macromoleculares homotípicas funcionais e imunogénicas, incluindo partículas de gripe subvirais e gripe quimérica VLP, em células vegetais transformadas.

Por conseguinte, a divulgação proporciona VLPs quiméricas e um modo para a produção de VLPs quiméricas num sistema de expressão de plantas, a partir da expressão de uma única proteína de envelope quimérica.

O ácido nucleico que codifica o HA quimérico de subtipos de gripe, por exemplo, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69, pode ser sintetizado por transcrição reversa e reação em cadeia da polimerase (PCR) usando ARN HA. Como exemplo, o ARN pode ser isolado de H1/ NC, H1/Bri ou H5/Indo, ou de células infectadas com estes ou outros tipos ou subtipos de vírus gripe. Para a transcrição reversa e PCR, podem ser utilizados iniciadores oligonucleotídicos específicos para o ARN HA. Adicionalmente, um ácido nucleico codificando uma HA quimérica pode ser sintetizado quimicamente utilizando métodos como seria conhecido de um especialista na técnica.

A presente invenção é ainda dirigida a uma construção genética compreendendo um ácido nucleico que codifica uma HA quimérica, como descrito acima, operativamente ligado a um elemento regulador que é operativo numa planta. Exemplos de elementos reguladores operacionais numa célula vegetal e que podem ser utilizados de acordo com a presente invenção incluem, mas não estão limitados a uma região reguladora de plastocianina (US 7125978), ou uma região reguladora da Ribulose 1,5-bisfosfato carboxilase/oxigenase (RuBisCO; US 4962028), proteína de ligação à clorofila a/b (CAB; Leutwiler et al., 1986), ST-LS1 (associada ao complexo do fotossistema II que envolve o oxigénio e descrita por Stockhaus et al. 1987, 1989).

A construção genética da presente invenção pode também compreender um promotor constitutivo que dirige a expressão de um gene que está operativamente ligado ao promotor ao longo das várias partes de uma planta e continuamente ao longo do desenvolvimento da planta. Um exemplo não limitativo de um promotor constitutivo é aquele associado ao transcrito CaMV 35S (Odell et al., 1985, Nature, 313: 810-81.).

Um exemplo de uma sequência compreendendo uma região reguladora da plastocianina é a sequência 5' para a sequenciada sublinhada que codifica um péptido de sinal da PDI da SEQ ID NO: 58. Um elemento regulador ou região reguladora pode melhorar a tradução de uma sequência nucleotídica à qual está ligada de forma operacional, onde

a sequência nucleotídica pode codificar uma proteína ou polipéptido. Outro exemplo de uma região reguladora, o derivado das regiões não traduzidas do vírus do mosaico do caupi (CPMV), que pode ser utilizado para traduzir preferencialmente a sequência de nucleótidos a que está ligada operativamente. Esta região reguladora do CPMV é explorada em um sistema CPMV hiper-traduzível (CPMV-HT; ver, por exemplo, Sainsbury et al, 2008, Plant Physiology 148: 1212-1218; Sainsbury et al., 2008 Plant Biotechnology Journal 6: 82-92).

Por conseguinte, um aspeto da invenção proporciona um ácido nucleico compreendendo uma região reguladora operativamente ligada a uma sequência que codifica uma HA quimérica de gripe. A região reguladora pode ser um elemento regulador de plastocianina, e a gripe HA quimérico pode compreender subdomínios tipos de gripe de H5/Indo, H1/Bri, H1/NC, subtipos ou estirpes. As sequências de ácido nucleico compreendendo um elemento regulador de plastocianina e um HA de gripe quimérico são aqui exemplificadas pelas SEQ ID NOs: 63 e 64. As sequências de ácido nucleico compreendendo um elemento regulador 35S e um HA de gripe quimérico s aqui exemplificadas pelas SEQ ID NOs: 68, 69 e 71-75.

Noutro aspeto, a invenção proporciona um ácido nucleico compreendendo uma região reguladora de CPMV e uma HA de gripe quimérica, compreendendo subdomínios dos tipos, subtipos ou estirpes de gripe H5/Indo, H1/Bri, H1/NC. As sequências de ácido nucleico compreendendo um elemento regulador de CPMP e um HA quimérico s aqui exemplificadas pelas SEQ ID NOs: 66-69 e 71-75.

VLPs de gripe quiméricas produzidas em plantas brotam a partir da membrana plasmática e a composição lipídica das VLPs quiméricas reflete aquela do tipo de tecido celular ou planta da planta do qual elas são produzidas. As VLPs produzidas de acordo com a presente invenção compreendem HA quimérico de dois tipos ou subtipos de gripe, complexados com lípidos derivados de plantas. Os lípidos vegetais podem estimular células imunitárias específicas e aumentar a resposta imunitária induzida.

Lípidos de plantas tais como PC (fosfatidilcolina) e PE (fosfatidil etanolamina), bem como glicosfingolípídeos podem ligar-se a moléculas CD1 expressas por células imunes de mamíferos tais como células apresentadoras de antígenos (APCs) como células dendríticas e macrófagos e outras células incluindo B e Linfócitos T no timo e fígado (revisto em Tsuji M,. 2006 Cell Mol Life Sci 63: 1889-98). As moléculas de CD1 são estruturalmente semelhantes às moléculas do complexo de histocompatibilidade principal (MHC) da classe I e seu papel é apresentar antígenos glicolipídicos às células NKT (células T Natural Killer). Após a ativação, as células NKT ativam as células imunes inatas, como as células NK e as células dendríticas, e também ativam as células imunes adaptativas, como as células B e as células T produtoras de anticorpos.

Os fitosteróis presentes numa VLP de gripe complexada com uma bicamada lipídica, tal como um envelope derivado da membrana plasmática, podem proporcionar uma composição de vacina vantajosa. Sem querer estar limitado pela teoria, as VLPs feitas de plantas, incluindo aquelas compreendendo HA quimérico, complexadas com uma bicamada lipídica, tal como um envelope derivado da membrana plasmática, podem induzir uma reação imune mais forte do que as VLPs feitas em outros sistemas de expressão. ser semelhante à reação imune induzida por vacinas de vírus inteiros vivos ou atenuados.

Por conseguinte, em algumas formas de realização, a invenção proporciona uma VLP compreendendo uma HA quimérica, complexada com uma bicamada lipídica derivada de plantas. Em algumas formas de realização, a bicamada lipídica derivada de plantas pode compreender o envelope da VLP.

A VLP produzida dentro de uma planta pode incluir um HA quimérico compreendendo N-glicanos específicos de plantas. Portanto, esta invenção também proporciona uma VLP compreendendo uma HA quimérica tendo N-glicanos específicos da planta.

Além disso, a modificação do N-glicano nas plantas é conhecida (ver, por exemplo, WO 2008/151 440) e HA quimérico tendo N-glicanos

modificados pode ser produzido. Pode ser obtido um HA quimérico compreendendo um padrão de glicosilação modificado, por exemplo com N-glicanos fucosilados, xilosilados ou ambos fucosilados e xilosilados, ou pode ser obtido HA quimérico possuindo um padrão de glicosilação modificado, em que a proteína não tem fucosilação, xilosilação ou ambos e compreende galatosilação aumentada. Além disso, a modulação de modificações pós-tradução, por exemplo, a adição de galactose terminal pode resultar numa redução da fucosilação e xilosilação do HA quimérico expresso quando comparado com uma planta de tipo selvagem que expressa HA quimérico.

Por exemplo, o que não deve ser considerado limitante, a síntese de HA quimérico tendo um padrão de glicosilação modificado pode ser conseguido co-expressando a proteína de interesse juntamente com uma sequência nucleotídica codificando beta-1,4galactosiltransferase (GalT), por exemplo, mas não se limitando a GalT de mamífero, ou GalT humano, no entanto, também pode ser utilizado GalT de outras fontes. O domínio catalítico de GalT também pode ser fundido a um domínio CTS (ou seja, cauda citoplasmática, domínio transmembrana, região tronco) da N-acetilglucosaminil transferase (GNT1), para produzir uma enzima híbrida GNT1-GalT, e a enzima híbrida pode ser co-expressa com HA. O HA pode também ser co-expresso juntamente com uma sequência nucleotídica codificando N-acetilglucosaminiltransferase III (GnT-III), por exemplo mas não limitada a GnT-III de mamífero ou GnT-III humana, A GnT-III de outras fontes também pode ser usada. Adicionalmente, também pode ser utilizada uma enzima híbrida GNT1-GnT-III, compreendendo o CTS de GNT1 fundido com GnT-III.

Portanto, a presente invenção também inclui VLP compreendendo HA quimérico tendo N-glicanos modificados.

Sem querer estar limitado pela teoria, a presença de N-glicanos vegetais num HA quimérico pode estimular a resposta imunitária promovendo a ligação de HA por células apresentadoras de antigénio. A estimulação da resposta imune utilizando a planta N glicano tem sido proposta por Saint-Jore-Dupas et al. (Trends Biotechnol 25: 317-23,

2007). Além disso, a conformação da VLP pode ser vantajosa para a apresentação do antígeno e aumentar o efeito adjuvante da VLP quando complexada com uma camada lipídica derivada da planta.

Por "região reguladora", "elemento regulador" ou "promotor" entende-se uma porção de ácido nucleico tipicamente, mas nem sempre, a montante da região codificadora de proteína de um gene, que pode ser constituído por ADN ou ARN, ou ambos ADN e ARN. Quando uma região reguladora está ativa, e em associação operativa, ou operativamente ligada, com um gene de interesse, isto pode resultar na expressão do gene de interesse. Um elemento regulador pode ser capaz de mediar a especificidade de órgãos ou controlar a ativação de genes temporais ou de desenvolvimento. Uma "região reguladora" inclui elementos promotores, elementos promotores centrais exibindo uma atividade promotora basal, elementos que são indutíveis em resposta a um estímulo externo, elementos que medeiam a atividade promotora, tais como elementos reguladores negativos ou estimuladores da transcrição. "Região reguladora", como aqui utilizada, também inclui elementos que são ativos após transcrição, por exemplo, elementos reguladores que modulam a expressão génica, tais como intensificadores de tradução e transcrição, repressores de tradução e transcrição, sequências de ativação a montante e determinantes da instabilidade de ARNm. Vários destes últimos elementos podem estar localizados próximo da região de codificação.

No contexto desta divulgação, o termo "elemento regulador" ou "região reguladora" refere-se tipicamente a uma sequência de ADN, normalmente, mas nem sempre, a montante (5') da sequência de codificação de um gene estrutural, que controla a expressão da região codificadora, proporcionando o reconhecimento da polimerase de ARN e/ou outros fatores necessários para a transcrição iniciar num local particular. No entanto, é para ser entendido que outras sequências nucleotídicas, localizadas dentro dos intrões, ou 3'da sequência podem também contribuir para a regulação da expressão de uma região codificadora de interesse. Um exemplo de um elemento regulador que

proporciona o reconhecimento da ARN polimerase ou de outros fatores transcricionais para assegurar a iniciação em um local específico é um elemento promotor. A maioria, mas não todas Os elementos promotores eucarióticos contêm uma caixa TATA, uma sequência de ácidos nucleicos conservada composta por pares de bases de nucleótidos de adenosina e timidina situados habitualmente aproximadamente 25 pares de bases a montante de um local de início da transcrição. Um elemento promotor compreende um elemento promotor basal, responsável pela iniciação da transcrição, bem como outros elementos reguladores (como listado acima) que modificam a expressão gênica.

Existem vários tipos de regiões reguladoras, incluindo aquelas que são reguladas pelo desenvolvimento, indutíveis ou constitutivas. Uma região reguladora que é regulada pelo desenvolvimento, ou controla a expressão diferencial de um gene sob seu controle, é ativada dentro de certos órgãos ou tecidos de um órgão em momentos específicos durante o desenvolvimento desse órgão ou tecido. No entanto, algumas regiões reguladoras que são reguladas pelo desenvolvimento podem ser preferencialmente ativas dentro de certos órgãos ou tecidos em estágios específicos de desenvolvimento, elas também podem ser ativas de maneira regulada pelo desenvolvimento, ou em nível basal em outros órgãos ou tecidos dentro da planta também. Exemplos de regiões reguladoras específicas do tecido, por exemplo, uma região reguladora específica da vista, incluem o promotor da napina e o promotor da cruciferina (Rask et al.f 1998, J. Plant Physiol. 152: 595-599; Bilodeau et al., 1994, Plant Cell 14: 125-130). Um exemplo de um promotor específico de folhas inclui o promotor de plastocianina (ver, por exemplo, a SEQ ID NO: 58); US 7125978, que é aqui incorporado por referência).

Uma região reguladora indutível é aquela que é capaz de direta ou indiretamente ativar a transcrição de uma ou mais sequências de ADN ou genes em resposta a um indutor. Na ausência de um indutor, as sequências de ADN ou genes não serão transcritos. Tipicamente, o fator de proteína que se liga especificamente a uma região reguladora

indutível para ativar a transcrição pode estar presente em uma forma inativa, que é então direta ou indiretamente convertida para a forma ativa pelo indutor. No entanto, o fator de proteína também pode estar ausente. O indutor pode ser um agente químico como uma proteína, metabólito, regulador de crescimento, herbicida ou composto fenólico ou um estresse fisiológico imposto diretamente por calor, frio, sal ou elementos tóxicos ou indiretamente pela ação de um patógeno ou agente causador de doenças, como um vírus. Uma célula de planta contendo uma região reguladora indutível pode ser exposta a um indutor aplicando externamente o indutor à célula ou planta tal como por pulverização, rega, aquecimento ou métodos semelhantes. Elementos reguladores indutíveis podem ser derivados de genes vegetais ou não-vegetais (por Gatz, C. e Lenk, IRP, 1998, Trends Plant Sci. 3, 352-358; que é incorporado por referência). Exemplos de potenciais promotores indutíveis incluem, mas não limitados a promotor indutível por tetraciclina (Gatz, C., 1997, Ann. Rev. Plant Physiol. Plant Mol. Biol. 48, 89-108; que é incorporado por referência), promotor induzido por esteroides (Aoyama, T. e Chua, NH, 1997, Plant J. 2, 397-404; que é incorporado por referência) e promotor indutível por etanol (Salter, MG, et al., 1998, Plant Journal 16, 127-132; Caddick, MX, et al., 1998, Nature Biotech. 16, 177-180, que são incorporados por referência) genes IB6 e CKI1 indutíveis pela citoquinina (Brandstatter, I. e Kieber, JJ, 1998, Plant Cell 10, 1009-1019; Kakimoto, T., 1996, Science 274, 982-985; incorporados por referência) e o elemento indutível de auxina, DR5 (Ulmasov, T., et al., 1997, Plant Cell 9, 1963-1971; que é incorporado por referência).

Uma região reguladora constitutiva dirige a expressão de um gene ao longo das várias partes de uma planta e continuamente durante o desenvolvimento da planta. Exemplos de elementos reguladores constitutivos conhecidos incluem promotores associados com o transcrito CaMV 35S. (Odell et al., 1985, Nature, 313: 810-812), a actina 1 do arroz (Zhang et al, 1991, Plant Cell, 3: 1155-1165), actina 2 (An et al., 1996, Plant J., 10: 107-121.) ou tms 2 (US 5428147, que é aqui incorporado por referência), e genes do

triosefosfato isomerase 1 (Xu et. al., 1994, Plant Physiol. 106: 459-467), o gene da ubiquitina 1 do milho (Cornejo et al, 1993, Plant Mol. Biol. 29: 637-646), os genes da *Arabidopsis* ubiquitina 1 e 6 (Holtorf et al, 1995, Plant Mol. Biol. 29: 637-646) e o gene do fator de iniciação translacional do tabaco 4A (Mandel et al, 1995 Plant Mol. Biol. 29: 995-1004). O termo "constitutivo", como aqui utilizado, não indica necessariamente que um gene sob o controle da região reguladora constitutiva é expresso ao mesmo nível em todos os tipos de células, mas que o gene é expresso numa vasta gama de tipos de células, embora a variação abundância é frequentemente observada. Os elementos reguladores constitutivos podem ser acoplados a outras sequências para aumentar ainda mais a transcrição e/ou tradução da sequência nucleotídica à qual estão operativamente ligados. Por exemplo, o sistema CPMV-HT é derivado das regiões não traduzidas do vírus do mosaico de Cowpea (CPMV) e demonstra uma tradução melhorada da sequência de codificação associada.

Por "nativo" significa-se que o ácido nucleico ou sequência de aminoácidos ocorre naturalmente ou "tipo selvagem".

Por "operacionalmente ligado" entende-se que as sequências particulares, por exemplo, um elemento regulador e uma região codificadora de interesse, interagem direta ou indiretamente para realizar uma função pretendida, tal como mediação ou modulação da expressão gênica. A interação de sequências operativamente ligadas pode, por exemplo, ser mediada por proteínas que interagem com as sequências ligadas operativamente.

A sequência nucleotídica da presente invenção pode ser expressa em qualquer hospedeiro vegetal adequado que seja transformado pela sequência nucleotídica ou construções ou vetores da presente invenção. Exemplos de hospedeiros adequados incluem, mas não estão limitados a culturas agrícolas incluindo alfafa, canola, *Brassica* spp., Milho, *Nicotiana* spp., Alfafa, batata, ginseng, ervilha, aveia, arroz, soja, trigo, cevada, girassol, algodão. e similar.

A construção genética quimérica da presente invenção pode ainda compreender uma região 3' não traduzida. Uma região 3' não traduzida refere-se à porção de um gene compreendendo um segmento de ADN que contém um sinal de poliadenilação e quaisquer outros sinais reguladores capazes de efectuar o processamento de ARNm ou a expressão génica. O sinal de poliadenilação é normalmente caracterizado por efetuar a adição de pistas de ácido poliadenílico à extremidade 3' do precursor de ARNm. Sinais de poliadenilação são comumente reconhecidos pela presença de homologia com a forma canônica 5' AATAAA-3', embora variações não sejam incomuns.

Exemplos não limitantes de regiões 3' adequadas são as regiões não traduzidas transcritas em 3' contendo um sinal de poliadenilação de genes de plasmídeos indutores de tumor (Ti) de *Agrobacterium*, tais como o gene da nopalina sintase (NOS), genes de plantas tais como a proteína de armazenamento de genes de soja, a pequena subunidade do gene da ribulose-1, 5-bisfosfato carboxilase (ssRUBISCO; US 4962028; que é aqui incorporado por referência), o promotor usado na regulação da expressão de plastocianina, descrito em US 7125978 (que é aqui incorporado por referência).

Uma ou mais das construções genéticas quiméricas da presente invenção também podem incluir outros intensificadores, sejam tradutores ou intensificadores de transcrição, conforme possa ser requerido. Os realçadores podem estar localizados 5' ou 3' para a sequência que está sendo transcrita. As regiões potenciadoras são bem conhecidas dos peritos na arte e podem incluir um codão de iniciação ATG, sequências adjacentes ou semelhantes. O codão de iniciação, se presente, pode estar em fase com o quadro de leitura ("no enquadramento") da sequência de codificação para proporcionar a tradução correta da sequência transcrita.

Para auxiliar na identificação de células vegetais transformadas, as construções desta invenção podem ser ainda manipuladas para incluir marcadores selecionáveis de plantas. Marcadores selecionáveis úteis incluem enzimas que proporcionam resistência a químicos tais como um

antibiótico, por exemplo, gentamicina, higromicina, canamicina ou herbicidas tais como fosfinotricina, glifosato, clorossulfurão e semelhantes. De modo semelhante, podem ser utilizadas enzimas que proporcionam a produção de um composto identificável por alteração de cor, tal como GUS (beta-glucuronidase), ou luminescência, tal como luciferase ou GFP.

Também são considerados parte desta invenção plantas transgênicas, células vegetais ou sementes contendo a construção de genes quiméricos da presente invenção. Métodos de regeneração de plantas inteiras a partir de células vegetais são também conhecidos na técnica. Em geral, as células vegetais transformadas são cultivadas num meio apropriado, o qual pode conter agentes seletivos tais como antibióticos, onde marcadores selecionáveis são utilizados para facilitar a identificação de células vegetais transformadas. Uma vez formadas pelos calos, a formação de rebentos pode ser encorajada empregando as hormonas vegetais apropriadas de acordo com métodos conhecidos e os rebentos transferidos para o meio de enraizamento para regeneração de plantas. As plantas podem então ser usadas para estabelecer gerações repetitivas, a partir de sementes ou usando técnicas de propagação vegetativa. Plantas transgênicas também podem ser geradas sem o uso de culturas de tecidos.

Também consideradas parte desta invenção são plantas transgênicas e árvores contendo a construção gênica quimérica compreendendo um ácido nucleico codificando HA ou HA0 quimérico recombinante para produção de VLP, de acordo com a presente invenção.

Os elementos reguladores da presente divulgação também podem ser combinados com a região codificante de interesse para expressão dentro de uma gama de organismos hospedeiros que são susceptíveis a transformação ou expressão transiente. Tais organismos incluem, mas não estão limitados a plantas, monocotiledóneas e dicotiledóneas, por exemplo mas não limitadas a milho, plantas de cereais, trigo, cevada, aveia, *Nicotiana* spp, *Brassica* spp, soja, feijão, ervilha, alfafa, batata, tomate, ginseng e *Arabidopsis*.

Métodos para transformação estável e regeneração destes organismos são estabelecidos na arte e conhecidos de um especialista na técnica. O método de obtenção de plantas transformadas e regeneradas não é crítico para a presente invenção.

Por "transformação" entende-se a transferência interespecífica de informação genética (sequência nucleotídica) que se manifesta genotipicamente, fenotipicamente ou ambas. A transferência interespecífica de informação genética de uma construção quimérica para um hospedeiro pode ser hereditária e a transferência de informação genética considerada estável, ou a transferência pode ser transitória e a transferência de informação genética não é hereditária.

Pelo termo "matéria vegetal", entende-se qualquer material derivado de uma planta. A matéria vegetal pode compreender uma planta inteira, tecido, células ou qualquer fração dela. Além disso, a matéria vegetal pode compreender componentes de plantas intracelulares, componentes de plantas extracelulares, extratos líquidos ou sólidos de plantas ou uma combinação destes. Além disso, a matéria vegetal pode compreender plantas, células vegetais, tecido, um extrato líquido ou uma sua combinação, a partir de folhas de plantas, caules, frutos, raízes ou uma sua combinação. A matéria vegetal pode compreender uma planta ou parte dela que não tenha sido submetida a quaisquer etapas de processamento. Uma parte de uma planta pode compreender matéria vegetal. No entanto, também é contemplado que o material vegetal pode ser submetido a etapas de processamento mínimas, conforme definido abaixo, ou a um processamento mais rigoroso, incluindo purificação de proteína parcial ou substancial utilizando técnicas vulgarmente conhecidas na técnica incluindo, mas não se limitando a cromatografia, eletroforese e semelhantes.

Pelo termo "processamento mínimo" entende-se a matéria vegetal, por exemplo, uma planta ou sua porção compreendendo uma proteína de interesse que é parcialmente purificada para produzir um extrato vegetal, homogeneizado, fração de homogenato de planta ou semelhantes

(isto é, minimamente processados). A purificação parcial pode compreender, mas não está limitada a rutura de estruturas celulares vegetais, criando assim uma composição compreendendo componentes vegetais solúveis e componentes vegetais insolúveis que podem ser separados, por exemplo, mas não limitados a centrifugação, filtração ou uma sua combinação. A este respeito, as proteínas segregadas no espaço extracelular da folha ou outros tecidos podem ser prontamente obtidas usando vácuo ou extração centrífuga, ou os tecidos podem ser extraídos sob pressão por passagem através de rolos ou trituração ou semelhantes para espremer ou libertar a proteína livre de dentro do espaço extracelular. O processamento mínimo também poderia envolver a preparação de extratos brutos de proteínas solúveis, uma vez que essas preparações teriam contaminação desprezível de produtos vegetais secundários. Além disso, o processamento mínimo pode envolver a extração aquosa da proteína solúvel das folhas, seguida de precipitação com qualquer sal adequado. Outros métodos podem incluir maceração em grande escala e extração de suco, a fim de permitir o uso direto do extrato.

A matéria vegetal, na forma de material vegetal ou tecido, pode ser administrada oralmente a um sujeito. A matéria vegetal pode ser administrada como parte de um suplemento dietético, junto com outros alimentos, ou encapsulada. A matéria vegetal ou tecido também pode ser concentrado para melhorar ou aumentar a palatabilidade, ou fornecido juntamente com outros materiais, ingredientes ou excipientes farmacêuticos, conforme necessário.

Exemplos de um sujeito ou organismo alvo que as VLPs da presente invenção podem ser administradas para incluir, mas não estão limitados a humanos, primatas, aves, aves aquáticas, aves migratórias, codornas, patos, gansos, aves domésticas, galinhas, suínos, ovelhas, equino, cavalo, camelo, canino, cães, felino, gatos, tigre, leopardo, civeta, marta, marta de pedra, furões, casa de animais de estimação, gado, coelhos, ratos, cobaias ou outros roedores, selo, baleia e o gosto. Tais organismos alvo são exemplares e não devem ser

considerados limitantes às aplicações e utilizações da presente invenção.

É contemplado que uma planta compreendendo a HA quimérica de acordo com algumas formas de realização da invenção, ou expressando a VLP compreendendo a HA quimérica de acordo com algumas formas de realização da invenção, pode ser administrada a um sujeito ou organismo alvo, numa variedade de maneiras dependendo sobre a necessidade e a situação. Por exemplo, o HA quimérico obtido a partir da planta pode ser extraído antes da sua utilização quer numa forma bruta, parcialmente purificada, quer purificada. Se o HA quimérico deve ser pelo menos parcialmente purificado, então pode ser produzido em plantas comestíveis ou não comestíveis. Além disso, se o HA quimérico for administrado oralmente, o tecido da planta pode ser colhido e ser alimentado diretamente ao indivíduo, ou o tecido colhido pode ser seco antes da alimentação, ou um animal pode ser permitido pastar na planta sem colheita prévia. Lugar, colocar. É também considerado dentro do âmbito desta invenção que os tecidos vegetais colhidos sejam fornecidos como um suplemento alimentar na alimentação animal. Se o tecido vegetal estiver a ser alimentado a um animal com pouco ou nenhum processamento adicional, é preferível que o tecido vegetal a ser administrado seja comestível.

O silenciamento génico pós-transcricional (PTGS) pode estar envolvido na limitação da expressão de transgenes em plantas, e a coexpressão de um supressor de silenciamento do vírus da batata Y (HcPro) pode ser usada para neutralizar a degradação específica de mARNs transgénicos (Brigneti et al., 1998). Supressores alternativos de silenciamento são bem conhecidos na técnica e podem ser usados como descrito aqui (Chiba et al., 2006, Virology 346: 7-14; que é aqui incorporado por referência), por exemplo mas não se limitando a, TEV-pl / HC-Pro (Vírus etch de tabaco-pl/HC-Pro), BYV-p21, p19 de Tomato stab virus (TBSV p19), cápside proteína do Tomato plinkle virus (TCV-PC), 2b do vírus do mosaico do pepino; CMV-2b), p25 do Potato virus X (PVX-p25), p11 do Potato virus M (PVM-p11), p11 do Potato virus S

(PVS-p11), p16 do Blueberry scorch virus, (BScV-p16), p23 do vírus Citrus tristeza (CTV-p23), p24 do vírus associado a Leafroll Grapevine-2, (GLRaV-2 p24), p10 do vírus Grapevine A, (GVA-p10), p14 do vírus Grapevine B (GVB-p14), p10 do vírus latente Heracleum (HLV-p10), ou p16 do vírus latente comum Garlic (GCLV-p16). Portanto, um supressor de silenciamento, por exemplo, mas não limitado a HcPro, TEV-p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS- p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 ou GVA-p10, podem ser co-expressos juntamente com a sequência de ácido nucleico que codifica a proteína de interesse para assegurar ainda mais elevados níveis de produção de proteína dentro de uma planta.

Além disso, as VLP produzidas como aqui descrito não compreendem neuraminidase (NA). No entanto, a NA pode ser co-expressa com HA se as VLPs compreendendo HA e NA forem desejadas.

Por conseguinte, a presente divulgação inclui ainda um vetor adequado compreendendo a sequência de HA quimérica adequada para utilização com sistemas de expressão estável ou transitória. A informação genética pode também ser fornecida dentro de um ou mais de uma construção. Por exemplo, uma sequência nucleotídica que codifica uma proteína de interesse pode ser introduzida numa construção e uma segunda sequência nucleotídica que codifica uma proteína que modifica a glicosilação da proteína de interesse pode ser introduzida utilizando uma construção separada. Estas sequências nucleotídicas podem então ser co-expressas dentro de uma planta. No entanto, uma construo compreendendo uma sequência nucleotídica que codifica a proteína de interesse e a proteína que modifica o perfil de glicosilação da proteína de interesse pode também ser utilizada.

Por "co-expresso" entende-se que duas ou mais do que duas sequências de nucleótidos são expressas aproximadamente ao mesmo tempo dentro da planta e dentro do mesmo tecido da planta. No entanto, as sequências de nucleotídeos não precisam ser expressas exatamente ao mesmo tempo. Pelo contrário, as duas ou mais sequências de nucleótidos são expressas de uma maneira tal que os produtos codificados têm uma

possibilidade de interagir. Por exemplo, a proteína que modifica a glicosilação da proteína de interesse pode ser expressa antes ou durante o período em que a proteína de interesse é expressa, de modo que a modificação da glicosilação das duas ou mais de duas sequências nucleotídicas podem ser co-expressas utilizando um sistema de expressão transiente, onde as duas ou mais sequências são introduzidas dentro da planta aproximadamente ao mesmo tempo sob condições em que ambas as sequências são expressas. Alternativamente, uma planta de plataforma compreendendo uma das sequências nucleotídicas, por exemplo, a sequência que codifica a proteína que modifica o perfil de glicosilação da proteína de interesse, pode ser transformada, transitoriamente ou de um modo estável, com uma sequência adicional que codifica a proteína de interesse. Neste caso, a sequência que codifica a proteína que modifica o perfil de glicosilação da proteína de interesse pode ser expressa dentro de um tecido desejado, durante um estágio de desenvolvimento desejado, ou a sua expressão pode ser induzida utilizando um promotor indutível e a sequência adicional codificando a proteína de interesse pode ser expressa em condições semelhantes e no mesmo tecido, para assegurar que as sequências nucleotídicas sejam co-expressas.

As construções da presente invenção podem ser introduzidas em células vegetais utilizando plasmídeos Ti, plasmídeos Ri, vetores de vírus de plantas, transformação direta de ADN, microinjeção, electroporação, infiltração e semelhantes. Para revisões de tais técnicas ver por exemplo, Weissbach e Weissbach, *Methods for Plant Molecular Biology*, Academia Press, New York VIII, pp. 421-463 (1988); Geierman e Corey, *Plant Molecular Biology*, 2d Ed. (1988); e Miki e Iyer, *Fundamentals of Gene Transfer in Plants*. In *Plant Metabolism*, 2d Ed. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison-Wesley, Langmans Ltd. London, pp. 561-579 (1997). Outros métodos incluem a absorção direta de ADN, a utilização de lipossomas, electroporação, por exemplo, utilizando protoplastos, microinjeção, microprojéteis ou "whiskers" e infiltração a vácuo. Veja, por exemplo, Bilang et al. (*Gene* 100: 247-250 (1991)) Scheid et al. (*Mol. Gen. Genet.* 228:

104-112, 1991) Guerche et al. (Plant Science 52: 111-116, 1987) Neuhauser et al. (Theor. Appl Genet. 75: 30-36, 1987) Klein et al., Nature 327: 70-73 (1987); Howell et al. (Ciência 208: 1265, 1980) Horsch et al. (Science 227: 1229-1231, 1985) DeBlock et al., Plant Physiology 91: 694-701, 1989), Liu e Lomonosoff (J. Virol Meth, 105: 343-348, 2002,), Patentes US 4945050; 5036006; 5100792; 6403865; 5625136, (todos os quais são aqui incorporados por referência).

Os métodos de expressão transiente podem ser usados para expressar as construções da presente invenção (ver Liu e Lomonosoff, 2002, Journal of Virological Methods, 105: 343-348; que é aqui incorporado por referência). Em alternativa, um método de expressão transiente baseado em vácuo, como descrito por Kapila et al. 1997 Plant Science 122: 101-108(aqui incorporado por referência) pode ser usado. Estes métodos podem incluir, por exemplo, mas não estão limitados a um método de Agro-inoculação ou de Agro-infiltração, no entanto, outros métodos transitórios podem também ser utilizados como mencionado acima. Com Agro-inoculação ou Agro-infiltração, uma mistura de *Agrobacteria* compreendendo o ácido nucleico desejado entra nos espaços intercelulares de um tecido, por exemplo, as folhas, porção aérea da planta (incluindo caule, folhas e flor), outra porção da planta (caule, raiz, flor), ou a planta inteira. Depois de atravessar a epiderme, o *Agrobacterium* infecta e transfere cópias de t-ADN para as células. O t-ADN é epissalograficamente transcrito e o mRNA é traduzido, levando à produção da proteína de interesse em células infectadas, no entanto, a passagem do t-ADN dentro do núcleo é transitória.

As VLP compreendendo HA quimérico proporcionado pela presente invenção podem ser utilizadas em conjunto com uma vacina de gripe existente, para suplementar a vacina, torná-la mais eficaz ou para reduzir as dosagens de administração necessárias. Tal como seria do conhecimento de um perito na arte, a vacina pode ser dirigida contra um ou mais do que um vírus da gripe. Exemplos de vacinas adequadas incluem, mas não estão limitados a, aqueles comercialmente disponíveis da Sanofi-

Pasteur, ID Biomedical, Merial, Sinovac, Quiron, Roche, MedImmune, GlaxoSmithKline, Novartis, Sanofi-Aventis, Serono, Shire Pharmaceuticals e semelhantes.

Se desejado, as VLPs da presente invenção podem ser misturadas com um adjuvante adequado, como seria do conhecimento de um especialista na técnica. Além disso, a VLP pode ser utilizada numa composição de vacina compreendendo uma dose eficaz da VLP para o tratamento de um organismo alvo, como definido acima. Al disso, a VLP produzida de acordo com a presente invenção pode ser combinada com VLPs obtidas utilizando diferentes proteínas de gripe, por exemplo, neuraminidase (NA).

Por conseguinte, a presente invenção proporciona um método para induzir imunidade à infecção por vírus da gripe num organismo animal ou alvo compreendendo a administração de uma dose eficaz de uma vacina compreendendo uma ou mais de uma VLP. A vacina pode ser administrada oralmente, intradermicamente, intranasalmente, intramuscularmente, intraperitonealmente, intravenosamente ou subcutaneamente.

As composições de acordo com a presente divulgação podem compreender VLPs de duas ou mais estirpes ou subtipos de gripe. "Dois ou mais" refere-se a duas, três, quatro, cinco, seis, sete, oito, nove, 10 ou mais estirpes ou subtipos. As estirpes ou subtipos representados podem ser de um único subtipo (por exemplo, todo o H1N1 ou todo o H5N1) ou podem ser uma combinação de subtipos. Exemplos de subtipos e estirpes incluem H5/Indo, H1/Bri, H1/NC. A escolha da combinação de estirpes e subtipos pode depender da área geográfica dos sujeitos suscetíveis de serem expostos à gripe, proximidade de espécies animais a uma população humana a ser imunizada (por exemplo, espécies de aves aquáticas, animais agrícolas como porcos, etc.) e as estirpes que eles carregam, estão expostos ou provavelmente serão expostos a predições de deriva antigênica dentro de subtipos ou estirpes, ou combinações desses fatores.

As duas ou mais VLPs podem ser expressas individualmente, e as VLPs purificadas ou semi-purificadas subsequentemente

combinadas. Alternativamente, as VLPs podem ser co-expressas no mesmo hospedeiro, por exemplo, uma planta, protão de planta ou célula vegetal. As VLPs podem ser combinadas ou produzidas numa proporção desejada, por exemplo, em rácios equivalentes, ou podem ser combinadas de tal maneira que um subtipo ou estirpe compreende a maioria das VLPs na composição.

Por conseguinte, a divulgação proporciona composições compreendendo VLPs de duas ou mais estirpes ou subtipos.

Também proporcionado um artigo de fabrico, compreendendo material de embalagem e uma composição compreendendo uma VLP compreendendo uma HA quimérica. A composição inclui um excipiente fisiologicamente ou farmacologicamente aceitável, e o material de embalagem pode incluir uma etiqueta que indica os ingredientes ativos da composição (por exemplo, a VLP).

Um kit compreendendo uma composição compreendendo um ácido nucleico codificando um HA quimérico como aqui proporcionado, juntamente com instruções para utilização do ácido nucleico para produção de HA quimérico, ou VLPs compreendendo o HA quimérico também proporcionado. O kit pode ser útil para a produção de VLPs compreendendo o HA quimérico e as instruções podem incluir, por exemplo, informação na expressão do ácido nucleico numa planta ou numa célula vegetal, instruções para colheita e obtenção das VLPs da planta ou tecido da planta.

Noutra forma de realização, é proporcionado um kit para a preparação de um medicamento, compreendendo uma VLP compreendendo uma HA quimérica, juntamente com instruções para a sua utilização. As instruções podem compreender uma série de passos para a preparação do medicamento, sendo o medicamento útil para induzir uma resposta imunitária terapêutica ou profilática num sujeito a quem é administrado. O kit pode ainda compreender instruções que abordam concentrações de dose, intervalos de dose, métodos de administração preferidos ou semelhantes.

SEQ ID NO	Descrição	SEQ ID NO:	Descrição
1	iniciador XmaI- pPlas.c	55	Seq de ácido nucleico de BglII sintético - PDI SP-H1 / NC-SacI / StuI (sem TMD / CT)
2	iniciador SacI-ATG- pPlas.r	56	Seq de ácido nucleico de KpnI - H1 / NC TmD / CT-SacI / StuI compreendendo a sequência codificadora C-ter H1 (A / Nova Caledónia / 20/99 (H1N1) incluindo o TmD e Ctail do local KpnI ao codão de paragem (flanqueado em 3 'por um site duplo SacI / StuI
3	iniciador SacI- PlasTer.c	57	Seq nucleotídico do ácido nucleico da proteína dissulfeto isomerase Os nucleótidos 32-103 codificam o peptídeo sinal do PDI
4	iniciador EcoRI- PlasTer.r	58	Seq do ácido nucleico do promotor do plasto - PDI SP-BglI
5	iniciador Plasto- 443c	59	SEQ de ido nucleico da construo 540 compreendendo promotor de plastocianina de alfafa e 5 'UTR, sequência codificante do péptido sinal da PDI e da H1 forma A / Nova Caldia/20/99 (H1N1), plastocianina 3' UTR e sequências de terminador da alfafa. H1 da sequência de codificação A / Nova Caledónia / 20/1999 está sublinhada.
6	iniciador SpHA (Ind) -Plasto.r	60	SEQ de ácido nucleico do promotor DraI - Plasto-H1 / Bri-SacI compreendendo a região codificante H1 completa (A / Brisbane / 59/07 (H1N1)) (incluindo o ptido sinal e o codão de terminação) flanqueada, em 5 'por plastocianina de alfafa sequências genéticas

			correspondentes aos primeiros 84 nucleótidos a montante do ATG inicial,
7	iniciador Plasto-SpHA	61	SEQ de ácido nucleico da construção 774 compreendendo o promotor de plastocianina de alfafa e 5' UTR, sequência de codificação de hemaglutinina de H1 forma A / Brisbane / 59/07 (H1N1), sequências de terminador de UTR e plastocianina de alfafa 3'
8	iniciador HA (Ind) - Sac.r	62	SEQ de ácido nucleico da construção 828, de PacI (a montante do promotor) para AscI (imediatamente a jusante do terminador NOS). O sequenciador UTR de HTV de 3CPMV foi sublinhado com ATG mutado em negrito.
9	iniciador pBinPlus.2613c	63	Seq do ácido nucleico da construção 690
10	iniciador Mut- <u>ATG115</u> .r	64	Seq de ácido nucleico da construção 691
11	iniciador Mut- <u>ATG161</u> .c	65	Seq de ácido nucleico da construção 696
12	iniciador LC-C5-1.110r	66	SEQ de ácido nucleico da construção 732 compreendendo o promotor CaMV 35S, CPMV-HT 5' UTR, sequência codificadora de hemaglutinina de H1 forma A / Brisbane / 59/07 (H1N1), CPMV-HT 3' UTR e sequências de terminador NOS. A sequência de codificação de H1 / Bri está sublinhada.

13	iniciador E1 H1B-E1 H5I.r	67	Seq do ácido nucleico da construção intermediária 787
14	iniciador E1 H5N-E1 H1B.c	68	Seq do ácido nucleico da construção 733
15	iniciador E2 H5I-RB H1B.r	69	Seq de ácido nucleico da construção 734
16	iniciador RB H1B-E2 H5I.c	70	SEQ de ácido nucleico do promotor DraIII-plasto - H3 / Bri-SacI compreendendo a região codificante completa de H3 (A / Brisbane / 10/07 (H3N2)) (incluindo o péptido sinal e o codão de terminação) flanqueada, em 5 'por plastocianina de alfafa sequências genéticas correspondentes aos primeiros 84 nucleótidos a montante do ATG inicial
17	iniciador E1 H1B- F'1 H5I.r	71	Seq do ácido nucleico da construção 736
18	iniciador F'1 H5N- E1 H1B.c	72	Seq do ácido nucleico da construção 737
19	iniciador F'2 H5I- E2 H1B.r	73	Seq de ácido nucleico do promotor DraIII-plasto - B / Flo-SacI
20	iniciador E2 H1B- F'2 H5I.c	74	Seq do ácido nucleico da construção 739
21	iniciador E1 H5I-E1 H1NC.r	75	Seq do ácido nucleico da construção 745
22	iniciador E1 H1NC- E1 H5I.c	76	Seq do ácido nucleico de Msj 1
23	iniciador E2 H1NC- RB H5I.r	77	Seq de ácido nucleico da construção 850
24	iniciador RB H5I-E2 H1NC.c	78	Seq do ácido nucleico da construção 860

25	iniciador HA-SacI.r	79	Seq de ácido nucleico da construção 870
26	iniciador H1B.c ApaI-	80	Sequência de aminoácidos compreendendo um subdomínio RB de H1 / Bri, um peptídeo sinal H5 / Indo e um complexo de domínio estaminal (SDC) compreendendo subdomínios H5 / Indo F'1, E1, E2, F'2 e F
27	iniciador H1B.r StuI-	81	sequência de aminoácidos compreendendo o complexo do domínio da cabeça de H1 / Bri (HDC) que inclui E1, RB, E2, um péptido de sinal H5 / Indo e complexo do domínio H5 / Indo-estaminal (SDC) compreendendo H5 / Indo F'1, F'2 e F subdomínios
28	iniciador H1B.c SpPDI-	82	sequência de aminoácidos compreendendo um subdomínio RB de H5 / Indo, um péptido de sinalização PDI e complexo domínio do domínio H1 / NC compreendendo F'1, E1, E2 e F'2
29	iniciador H1B.r SacI-	83	sequência de aminoácidos compreendendo um peptídeo sinal de PDI, um ectodomínio de H3 A / Brisbane / 10/2007 e um TMD / CT de H5 A / Indonesia / 5/2005
30	iniciador SpPDI.c ApaI-	84	sequência de aminoácidos compreendendo um ectodomínio de B / Florida / 4/2006 e um TMD / CT de H5 A / Indonesia / 5/2005 codificado pela estrutura de

			leitura aberta no número de construção 745
31	iniciador ApaI-H5 (A-Indo) .1c	85	Seq de ácido nucleico de SacI - plastocianina 3'UTR
32	iniciador H5 (A-Indo) -StuI.1707r	86	Seq de nucleotídeo PDI SP-H1 A / Califórnia / 04/09 A sequência codificante do peptídeo sinal da dissulfeto isomerase da proteína da alfalfa é sublinhada e a sequência codificadora H1 madura é destacada em negrito.
33	iniciador H3B-SpPDI.r	87	sequência de aminoácidos de PDI SP-H1 A / Califórnia / 04/09 Péptido sinal de dissulfeto isomerase de proteína de alfafa é sublinhado
34	iniciador SpPDI-H3B.c	88	Promotor 2X35S.
35	iniciador StuI-H3B.r	89	iniciador PacI - <u>MCS</u> -2X35S.c
36	iniciador TmD H5I-H3B.r	90	iniciador CPMV 5'UTR- 2X35S . r
37	iniciador H3B-TmD H5I.c	91	iniciador 2X35S -CPMV 5'UTR.c
38	iniciador HBF-SpPDI.r	92	iniciador ApaI -M prot.r
39	iniciador SpPDI-HBF.c	93	sequência de ácido nucleico da construção 747. A sequência de codificação da HA quimérica está sublinhada. A sequência do promotor 2X35S é indicada em itálico

40	iniciador StuI-HBF.r	94	sequência de aminoácidos da B-Florida
41	iniciador TmD H5I-BFlo.r	95	sequência de aminoácidos da B-Malásia
42	iniciador B Flo-TmD H5I.c	96	sequência de aminoácidos de H1-Brisbane
43	iniciador Hsp40Luz.1c	97	sequência de aminoácidos de H1-Sol.Isl.
44	iniciador Hsp40Luz-SacI.1272r	98	sequência de aminoácidos de H1-New Cal.
45	iniciador Hsp40Luz-Plasto.r	99	sequência de aminoácidos de H2-Singapore
46	iniciador Hsp70Ara.1c	100	sequência de aminoácidos de H3-Brisbane
47	iniciador Hsp70Ara-SacI.1956r	101	sequência de aminoácidos de H3A-WCN
48	iniciador Hsp70Ara-Plasto.r	102	Sequência de aminoácidos de H5A-Anhui
49	iniciador supP19-plasto.r	103	Sequência de aminoácidos do H5A-Vietnam
50	iniciador supP19-1c	104	sequência de aminoácidos do H5-Indo
51	iniciador SupP19-SacI.r		
52	sequência de ácidos nucleicos da região codificante completa de H5 (A / Indonésia / 5/05 (H5N1)) (incluindo o péptido sinal e o codão stop)		

	flanqueada, em 5 'por um local HindIII e, em 3', por um local SacI		
53	SEQ de ácido nucleico da construção 660 - uma cassete de expressão de HA compreendendo um promotor de plastocianina de alfafa e 5'UTR, sequência de codificação de hemaglutinina de H5 forma A / Indoncia / 5/05 (H5N1), sequências 3'UTR e terminadora de plastocianina de alfafa		
54	Seq de ácido nucleico de H1 / NC nativo (sem TMD / CT)		

A presente invenção será ainda ilustrada nos exemplos seguintes.

As sequências aqui descritas estão resumidas abaixo.

Métodos e Materiais

1. Montagem de cassetes de expressão de HA

A- pCAMBIAPlasto

Todas as manipulações foram realizadas utilizando os protocolos gerais de biologia molecular de Sambrook e Russell (2001, que é aqui incorporado por referência). A Tabela 1 apresenta iniciadores oligonucleotícos utilizados para montagem de cassetes de expressão. O primeiro passo de clonagem consistiu na montagem de um plasmídeo recetor contendo elementos reguladores a montante e a jusante do gene da plastocianina de alfafe. O promotor de plastocianina e as sequências 5'UTR foram amplificados a partir do ADN genómico de alfafe utilizando os iniciadores oligonucleotídicos XmaI-pPlas.c (SEQ ID NO: 1) e SacI-ATG-pPlas.r (SEQ ID NO: 2). O produto de amplificação resultante foi digerido com XmaI e SacI e ligado a pCAMBIA2300 (Cambia, Canberra, Áustria), previamente digerido com as mesmas enzimas, para criar pCAMBIApromoPlasto. Similarmente, as sequências de 3'UTR e o terminador do gene da plastocianina foram amplificados a partir do ADN genómico de alfafe utilizando os seguintes iniciadores: SacI-PlasTer.c (SEQ ID NO: 3) e EcoRI-PlasTer.r (SEQ ID NO: 4), e o produto foi digerido com SacI e EcoRI antes de ser inserido nos mesmos locais de pCAMBIApromoPlasto para criar pCAMBIAPlasto.**B-**

Plasto-Nativo SP-H5 A/ Indonésia/5/05 (número de construção 660)

Um fragmento que codifica hemaglutinina da estirpe de gripe A / Indonésia / 5/05 (H5N1; Acc. No. LANL ISDN125873) foi sintetizado por Epoch Biolabs (Sugar Land, TX, EUA). O fragmento produzido, contendo a região codificadora H5 completa incluindo o péptido sinal nativo flanqueado por um sítio HindIII imediatamente a montante do ATG inicial, e um local SacI imediatamente a jusante do codão stop (TAA), é apresentado em (SEQ ID NO: 52 Figura 17). A região codificadora de H5 foi clonada numa cassette de expressão baseada em plastocianina pelo modo de ligação baseado em PCR apresentado em Darveau et al. (1995). Resumidamente, foi obtida uma primeira amplificação por PCR utilizando os iniciadores Plasto-443c (SEQ ID NO: 5;) e SpHA (Ind) -Plasto.r (SEQ ID NO: 6) e pCAMBIApromoPlasto como modelo. Em paralelo, foi realizada uma segunda amplificação com os iniciadores Plasto-SpHA (SEQ ID NO: 7) e HA (Ind) -Sac. r (SEQ ID NO: 8) com

fragmento codificador de H5 (SEQ ID NO: 52; Figura 17) como modelo. A amplificação obtida de ambas as reações foi misturada e a mistura serviu como molde para uma terceira reação (reação de montagem) utilizando Plasto-443c (SEQ ID NO: 5) e HA (Ind) -Sac.r (SEQ ID NO: 8) como iniciadores. O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (na extremidade 3' do fragmento) e clonado em pCAMBIAPlasto previamente digerido com as mesmas enzimas. O plasmídeo resultante, denominado 660, apresentado na Figura 18 (SEQ ID NO: 53). O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (na extremidade 3' do fragmento) e clonado em pCAMBIAPlasto previamente digerido com as mesmas enzimas. O plasmídeo resultante, denominado 660, apresentado na Figura 18 (SEQ ID NO: 53). O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (na extremidade 3' do fragmento) e clonado em pCAMBIAPlasto previamente digerido com as mesmas enzimas. O plasmídeo resultante, denominado 660, apresentado na Figura 18 (SEQ ID NO: 53).

C-Plasto-PDI SP-H1 A/Nova Caledónia/20/99 (número de construção 540)

O quadro de leitura aberto do gene H1 da estirpe A / New Caledonia / 20/99 (H1N1) foi sintetizado em dois fragmentos (Plant Biotechnology Institute, National Research Council, Saskatoon, Canadá). Um primeiro fragmento sintetizado corresponde à sequência de codificação H1 de tipo selvagem (nº de acesso GenBank AY289929; SEQ ID N°: 54; Figura 19) sem a sequência codificadora do peptídeo sinal na extremidade 5' e a sequência codificadora do domínio transmembrana na extremidade 3'. A extremidade 5' do fragmento é composta pelos últimos nucleótidos que codificam PDISP (incluindo um local de restrição BglIII) e foi adicionado um local duplo SacI / StuI imediatamente a jusante do codão de terminação na extremidade terminal 3' do fragmento, para produzir SEQ. ID No: 55 (Figura 20).

O primeiro fragmento H1 foi digerido com BglIII e SacI e clonado nos mesmos locais de um vetor binário (pCAMBIAPlasto) contendo o promotor da plastocianina e 5'UTR fundido ao peptídeo sinal do gene da

dissulfeto isomerase da proteína da alfafa (PDI) (nucleotídeos 32-103; Acesso No. Z11499; SEQ ID No: 57; Figura 22) resultando num gene quimérico PDI-H1 a jusante dos elementos reguladores da plastocianina. A sequência da cassete base de plastocianina, contendo o promotor e o péptido de sinalização PDI até ao local de restrição BglIII e o terminador de plastocianina a jusante de um local SacI, apresentada na SEQ ID NO. 58 (Figura 23). A adição da extremidade C-terminal da região de codificação H1 (codificando o domínio transmembranar e a cauda citoplasmática) foi obtida inserindo o fragmento sintetizado (SEQ ID NO. 56; Figura 21) previamente digerida com KpnI e SacI, no plasmídeo de expressão H1. A construção resultante, denominada 540, é apresentada na SEQ ID NO. 59 (Figura 24).

D-Plasto-nativo SP-H1 A/Brisbane/59/07 (número de construção 774)

A cassete de expressão número 774, conduzindo a expressão de H1 de A / Brisbane / 59/07, foi montada como se segue. Foi sintetizado um fragmento sintético compreendendo a sequência completa de codificação de hemaglutinina (de ATG para parar) flanqueada em 3' por sequências de genes de plastocianina de alfafa correspondendo aos primeiros 84 nucleótidos a montante da plastocianina ATG começando com um local de restrição DraIII. Os fragmentos sintéticos também compreendiam um sítio SacI imediatamente a jusante do codão de terminação.

O fragmento sintético foi sintetizado pela Top Gene Technologies (Montreal, QC, Canadá). O fragmento sintetizado é apresentado na SEQ ID NO. 60 (Figura 25). Para a montagem da cassete de expressão completa, o fragmento sintético foi digerido com DraIII e SacI e clonado em pCAMBIAPlasto previamente digerido com as mesmas enzimas para dar a construção 774 (SEQ ID NO. 61; Figura 26).

E-CPMV HT-LC C51 (número de construção 828)

As cassetes de expressão CPMV-HT utilizam o promotor 35S para controlar a expressão de um mRNA compreendendo uma sequência codificadora de interesse flanqueada, em 5', pelos nucleotídeos 1-512 do RNA2 do vírus do mosaico do feijão-caupi (CPMV) com ATG mutado

nas posições 115 e 161 e em 3', pelos nucleótidos 3330-3481 do CPMV RNA2 (correspondendo ao 3' UTR) seguido pelo terminador NOS. Plasmídeo pBD-C5-1LC, (Sainsbury et al. 2008; Plant Biotechnology Journal 6: 82-92 e Publicação PCT WO 2007/135480), foi usado para a montagem de cassetes de expressão de hemaglutinina baseados em CPMV-HT. A mutação dos ATGs na posição 115 e 161 do CPMV RNA2 foi feita usando um método de ligação baseado em PCR Darveau et al. (Methods in Neuroscience 26: 77-85 (1995))). Dois PCRs separados foram realizados usando pBD-C5-1LC como modelo. Os iniciadores para a primeira amplificação foram pBinPlus.2613c (SEQ ID NO: 9) e Mut- ATG115.r (SEQ ID NO: 10). Os iniciadores para a segunda amplificação foram Mut- ATG161.c (SEQ ID NO: 11) e LC-C5-1.110r (SEQ ID NO: 12). Os dois fragmentos obtidos foram misturados e utilizados como molde para uma terceira amplificação utilizando pBinPlus.2613c (SEQ ID NO: 9) e LC-C5-1.110r (SEQ ID NO: 12) como iniciadores. O fragmento resultante foi digerido com PacI e ApaI e clonado em pBD-C5-1LC digerido com as mesmas enzimas. A construção gerada, denominado 828, é apresentado na Figura 27 (SEQ ID NO: 62).

Domínio de ligação ao recetor F-H1 A/Brisbane/59/07 (RB) na estrutura H5 A/Indonésia/5/05 (número de construção 690)

Um HA quimérico foi feito substituindo o domínio RB no H5 A / Indonésia / 5/05 pelo de H1 A / Brisbane / 59/07 usando o método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995))). Numa primeira ronda de PCR, um segmento do promotor de plastocianina fundido com o péptido sinal natural, os domínios F'1 e E1 do H5 A / Indonésia / 5/05 foram amplificados utilizando os iniciadores Plasto-443c (SEQ ID NO: 5) e E1H1B-E1H5I.r (SEQ ID NO: 13) com o número de construção 660 (SEQ ID NO: 53, Figura 18) como modelo. Um segundo fragmento, compreendendo a sequência de codificação do domínio R1 H1A / Brisbane / 59/07, foi amplificado com os iniciadores E1 H5N-E1H1B.c (SEQ ID NO: 14) e E2H5I-RB HIB.r (SEQ ID NO: 15) utilizando o número de construção 774 (SEQ ID NO: 61; Figura 26) como modelo. Um terceiro fragmento compreendendo os domínios E2, F'2, F, transmembranar e citoplasmático de H5 A /

Indonésia/ 5/05 foi amplificado utilizando os iniciadores RB H1B-E2H5I.c (SEQ ID NO: 16) e HA (Ind) -SacI.r (SEQ ID NO: 8) com o número de construção 660 (SEQ ID NO: 53; Figura 18) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ ID NO: 5) e HA (Ind) -SacI.r (SEQ ID NO: 8). O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (após o codão de terminação) e clonado no número de construção 660 (SEQ ID NO: 53; Figura 18), previamente digerido com as mesmas enzimas de restrição para originar o número 690 (SEQ ID NO: 63). A construção é apresentada na Figura 28. previamente digerido com as mesmas enzimas de restrição para dar o número 690 (SEQ ID NO: 63). A construção é apresentada na Figura 28. previamente digerido com as mesmas enzimas de restrição para dar o número 690 (SEQ ID NO: 63). A construção é apresentada na Figura 28.

Domínios G-H1 A / Brisbane / 59/07 esterase e de ligação ao recetor (E1-RB-E2) na estrutura H5 A / Indonésia / 5/05 (número de construção 691)

Um HA quimérico foi montado substituindo os domínios E1-RB-E2 em H5 A / Indonésia / 5/05 por aqueles de H1 A / Brisbane / 59/07 usando o método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Numa primeira ronda de PCR, um segmento do promotor de plastocianina fundido com o péptido sinal natural e o domínio F '1 de H5 A / Indonésia / 5/05 foi amplificado utilizando os iniciadores Plasto-443c (SEQ ID NO: 5) e E1 H1B-F'1 H5I.r (SEQ ID NO: 17) com o número de construção 660 (SEQ ID NO: 53; Figura 18) como modelo. Em paralelo, dois outros fragmentos foram amplificados. O segundo fragmento, contendo a sequência de codificação dos domínios E1-RB-E2 de H1A / Brisbane / 59/07, foi amplificado com os iniciadores F'1H5N-E1H1B.c (SEQ ID NO: 18) e F'2H5I-E2. HIB.r (SEQ ID NO: 19) utilizando o número de construção 774 (SEQ ID NO: 61; Figura 26) como modelo. Para o terceiro fragmento, os domínios F'2, F, transmembrana e citoplasmático de H5 A / Indonésia

/ 5/05 foram amplificados usando os iniciadores E2 HIB-F'2H5I.c (SEQ ID NO: 20) e HA (Ind) - SacI.r (SEQ ID NO: 8) com o número de construção 660 (SEQ ID NO: 53; Figura 18) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ ID NO: 5) e HA (Ind) -SacI.r (SEQ ID NO: 8). O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (após o codão de terminação) e clonado no número de construção 660 (SEQ ID NO: 53; Figura 18), previamente digerido com as mesmas enzimas de restrição para originar o número 691 (SEQ ID NO: 64). A construção é apresentada na Figura 29.

Domínio de ligação ao recetor H-H5 A / Indonésia / 5/05 (RB) na estrutura H1 A / Nova Caledônia / 20/99 (número de construção 696)

Um HA quimérico foi feito substituindo o domínio RB no H1 A / Nova Caledônia / 20/99 pelo de H5 A / Indonésia / 5/05, usando o método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995))). Num primeiro ciclo de PCR, um segmento do promotor de plastocianina fundido ao péptido sinal da dissulfureto isomerase da proteína de alfafa (PDISP; N° de Acesso Z11499; nucleótidos 32-103 da SEQ ID NO: 57; Figura 22), o F'1 e os domios E1 de H1A / New Caledonia / 20/99 foram amplificados utilizando os iniciadores Plasto-443c (SEQ ID NO: 5) e E1H5I-E1H1NC.r (SEQ ID NO: 21) com o número de construção 540 (SEQ ID NO: 59; Figura 24) como modelo. Um segundo fragmento, compreendendo a sequência de codificação do domínio H5 A / Indonésia / 5/05 RB, foi amplificado com os iniciadores E1H1NC-E1H5I.c (SEQ ID NO: 22) e E2H1NC-RBH5I.r (SEQ ID NO: 23) utilizando o número de construção 660 (SEQ ID NO: 53; Figura 18) como modelo. O fragmento Athird compreendendo os domios E2, F'2, F, transmembranar e citoplasmico de H1 A / New Caledonia / 20/99 foi amplificado utilizando os iniciadores RB H5I-E2 HINC.c (SEQ ID NO: 24) e HA-SacI. r (SEQ ID NO: 25) com o número de construção 540 (SEQ ID NO: 59; Figura 24) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ

ID NO: 5) e HA-SacI.r (SEQ ID NO: 25). O fragmento resultante foi digerido com BglII e SacI e clonado no número de construção 540 (SEQ ID N59; Figura 24) previamente digerido com as mesmas enzimas de restrição para dar o número de construção 696 (SEQ ID NO: 65). A construção é apresentada na Figura 30.

I- Montagem de H1 A / Brisbane / 59/2007 na cassete de expressão CPMV-HT (número de construção 732).

A sequência de codificação de HA de H1 A / Brisbane / 59/2007 foi clonada em CPMV-HT como se segue. Os sítios de restrição ApaI (imediatamente a montante de ATG) e StuI (imediatamente a jusante do codão stop) foram adicionados à sequência de codificação da hemaglutinina através da realização de uma amplificação por PCR com os iniciadores ApaI-H1B.c (SEQ ID NO: 26) e StuI-H1B. r (SEQ ID NO: 27) utilizando o número de construção 774 (SEQ ID NO: 61; Figura 26) como modelo. O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO: 62; Figura 27) digerido com as mesmas enzimas. A cassete resultante foi denominada número de construção 732 (SEQ ID NO: 66; Figura 31).

J- Montagem de SpPDI-H1 A / Brisbane / 59/2007 na cassete de expressão CPMV-HT (número de construção 733).

Uma sequência que codifica o péptido de sinal da dissulfureto isomerase da proteína alfafa (PDISP; nucleótidos 32-103 da SEQ ID NO: 57 Figura 22; Acesso N Z11499) foi fundida com a sequência de codificação HA0 de H1 de A / Brisbane / 59/2007, e o fragmento resultante foi clonado em CPMV-HT do seguinte modo. A sequência de codificação H1 foi amplificada com os iniciadores SpPDI-H1B.c (SEQ ID NO: 28) e SacI-H1B.r (SEQ ID NO: 29) utilizando a construção 774 (SEQ ID NO: 61; Figura 26) como modelo. O fragmento resultante consistia na sequência de codificação H1 flanqueada, em 5', pelos últimos nucleótidos que codificam PDISP (incluindo um local de restrição BglII) e, em 3', por um sítio de restrição SacI. O fragmento foi digerido com BglII e SacI e clonado no número de construção 540 (SEQ

ID NO: 59; Figura 24) previamente digerido com as mesmas enzimas de restrição. A sequência de codificação da cassete intermédia, denominado número de construção 787 (SEQ ID NO: 67), é apresentada na Figura 32. Os sítios de restrição ApaI (imediatamente a montante ATG) e StuI (imediatamente a jusante do codão de paragem) foram adicionados à sequência de codificação da hemaglutinina através da realização de uma amplificação por PCR com os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e StuI-H1B.r (SEQ ID N27) utilizando o nero de construo 787 (SEQ ID N 67; Figura 32) como modelo. O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO: 62; Figura 27) digerido com as mesmas enzimas. A cassete resultante foi denominada número de construção 733 (SEQ ID NO: 68; Figura 33).

K- Montagem do domínio de ligação ao recetor H1A / Brisbane / 59/07 (RB) no esqueleto de H5A / Indonésia / 5/05 na cassete de expressão de CPMV-HT (número de construção 734).

A sequência de codificação de HA quimérico consistindo no domínio RB de H1A / Brisbane / 59/07 em estrutura H5A / Indonésia / 5/05 foi clonada em CPMV-HT como se segue. Os locais de restrição ApaI (imediatamente a montante de ATG) e StuI (imediatamente a jusante do codão de terminação) foram adicionados sequência de codificação da hemaglutinina quimérica através da realização de uma amplificação por PCR com os iniciadores ApaI-H5 (A-Indo) .1c (SEQ ID NO: 31) e H5 (A-Indo) -StuI.1707r (SEQ ID NO: 32) utilizando o número de construção 690 (SEQ ID NO: 63; Figura 28) como modelo. O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO: 62; Figura 27) digerido com as mesmas enzimas. A cassete resultante foi denominada número de construção 734 (SEQ ID NO: 69; Figura 34).

L-Montagem de SpPDI-H3 A / Brisbane / 10/2007 em cassete de expressão CPMV- HT (número de construção 736).

Uma sequência codificando o péptido sinal de alfafa PDI fundido a HAO a partir de H3A / Brisbane / 10/2007 foi clonada em CPMV- HT como se segue. Primeiro, sintetizou-se um fragmento sintético compreendendo a sequência codificadora da hemaglutinina completa (de ATG para parar) flanqueada em 3 'pela sequência do gene da plastocianina de alfafa correspondente aos primeiros 84 nucleótidos (começando com um local de restrição DraIII) a montante da plastocianina ATG. O fragmento sintético também continha um local SacI imediatamente após o códon de parada. Fragmento sintético foi sintetizado pela Top Gene Technologies (Montreal, QC, Canadá). O fragmento sintetizado é apresentado na SEQ ID NO: 70 (Figura 35) e foi utilizado como molde para posterior ligação baseada em PCR.

Em segundo lugar, o péptido sinal da dissulfureto isomerase da proteína de alfafa (PDISP) (nucleótidos 32-103; N.º de acesso Z11499; SEQ ID NO: 57; Figura 22) foi ligado à sequência de codificação HAO de H3 de A / Brisbane / 10/2007 ao longo com o sítio de restrição ApaI imediatamente a montante do local de restrição ATG e StuI a jusante do codão de paragem como se segue. A PDISP foi ligada à sequência de codificação H3 pelo método de ligação baseado em PCRDarveau et al. (Methods in Neuroscience 26: 77-85 (1995))). Num primeiro ciclo de PCR, amplificou-se o péptido sinal PDISP utilizando os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e H3B-SpPDI.r (SEQ ID NO: 33) com o número de construção 540 (SEQ ID NO: 59; Figura 24) como modelo. Paralelamente, outro fragmento contendo uma porção da sequência de codificação de H3A / Brisbane / 10/2007 (do codão 17 ao codão de paragem) foi amplificado com os iniciadores SpPDI-H3B.c (SEQ ID NO: 34) e StuI-H3B.r (SEQ ID NO: 35) utilizando fragmentos previamente sintetizados (SEQ ID NO: 70; Figura 35) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e StuI-H3B.r (SEQ ID NO: 35). O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO:

62; Figura 27) digerida com as mesmas enzimas. A cassette resultante foi denominada número de construção 736 (SEQ ID NO: 71; Figura 36).

M-Montagem de SpPDI-H3 A / Brisbane / 10/2007 quimérico (ectodomínio) + H5 A / Indonésia / 5/2005 (TmD + cauda Cyto) na cassette de expressão CPMV-HT (número de construção 737).

Uma sequência codificando o péptido sinal de alfalfa PDI fundido ao ectodomínio de H3A / Brisbane / 10/2007 e aos domínios transmembranar e citoplasmático de H5A / Indonésia / 5/2005 foi clonada em CPMV-HT como se segue. A sequência codificadora da PDISP-H3 foi fundida ao domínio transmembranar H5 pelo método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Num primeiro ciclo de PCR, um fragmento compreendendo péptido sinal PDISP e ectodomínio de H3 Brisbane foi gerado por amplificação (com local de restrição ApaI a montante do ATG inicial PDISP) utilizando os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e TmD H5I-H3B.r (SEQ ID NO: 36) com o número de construção 736 (SEQ ID NO: 71; Figura 36) como modelo. Paralelamente, outro fragmento contendo os domínios transmembranar e citoplasmático de H5 da Indonésia foi amplificado com os iniciadores H3B-TmD H5I.c (SEQ ID NO: 37) e H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32) utilizando construo número 660 (SEQ ID NO. 53; Figura 18) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32). O fragmento resultante foi digerido com enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO: 62; Figura 27) digerido com as mesmas enzimas. A cassette resultante foi denominada número de construção 737 (SEQ ID NO: 72; Figura 37).

N-Montagem de SpPDI-HA B / Florida / 4/2006 na cassette de expressão CPMV- HT (número de construção 739).

Uma sequência codificando o péptido sinal de alfalfa PDI fundido a HA0 de HA B / Florida / 4/2006 foi clonada em CPMV- HT como se segue. Primeiro, sintetizou-se um fragmento sintético compreendendo a

sequência codificadora da hemaglutinina completa (de ATG para parar) flanqueada em 3' pela sequência do gene da plastocianina de alfafa correspondente aos primeiros 84 nucleótidos (começando com um local de restrição DraIII) a montante da plastocianina ATG. O fragmento sintético também continha um local de restrição SacI imediatamente após o códon de terminação. O fragmento sintético foi sintetizado pela Epoch Biolabs (Sugar Land, Texas, EUA). O fragmento sintetizado é apresentado na SEQ ID NO: 73 (Figura 38) e foi utilizado como molde para posterior ligação baseada em PCR.

Segundo, o peptídeo sinal da dissulfeto isomerase da proteína da alfafa (PDISP) (nucleotídeos 32-103 da SEQ ID NO: 57; Figura 22; Acesso No. Z11499) foi ligado à sequência de codificação HA de HA de B / Florida / 4/2006 juntamente com o sítio de restrição ApaI imediatamente a montante de ATG e o sítio de restrição a jusante de sítio de restrição StuI como se segue. A PDISP foi ligada à sequência de codificação de HA pelo método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Num primeiro ciclo de PCR, amplificou-se o péptido sinal PDISP utilizando os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e HBF-SpPDI.r (SEQ ID NO: 38) com o número de construção 540 (SEQ ID NO: 59; Figura 24) como modelo. Paralelamente, outro fragmento contendo uma porção da sequência de codificação de HA de B / Florida / 4/2006 (do códon 16 ao códon de parada) foi amplificado com os iniciadores SpPDI-HBF.c (SEQ ID NO: 39) e StuI-HBF.r (SEQ ID NO: 40) utilizando fragmentos previamente sintetizados (SEQ ID NO: 73; Figura 38) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e StuI-HBF.r (SEQ ID NO: 40). O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no vetor de construção 828 (SEQ ID NO: 62; Figura 27) digerida com as mesmas enzimas. A cassete resultante foi denominada número de construção 739 (SEQ ID NO: 74; Figura 39).

O-Montagem de SpPDI-HA quimérico B / Florida / 4/2006 (ectodomínio) + H5 A / Indonésia / 5/2005 (TmD + cauda Cyto) na cassette de expressão CPMV- HT (número de construção 745) .

Uma sequência que codifica o péptido de sinal PDI de alfafa fundido com o ectodomínio de HA B / Florida / 4/2006 e aos domínios transmembranar e citoplasmático de H5 A / Indonésia/ 5/2005 foi clonada em CPMV- HT como se segue. A sequência codificadora do ectodomínio PDISP-B / Florida / 4/2006 foi fundida aos domínios transmembrana e citoplasmática H5 pelo método de ligação baseado em PCR Darveau et al. (Methods in Neuroscience 26: 77-85 (1995))). Num primeiro ciclo de PCR, um fragmento compreendendo péptido sinal PDISP fundido com o ectodomínio de HA B / Florida / 4/2006 foi gerado por amplificação utilizando os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e TmD H5I-B Flo.r (SEQ ID NO: 41) com o número de construção 739 (SEQ ID NO: 74; Figura 39) como modelo. Paralelamente, outro fragmento contendo os domínios transmembranar e citoplasmático H5 da Indonésia foi amplificado com os iniciadores B Flo-TmD H5I.c (SEQ ID NO: 42) e H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32) utilizando construo número 660 (SEQ ID NO. 53; Figura 18) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores ApaI-SpPDI.c (SEQ ID NO: 30) e H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32). O fragmento resultante foi digerido com as enzimas de restrição ApaI e StuI e clonado no número de construção 828 (SEQ ID NO: 62; Figura 27) digerido com as mesmas enzimas. A cassette resultante foi denominada número de construção 745 (SEQ ID NO: 75; Figura 40).

Montagem em P de SpPDI-HA quimérico B / Florida / 4/2006 + H5 A / Indonésia / 5/2005 (TmD + cauda Cyto) na cassette de expressão 2X35S-CPMV- HT (número de construo 747) .

Uma sequência codificando o péptido sinal de alfalfa PDI fundido com HA0 de HA B / Florida / 4/2006 e ao domínio transmembranar e citoplasmático de H5 A / Indonésia/ 5/2005 foi clonada em 2X35S-CPMV- HT como se segue. A troca do promotor foi realizada utilizando o método de ligação baseado em PCR Darveau et al. (Methods in

Neuroscience 26: 77-85 (1995))). Um primeiro fragmento contendo o promotor 2X35S (SEQ ID NO: 88; Figura 50A) foi amplificado por PCR com os iniciadores **PacI** - MCS -2X35S.c (SEQ ID NO: 89) e CPMV 5'UTR- 2X35S.r (SEQ ID NO: 90):

PacI - MCS -2X35S.c (ID SEQ NO: 89)
 ATAGTA **TTAATTAA** GTCGACAAGCTTGCATGCCTGCAGG TCAAC

CPMV 5'UTR- 2X35S .r (SEQ ID NO: 90)
 TCAAAACCTATTAAGATTTTAATA CCTCTCCAAATGAAATGAACTTCC

utilizando um plasmídeo contendo o promotor 2X35S como modelo. Paralelamente, foi realizada uma segunda PCR utilizando os iniciadores 2X35S- CPMV 5'UTR.c (SEQ ID NO: 91) e ApaI- M prot.r (SEQ ID NO: 92):

2X35S-CPMV 5'UTR.c (ID SEQ NO: 91)
TTGGAGAGG TATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGG

ApaI-M prot.r (ID SEQ NO: 92)
 TCTCCAT GGGCCC GACAAATTTGGGCAGAAATATACAGAAGCTTA

utilizando a construção 745 (SEQ ID NO 75; Figura 40) como modelo. Os dois fragmentos obtidos foram então misturados e utilizados como molde para um segundo ciclo de PCR (reação de montagem) com **PacI** - MCS- 2X35S.c (SEQ ID NO: 89 e ApaI -M prot.r (SEQ ID NO: 92) como iniciadores. O fragmento resultante foi digerido com PacI e ApaI e clonado na construção 745 (SEQ ID NO: 75; Figura 40) digerido com as mesmas enzimas de restrição da sequência da cassette de expressão, denominada construção 747 (SEQ ID NO: 93), é apresentado na Figura 50B.

2. Montagem de cassetes de expressão acompanhantes

Duas cassetes de expressão de proteína de choque térmico (Hsp) foram montadas. Em uma primeira cassette, a expressão da HSP70 citosólica de *Arabidopsis thaliana* (ecótipo Columbia) (Athsp70-1 em Lin et al. (2001) Cell Stress and Chaperones 6: 201-208) é

controlado por um promotor quimérico que combina elementos dos promotores de Plastocianina de Nitrito Redutase (Nir) e alfafa (Nir/Plasto). Uma segunda cassette que compreende a região de codificação da HSP40 citosólica da alfafa (MsJ1; Frugis et al. (1999) *Plant Molecular Biology* 40: 397-408) sob o controle do promotor quimérico Nir/Plasto foi também montado.

Um plasmídeo aceitador contendo o promotor da redutase de nitrato de alfafa (Nir), o gene repórter GUS e o terminador NOS no vetor binário da planta foi primeiro montado. Plasmídeo pNir3K51 (descrito anteriormente em Patente US 6420548) foi digerido com HindIII e EcoRI. O fragmento resultante foi clonado em pCAMBIA2300 (Cambia, Canberra, Austrália) digerido pela mesma enzima de restrição para dar pCAMBIA-Nir3K51.

As sequências de codificação para Hsp70 e Hsp40 foram clonadas separadamente no plasmídeo aceitador pCAMBIANir3K51 pelo método de ligação baseado em PCR apresentado em Darveau et al. (*Methods in Neuroscience* 26: 77-85 (1995))).

Para Hsp40, a sequência de codificação Msj1 (SEQ ID NO: 76; Figura 41) foi amplificada por RT-PCR a partir de ARN total de folhas de alfafa (ecotipo Rangelander) utilizando os iniciadores Hsp40Luz.1c (SEQ ID NO: 43) e Hsp40Luz-SacI.1272r (SEQ ID NO: 44). Uma segunda amplificação foi realizada com os iniciadores Plasto-443c (SEQ ID NO: 5) e Hsp40Luz-Plasto.r (SEQ ID NO: 45) com a construção 660 (SEQ ID NO: 53; Figura 18) como modelo. Os produtos de PCR foram então misturados e utilizados como molde para uma terceira amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ ID NO: 5) e Hsp40Luz-SacI.1272r (SEQ ID NO: 44). O fragmento resultante foi digerido com HpaI (no promotor de plastocianina) e clonado em pCAMBIANir3K51, previamente digerido com HpaI (no promotor de Nir) e SacI e arquivado com ADN-polimerase de T4 para gerar extremidades cegas. Os clones obtidos foram triados para orientação correta e sequenciados para integridade da sequência. O plasmídeo resultante, denominado R850, é apresentado na Figura 42 (SEQ ID NO: 77). A região

codificadora do Athsp70-1 foi amplificada por RT-PCR a partir do ARN da folha de *Arabidopsis* utilizando os iniciadores Hsp70Ara.1c (SEQ ID NO: 46) e Hsp70Ara-SacI.1956r (SEQ ID NO: 47). Uma segunda amplificação foi realizada com os iniciadores Plasto-443c (SEQ ID NO: 5) e Hsp70Ara-Plasto.r (SEQ ID NO: 48) com a construção 660 (SEQ ID NO: 53; Figura 18) como modelo. Os produtos de PCR foram então misturados e utilizados como molde para uma terceira amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ ID NO: 5) e Hsp70ARA-SacI.1956r (SEQ ID NO: 47). O fragmento resultante foi digerido com HpaI (no promotor de plastocianina) e clonado em pCAMBIANir3K51 digerido com HpaI (no promotor de Nir) e SacI e depositado com ADN-polimerase de T4 para gerar extremidades cegas. Os clones obtidos foram triados para orientação correta e sequenciados para integridade da sequência. O plasmídeo resultante, denominado R860, é apresentado na Figura 43 (SEQ ID NO: 78).

Um plasmídeo de expressão Hsp duplo foi montado como se segue. A R860 (SEQ ID NO: 78; Figura 43) foi digerida com BsrBI (a jusante do terminador NOS), tratada com polimerase de ADN T4 para gerar uma extremidade romba e digerida com SbfI (a montante do promotor NIR / Plasto quimérico). O fragmento resultante (promotor quimérico Nir / Plasto-sequência de codificação de HSP70-terminador Nos) foi clonado em R850 (SEQ ID NO: 77; Figura 42) previamente digerido com SbfI e SmaI (ambos localizados no sítio de clonagem múltipla a montante Nir / Plasto quimérico promotor). O plasmídeo resultante, denominado R870, é apresentado na Figura 44 (SEQ ID NO: 79).

3. Montagem de outras cassetes de expressão

Cassete de expressão HcPro

Uma construção HcPro (35HcPro) foi preparada como descrito em Hamilton et al. (2002). Todos os clones foram sequenciados para confirmar a integridade das construções. Os plasmídeos foram utilizados para transformar *Agrobacterium tumefaciens* (AGL1; ATCC, Manassas, VA 20108, EUA) por electroporação (Mattanovich et al., 1989). A integridade de

todas as estirpes de *A. tumefaciens* foi confirmada pelo mapeamento de restrição.

Cassete de expressão P19

A sequência codificadora da proteína p19 do vírus do tomateiro de tomateiro (TBSV) foi ligada à cassete de expressão de plastocianina de alfafa pelo método de ligação baseado em PCR apresentado em Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Numa primeira ronda de PCR, um segmento do promotor de plastocianina foi amplificado utilizando os iniciadores Plasto-443c (SEQ ID NO: 5) e supP19-plasto.r (SEQ ID NO: 49) com a construção 660 (SEQ ID NO: 53) como modelo. Paralelamente, um outro fragmento contendo a sequência de codificação de p19 foi amplificado com os iniciadores supP19-1c (SEQ ID NO: 50) e SupP19-SacI.r (SEQ ID NO: 51) utilizando a construção 35S: p19 como descrito em Voinnet et al. (The Plant Journal 33: 949-956 (2003)) como modelo. Os produtos de amplificação foram então misturados e utilizados como molde para um segundo ciclo de amplificação (reação de montagem) com os iniciadores Plasto-443c (SEQ ID NO: 5) e SupP19-SacI.r (SEQ ID NO: 51). O fragmento resultante foi digerido com BamHI (no promotor de plastocianina) e SacI (no final da sequência de codificação de p19) e clonado no número de construção 660 (SEQ ID NO: 53; Figura 18), previamente digerido com as mesmas enzimas de restrição para dar o número de construção R472. O plasmídeo R472 é apresentado na Figura 45.

Número de construção 443

O número de construção 443 corresponde a pCAMBIA2300 (vetor vazio).

Tabela 1: Iniciadores oligonucleotídicos para montagem de cassetes de expressão

SEQ ID	Título	Sequência oligonucleotídica (5' ... 3')
1	Xmal-pPlas.c	AGTCCCGGGCGTGTATATTATATGTTCTC
2	SacI-ATG-pPlas.r	AATAGAGCTCCATTTTCTCTCAAGATGATTAATTAATTAGTC
3	SacI-PlasTer.c	AATAGAGCTCGTTAAATGCTTCTGCTCTCATTTATATATG
4	EcoRI-PlasTer.r	TTACGAATTCCTCTTCTTAATTTGGTGTACTATCAITTAATCAAGGGGA
5	Plasfo-443c	GTATTAGTAATTAGAAITTTGGTGTCTC
6	SphAI(Ind)-Plasfo.r	GCAAGAGAAGAGCAGTATTTTCTCCATTTTCTCTCAAGATGATTA
7	Plasfo-SphA.c	TTATCATCTTGAGAGCAAAATGGAGAAATAGTGCCTTCTTCTGC
8	HA(Ind)-SacI	ACTTTGAGCTCTTAATGCAAAATTCGCAATTGTAAACGA
9	pBinPlus.2613c	AGGAAGCGAAGAAAGCGAAAGGAG
10	Mut-ATG115.r.r	GTGCCGAAGCAGCATCTGACAAACGTTGAAGATCGCTCACCAGAGAAA GACAAGAGA
11	Mut-ATG161.c	GTTGTCAAGATCGTGCTTCGGCACCAGTACAAACGTTTCTTTCACCTGAA CCGA
12	LC-C5-1.110r	TCCTCTGGAGTCAAGACAGGGGTGG
13	E1 H1B-E1 H51.r	TCAATAGTCAGCGAAATGCCCCGTGGTAACAGAGCTCATTTGGTTGATT GGCCT
14	E1 H5N-E1 H1B.c	ATGACCCTCTGTTACCCAGGGCATTTCCGCTGACTATGAGGAAGTGGAG G
15	E2 H51-RB H1B.r	CCAAATTCACTTTTCATTAATTCCTGATCCAAAGCCTCTACTCAGTGGGA
16	RB H1B-E2 H51.c	GGCTTTGGATCAAGGAATTATGAAAAGTGAATTGGAATATGGTAACCTG CAAC
17	E1 H1B-F1 H51.r	GGCTATTCCTTTTAATAGGAGAGAGCTTCCCGTTGTGTCTTTTCCAG
18	F1 H5N-E1 H1B.c	AACGGGAAGCTCTGCTTAATTAATAAGGAATAGCCCCACACTACAATTGGG T
19	F2 H51-E2 H1B.r	GGAGTTTGACACTGTGTGTTGCAITTTATCCATTGGTGCAITTTGAGTTG
20	E2 H1B-F2 H51.c	AAATGCACCAATGGATAAATGCAACACCAAGTGTCAAACTCCAAATGGG G

SEQ ID	Título	Sequência oligonucleotídica (5'....3')
SEQ ID	Título	Oligonucleotide sequence (5'...3')
21	E1 HSI-E1 HINC.r	TCCTCATAGTCGTTGAACCTCCCTGGGTAAACATGTTCCATTCACAGGA
22	E1 HINC-E1 HSI.c	CTGAGATGGACATGTTTACCACGAGGAGTTTCAACGACTATGAAGAA C
23	E2 HINC-RB HSI.r	ATTGAGGTGATGATTCCTGAGTCCCTTTCTTGACAATTTGTATGC ATA
24	RB HSI-E2 HINC.c	GTCAAGAAAGGGAGCTCAGCAATCATCACCCTCAATGCACCAATGGA T
25	HA-SacI.r	TTAACTTAGAGCTCTAGATGCAATATCTACACTGCMAAGACC
26	Apal-H1B.c	TGTCGGGCCCATGAAAGTAAACTACTGGTCCTGTATGCACATT
27	Sul-H1B.r	AAATAGGCGCTTAGATGCAATATCTACACTGTAAAGACCCATTGGA
28	SppDI-H1B.c	TTCTCAGATCTTCGCTGACACCAATATGTATAGCTTACCATGCTAACA AC
29	SacI-H1B.r	CTTAGAGCTCTAGATGCAATATCTACACTGTAAAGACCCATTGGA
30	Apal-SppDI.c	TTGTGGGGCCCATGGCGAAAAAGCTTGCCGATTTTGGGCTTATTGT
31	Apal-H5 (A-Indo).1.c	TGTCGGGGCCCATGGCGAAAAATAGTCTCTCTCTTCTTGCAT
32	H5 (A-Indo)-Sul.1707r	AAATAGGCGCTTAAATGCATAATCTGCAITGTACGA
33	H3B-SppDI.r	TGTCATTTCCGGGAAGTTTTGAGCGAAGATCTGAGAAAGGAACCA
34	SppDI-H3B.c	TCTCAGATCTTGGCTCAAAAACCTTCCCGGAAATGACAAACAGCAGC
35	Sul-H3B.r	AAATAGGCGCTTCAATGCAATATGTTGCACCTAATGTTGCCTTT
36	Tmd HSI-H3B.r	ATTGTGATGTTCTTATGACTTCAGCTCAAGGCCCTTGATCTGG
37	H3B-Tmd HSI.c	TGAGCTGAAGTCAATAGGAACCTTACCAATACTGTCAATTTATTCAA C
38	HBf-SppDI.r	GTTATTCAGATGCAGATTCGATCAGCGAAGATCTGAGAAAGGAACCAA CAC
39	SppDI-HBf.c	CAGATCTTCGCTGATCGAATCTGCACCTGGAATTAACAATCTTCAAACTC ACC
40	Sul-HBf.r	AAATAGGCGCTTATAGACAGATGAGGCAATGAACGTTGTCTCTGG

SEQ ID	Título	Sequência oligonucleotídica (5'...3')
SEQ ID	Title	Oligonucleotide sequence (5'---3')
41	TMD H5l-B Flo.r	TGACAGTATTGGTAGTTATCCATCCATCATCTTAAGATGC
42	B Flo-TMD H5l.c	GGATTGGATAACTACCAAAATACTGTCAATTATTCACAGTGGCGAG TTC
43	Hsp-40Luz.1c	ATGTTTGGCGCGGACCAAC
44	Hsp40Luz-Sacl.1272r	AGCTGAGCTCCTACTGTTGAGCGCATTGCAC
45	Hsp40Luz-Plasto.r	GTTGGTCCGCGCCCAACATTTTCTCTCAAGATGAT
46	Hsp70Ara.1 c	ATGTCGGTAAGGAGAAGGA
47	Hsp70Ara-Sacl.1956r	AGCTGAGCTCTTAGTCGACCTCCTCGAICTTAG
48	Hsp70Ara-Plasto.r	TCCTTCTCCTTACCCGACATTTTCTCTCAAGATGAT
49	supP19-plasto.r	CCTTGTATAGCTCGTTCCATTTTCTCTCAAGATG
50	supP19-1c	ATGGAACGAGCTATACAAG
51	SupP19-Sacl.r	AGTCGAGCTCTTACTCGCTTCTTTTTCGAAG
89	Pacl-MCS-2X35S.c	AATTGTTAATTAAGTCGACACAGCTTGCAATGCCCTGCAGGTCAC
90	CPMV 5'UTR- <u>2X35S.r</u>	TCAAAACCTATTAGATTTTAATACCTCTCCAAATGAATGAACCTTCC
91	<u>2X35S</u> -CPMV 5'UTR.c	<u>TTTGAGAGAGG</u> CTATTAAAAATCTTAATAGCTTTTGATATAAAGCGAACGTC G
92	<u>Apal</u> -M prot.r	TCTCCATGGGCCCCGACAAATTTGGGACAGAAATATACAGAACGCTTA

Tabela 2: Estirpes de Agrobacterium usadas para expressão de hemaglutininas de gripe com peptídeos sinalizadores nativos ou PDI

Estirpe Agro	HA expresso	Peptídeo de sinal	Cassete de expressão
AGL1 / 540	H1 (A / Nova Caledônia / 20/99)	PDI	Plastocianina
AGL1 / 774	H1 (A / Brisbane / 59/2007)	nativo	Plastocianina
AGL1 / 787	H1 (A / Brisbane / 59/2007)	PDI	Plastocianina
AGL1 / 732	H1 (A / Brisbane / 59/2007)	nativo	35S / CPMV-HT
AGL1 / 736	H3 (A / Brisbane / 10/2007)	PDI	35S / CPMV-HT
AGL1 / 660	H5 (A / Indonésia / 5/2005)	nativo	Plastocianina
AGL1 / 739	B (B / Flórida / 4/2006)	PDI	35S / CPMV-HT
AGLI / 828	CPMV HT-LC C51	C51LC	35S / CPMV-HT
AGLI / 690	H1 / Bris RB + H5 / Indo SDC	nativo	Plasto
AGLI / 691	H1 / Bri E1-RB-E2 + H5 SDC	nativo	Plasto
AGLI / 696	H5 / Indo RB + H1 / NC SDC	PDI	Plasto
AGLI / 733	H1 / Bri	PDI	35S / CPMV-HT
AGLI / 734	H1 / Bri RB + H5 / Indo SDC	nativo	35S / CPMV-HT
AGLI / 737	Ectodomínio H3 / Bri + H5 / Indo TDC	PDI	35S / CPMV-HT
AGLI / 745	Ectodomínio B / Flo H5 / Indo TDC	PDI	35S / CPMV-HT
AGL1 / 747	Ectodomínio B / Flo H5 / Indo TDC	PDI	2X35S / CPMV-HT

4. Preparação da biomassa vegetal, inóculo, agro-infiltração e colheita

Plantas de *Nicotiana benthamiana* foram cultivadas a partir de sementes em apartamentos preenchidos com substrato comercial de turfa. As plantas foram deixadas crescer na estufa sob um fotoperíodo de 16/8 e um regime de temperatura de 25°C dia/20°C noite. Três semanas após a sementeira, as plântulas individuais foram colhidas, transplantadas em vasos e deixadas a crescer na estufa durante mais três semanas sob as mesmas condições ambientais. Antes da transformação, os botões apicais e axilares foram removidos em vários momentos, conforme indicado abaixo, seja comprimindo os botões da planta, ou tratando quimicamente a planta.

As *agrobacterias* transfectadas com cada construção foram cultivadas num meio YEB suplementado com ácido 2- [N-morfolino] etanossulfónico (MES), acetoseringona 20, 50 / ml de canamicina e 25 / ml de carbenicilina pH5,6 até atingirem um OD₆₀₀ entre 0,6 e 1,6. As suspensões de *Agrobacterium* foram centrifugadas antes da utilização e ressuspenderam-se em meio de infiltração (10 mM de MgCl₂ e MES 10 mM pH 5,6). A infiltração com seringa foi realizada como descrito por Liu e Lomonossoff (2002, Journal of Virological Methods, 105: 343-348). Para a infiltração por vácuo, as suspensões de *A. tumefaciens* foram centrifugadas, ressuspensas no meio de infiltração e armazenadas durante a noite a 4°C. No dia da infiltração, os lotes de cultura foram diluídos em 2,5 volumes de cultura e deixados a aquecer antes da utilização. Plantas inteiras de *N. benthamiana* ou *N. tabacum* foram colocados de cabeça para baixo na suspensão bacteriana em um tanque de aço inoxidável hermético sob um vácuo de 20-40 Torr por 2 min. Após a seringa ou infiltração a vácuo, as plantas foram devolvidas à estufa durante um período de incubação de 4-5 dias até à colheita. Salvo especificação em contrário, todas as infiltrações foram realizadas como co-infiltração com AGL1 / 35S-HcPro numa proporção de 1: 1, exceto para as estirpes contendo cassetes CPMV-HT

que foram co-infiltradas com a estirpe AGL1 / R472 numa proporção de 1:1.

5. Amostragem de folhas e extração total de proteínas

Após a incubação, a parte aérea das plantas foi colhida, congelada a -80°C , esmagada em pedaços. As proteínas solúveis totais foram extraídas por homogeneização (Polytron) de cada amostra de material vegetal triturado congelado em 3 volumes de Tris 50 mM, pH 8, NaCl 0,15 M, metabissulfito de sódio 0,04% e fluoreto de fenilmetanossulfonilo 1 mM. Após a homogeneização, as suspensões foram centrifugadas a 20.000 g por 20 min a 4°C e esses extratos brutos clarificados (sobrenadante) foram mantidos para análise. O teor de proteína total dos extratos brutos clarificados foi determinado pelo ensaio de Bradford (Bio-Rad, Hercules, CA) utilizando albumina de soro bovino como padrão de referência.

6. Análise de Proteína e Immunoblotting

As concentrações de proteína foram determinadas pelo ensaio de proteína BCA (Pierce Biochemicals, Rockport IL). As proteínas foram separadas por SDS-PAGE sob condições redutoras e coradas com Azul de Coomassie. Os géis corados foram digitalizados e a análise densitométrica realizada usando o ImageJ Software (NIH).

Proteínas da fração de eluição de SEC foram precipitadas com acetona (Bollag et al., 1996), ressuspensas em 1/5 do volume em equilíbrio/tampão de eluição e separadas por SDS-PAGE sob condições redutoras e eletrotransferidas em membranas de polivinil difluoreto (PVDF) (Roche Diagnostics Corporation, Indianapolis, IN) para imunodeteção. Antes da imunotransferência, as membranas foram bloqueadas com leite desnatado a 5% e Tween-20 a 0,1% em solução salina tamponada com Tris (TBS-T) durante 16-18 h a 4°C .

A imunotransferência foi realizada por incubação com um anticorpo adequado (Tabela 6), em 2 $\mu\text{g/ml}$ em leite desnatado a 2% em TBS-Tween 20 a 0,1%. Os anticorpos secundários utilizados para detecção de quimiluminescência foram os indicados na Tabela 4, diluições como indicadas em leite desnatado a 2% em TBS-Tween 20 a 0,1%. Complexos

imunorreativos foram detectados por quimiluminescência usando luminol como substrato (Roche Diagnostics Corporation). A conjugação de peroxidase de rábano-enzima do anticorpo IgG humano foi realizada utilizando o kit de conjugação de peroxidase ativada EZ-Link Plus® (Pierce, Rockford, IL). Vírus inativados inteiros (WIV), utilizados como controles de detecção para os subtipos H1, H3 e B, foram adquiridos ao Instituto Nacional de Normas Biológicas e Controle (NIBSC).

Tabela 3: Condições de eletroforese, anticorpos e diluições para imunotransferência de proteínas expressas.

Subtipo de HA	Estirpe da gripe	Condição de eletroforese	Anticorpo primário	Diluição	Anticorpo secundário	Diluição
H1	A / Brisbane / 59/2 007 (H1N1)	Reduzindo	FII 10-150	4 µg / ml	Antirato de cabra (JIR 115-035-146)	1:10000
H1	A / Nova Caledônia / 20/99 (H1N1)	Reduzindo	FII 10-150	4 µg / ml	Antirato de cabra (JIR 115-035-146)	1:10 000
H3	A / Brisbane / 10/2 007 (H3N2)	Não Reduzindo	NIBSC 08/124	1: 4000	Coelho anti- ovelha (JIR 313-035-045)	1:10 000
H5	A / Indonésia / 5/2 005 (H5N1)	Reduzindo	ITC IT-003- 005V	1: 4000	Cabra anti- coelho (JIR 111-035-144)	1:10 000
B	B / Flórida / 4/200 6	Não Reduzindo	NIBSC 07/356	1: 2000	Coelho anti- ovelha (JIR 313-035-045)	1:10 000

FII:	Fitzgerald	Industries	International,	Concord,	MA,	EUA;	
NIBSC:	Instituto	Nacional	de	Padrões	Biológicos	e	Controle;
JIR:	Jackson	ImmunoResearch,	West	Grove,	PA,	EUA;	
ITC:	Immune Technology Corporation, Woodside, NY, EUA;						

7. Esclarecimento e concentração antes da SEC

Para melhorar a resolução e aumentar o sinal nas frações de eluição, os extratos a serem carregados na cromatografia por exclusão de tamanho, os extratos de proteína bruta foram clarificados e concentrados utilizando o método seguinte. Os extratos foram centrifugados a 70 000 g, 4 durante 20 min e o sedimento foi lavado duas vezes por ressuspensos em 1 volume (comparado com o volume inicial do extrato) de tampão de extração (50 mM Tris pH 8, 0,15 M NaCl) e centrifugação a 70 000 g, 4 durante 20 min. O sedimento resultante foi ressuspensão em 1/3 do volume (comparado com o volume inicial do extrato) e as proteínas (incluindo VLPs) foram precipitadas pela adição de 20% (p/v) de PEG 3350 seguido por incubação em gelo durante 1 h. As proteínas precipitadas foram recuperadas por centrifugação a 10 000 g, 4°C, 20 min e ressuspensas em 1/15 do volume (comparado com o volume inicial de extrato) do tampão de extração.

8. Cromatografia de exclusão de tamanho do extrato proteico

Colunas de cromatografia de exclusão por tamanhos (SEC) de 32 ml de esferas de alta resolução Sephacryl™ S-500 (S-500 HR: GE Healthcare, Uppsala, Suíça, Cat. No. 17-0613-10) foram empacotadas e equilibradas com tampão de equilíbrio/eluição (Tris pH 8 mM, NaCl 150 mM). Um mililitro e meio de extrato de proteína bruta foi carregado na coluna seguido de um passo de eluição com 45 mL de tampão de equilíbrio/eluição. A eluição foi recolhida em frações de 1,5 mL. O conteúdo proteico relativo das frações eluídas foi monitorizado misturando 10 µl da fração com 200 µl de reagente corante de proteína de diluição Bio-Rad (Bio-Rad, Hercules, CA). A coluna foi lavada com 2 colunas volumes de NaOH 0,2 N seguidos de 10 volumes de coluna de Tris 50 mM pH 8, NaCl 150 mM, 20% de etanol. Cada separação foi seguida por uma calibração da coluna com Blue Dextran 2000 (GE Healthcare Bio-

Science Corp., Piscataway, NJ, EUA). Os perfis de eluição de Blue Dextran 2000 e as proteínas solúveis do hospedeiro foram comparados entre cada separação para assegurar a uniformidade dos perfis de eluição entre as colunas utilizadas.

Exemplo 1: Estratégia de troca de domínio para domínios RB e/ou Esterase em hastes do subtipo de gripe.

O subdomínio RB de H5 / Indo pode ser substituído por um subdomínio RB de H1, H3 ou B HA. O HA quimérico resultante proporciona um SDC H5 / Indo para formar VLPs e apresenta o subdomínio RB compreendendo sítios imunogénicos de H1, H3 ou B. O subdomínio H5 / Indo RB pode ser inserido numa haste H1 (H1/NC). As Figuras 15A e 15B ilustram as sequências de aminoácidos nas fusões dos subdomínios indicados, e as sequências de aminoácidos dos respetivos subdomínios são fornecidas na Figura 2 (construções 690, 734, 696 e 691) e na Tabela 4 (construções 900 e 745) e 5 (construções 910, 920 e 930). As sequências de aminoácidos ilustradas na Figura 2 e as Tabelas 4 e 5 não incluem sequências peptídicas de sinal.

Tabela 4 Subdomínios e Gripe HA quimérica. Gripe HA quimérica compreendendo subdomínio RB heterólogo.

Const. Nº	haste N-terminal: F1+E1	cabeça RB	haste C-terminal: E2+F'2-stop
Const No	N-terminal Stem: -F'1+E1	RB head	C-terminal stem: E2+F'2-Stop
900 (SEQ ID NO: 105)	H5/Indo DQICGYHANNSTEQ VDTMEKNVTVTHA QDLEKTHNGKLCDL DGVKPLILRDCSVAG WLLGNPMCDFEIN PEWSYVEKANFTND LCYP	H3/Brisbane YDVPDYASLSLVAASSGTLLEN NESFNWVGVTQNGTSSACRRS NNSFSLNLWLILKFKYPALN VTNPNNKEKDKLYINGVHHPO TUNDQFLYAQASGRFTVSTKRS QQYVTPNGSRPRYRNPERSIV WTVKPGDILINSTGNLAPRGY FKRSCKSS	H5/Indo IMKSELEYGNCNTKQTPMGAINSSMPPHNHPLTGCEPKYV VKSRLVLATGILNSPQRESRRKRGIFGAIAGFEGGWQGMV MYDGYGYVHHSNEQSGYAADKESTQKADGVTNKVNSH DKMNTQFEAVGREFNLERHENLNKKMEDGFLDVTWYNALV ELLVLMENERTLDFHDSNVKYLVDKVLQLRDNAKELGNG CFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREISGVKLE SIGTYQLSTVSTYASSLALIMMAGLSLWMCNSGSLQCRGI
745 (SEQ ID NO: 106)	H5/Indo DQICGYHANNSTEQ VDTMEKNVTVTHA QDLEKTHNGKLCDL DGVKPLILRDCSVAG WLLGNPMCDFEIN PEWSYVEKANFTND LCYP	B/Florida IMHDLIKIRQLPNLIRGYENRL STQNVDAEAKAPGQPYRLQISG SCPNATSKSGFEATMAWAVPK DNKNATNPILVEVPYCTEGE DQITVWGFHSDNKTQMKNLVG DSNPQKFTBSANGVTTHYVSQI GSPFDQTEGGLPQSGRIYVDY MMQKPKRTGTVYQRCVLLPQ KVVCAQGRK	H5/Indo IMKSELEYGNCNTKQTPMGAINSSMPPHNHPLTGCEPKYV KSNRLVLATGILNSPQRESRRKRGIFGAIAGFEGGWQGMV DQWYGYHHSNEQSGYAADKESTQKADGVTNKVNSHDKM NTQFEAVGREFNLERHENLNKKMEDGFLDVTWYNALV LMENERTLDFHDSNVKYLVDKVLQLRDNAKELGNGCFEY HKCDNECMESIRNGTYNYPQYSEEARLKREISGVKLESIGTY QLSYSTVASSLALIMMAGLSLWMCNSGSLQCRGI

Os aminoácidos 1-92 da SEQ ID NO: 105 são um domínio F'1 + E1 de H5 / Indo; os aminoácidos 93-259 são o domínio da cabeça RB de H3 / Brisbane; os aminoácidos 260-548 são o domínio E2 + F'2 do H5 / Indo.

Os aminoácidos 1-92 da SEQ ID NO: 106 são um domínio F'1 + E1 de H5 / Indo; os aminoácidos 93-276 s o domínio principal RB de B / Florida; os aminoácidos 277-565 são o domínio E2 + F'2 de H5 / Indo.

Tabela 5 Subdomínios e Gripe HA guinéica. Gripe HA guinéica compreendendo subdomínio RB heterólogo

Const. Nº	haste N-terminal: F1	cabeçaE1-RB-E2	haste C-terminal: F'2-stop
Const. cl. No.	N-terminal Stem: F'1	E1-RB-E2 head	C-terminal stem: F'2-Stop
910 (SEQ ID NO: 107)	H5/Indo DQICIGYHANNST EQVDIMEKNVTV THAQDILEKIHNG KLC	H3/Brisbane DSPHQLDQENCTLDALLGDPOCDG FQNKRWDLFVERSKAYSNCPYDV PDYASLSLVAASSGTLEFNNESENV TOVTQNGISSACIRSSNSPFLNW LTHLKRYPALNVTMPNNKFPDKLY RWGVHHPQIDMDQFLYAQASGRIT VSTKRSQOTVIRPNIGSRPRVRNPSRI SYWWTVRKQDMLINSTONILAPROY FKRSOKSSMRSDAPIGK	H5/Indo CQTPMGAINSSMPFNHPLTJGCEPRVYKSNRLVELATG LNSPQRESRRKKRLFGAAGREGWQGMVDGWYG YHHSNQSGGYAADKESTQKADGVYTNKVNISDKMN TQFEAVGREFNLERENLNKKMEDGFLDVWYTNAE LVLMENERTLDFHDSNKNLYDKVRLQKDNARELG NGCFEYHRCNNECMESIRNGTYNYPQYSEARLKREE ISGVKLESQTYQLSNSTVASSLALANMAGLSLWMC SNGLQCRICI
920 (SEQ ID NO: 108)	H5/Indo DQICIGYHANNST EQVDIMEKNVTV THAQDILEKIHNG KLC	B/Florida PDCLNCTDLDAALGRPMCVGTPSA KASLHEVKNPVTSCTPRMHDKTSAR QLPNLLKQVENHLSQNVDAEKAP GQPYRLCTSCSPNATSKSGFEATM AWAVPRDNMNNATNPLTVEVPVCT EGEDQITVWCFHSDNKTQMKNLYG DSNPQKFTSSANGVTHYVSQIESHP DQTEDEGLFQSGRINVDYMMOKPG KTGTIVYQRGVILFQKVCASQGRSK VRGSLPLGEAD	H5/Indo CQTPMGAINSSMPFNHPLTJGCEPRVYKSNRLVELATG LNSPQRESRRKKRLFGAAGREGWQGMVDGWYG YHHSNQSGGYAADKESTQKADGVYTNKVNISDKMN TQFEAVGREFNLERENLNKKMEDGFLDVWYTNAE LVLMENERTLDFHDSNKNLYDKVRLQKDNARELG NGCFEYHRCNNECMESIRNGTYNYPQYSEARLKREE ISGVKLESQTYQLSNSTVASSLALANMAGLSLWMC SNGLQCRICI
930 (SEQ ID NO: 109)	H1/NC DTICIGYHANNST DTVDIVLEKNVTV THSVNLEDSHNG KLC	H5/Indo DLIDGVNPLLRDCSVAQWLLGNPMAC DEFINVPESVYVERKANFTNDLCYFG SFNDYEEELKHLSRNHFEKQIHPKSS WSDHEASSGVSSACPYLQSPSEFRN VWVLEKCNSTYPTIKKSYNNNTQED LLVLWGHHPNDAAFQTRLVQNPTT VBSIGSTLNQRLVPKIAITRSKVNQOS GRMEFTVTLKPNDAINFESNGNFA PEYAVYKIVKKGDSAMKSELEYGN	H1/NC CDARQTPQAGAINSSLFQNVHPVTJGCEPRVYVSAXL RMVLTGLRNPSIQRSKQIFGAGCFEGQWGTGVYDGYG YHHSNQSGGYAADKESTQKADGVYTNKVNISDKMN TQFEAVGREFNLERENLNKKMEDGFLDVWYTNAE LVLMENERTLDFHDSNKNLYDKVRLQKDNARELG NGCFEYHRCNNECMESIRNGTYNYPQYSEARLKREE ISGVKLESQTYQLSNSTVASSLALANMAGLSLWMC SNGLQCRICI

Os aminoácidos 1-42 da SEQ ID NO: 107 são um domínio F '1 do terminal N de H5 / Indo; os aminoácidos 43-228 são o domínio principal E1-RB-E2 de H3 / Brisbane; os aminoácidos 229-507 são o domínio F'2 do H5 / Indo.

Os aminoácidos 1-42 da SEQ ID NO: 108 são um domínio F '1 do terminal N de H5 / Indo; os aminoácidos 43-281 são o domínio principal E1-RB-E2 de B / Florida; os aminoácidos 282-556 são o domínio F'2 do H5 / Indo.

Os aminoácidos 1-42 da SEQ ID NO: 109 são um domínio F'1 no terminal N de H1 / NC; os aminoácidos 43-273 são o domínio da cabeça E1-RB-E2 de H5 / Indo; os aminoácidos 274-548 s o domínio F'2 de H1 / NC.

Os pontos de fusão para as várias quimeras foram selecionados de modo a estarem tão próximos (mas não necessariamente diretamente) dos terminais N e C dos vários subdomínios - sem querer ficarem limitados pela teoria, essas fusões foram selecionadas de forma a maximizar a estabilidade do HA quimérico. Por exemplo, a conservação de estrutura e sequência é observada no terminal N do subdomínio RB (Ha et al. 2002, EMBO J. 21: 865-875). Uma região menos variável na sequência primária é encontrada na tríade CF / YP localizada em aproximadamente 15 aminoácidos antes, no subdomínio E1. Esta cisteína está envolvida na ponte dissulfureto # 3, que conservada entre as HAs (ver Figuras 46 e 47). Uma junção nesta Cys pode proporcionar uma estabilidade adequada ou superior à HA quimérica em relação à sequência nativa. A extremidade do terminal C do RB fornece recursos conservados: por exemplo, um resíduo Ser conservado na posição -1 e o subdomínio E2 começa com uma folha beta observada em todo o HA no alinhamento (Ha et al. 2002, EMBO J. 21: 865-875). Portanto, o terminal C deste RB pode ser fundido ao aminoácido de iniciação desta estrutura de folha beta do subdomínio E2. Além disso, o padrão de ponte dissulfureto não é alterado, ou não é alterado substancialmente, para as quimeras compreendendo subdomínios RB de H1 / NC, HIBri, H3 / Bri ou B / Flo em um H5 / Indo SDC e para H5 / Indo RB subdomínio em H1 SDC (total de 6), mas uma ponte dissulfeto será adicionada (ponte nº 8) no híbrido

HA de BBR na haste H5. Esta adição de ponte dissulfureto não deve interferir com o enrolamento do HA (porque está localizado dentro do domínio RB e o Cys é adjacente na sequência), e pode produzir um híbrido HA ainda mais estável.

Os subdomínios E1-RB-E2 de um primeiro tipo de gripe foram substituídos por subdomínios E1-RB-E2 de um segundo tipo de gripe. Tal disposição pode apresentar um maior número de aminoácidos do segundo tipo na superfície do H5-VLP. Neste exemplo, o HDC de H1, H3 ou B foi colocado num H5 / Indo SDC, e um HDC de H5 / Indo num H1 / NC SDC (Tabela 5).

A junção do HDC foi definida com um resíduo de cisteína conservada (compreendendo a ponte dissulfureto # 6 de HA tipo A e # 7 em HA tipo B). A junção da HDC no terminal C do subdomínio E2 foi definida com outro resíduo de cisteína conservada compreendendo a ponte dissulfureto # 6 (o segundo aminoácido do subdomínio F'2) do tipo H1 ou H3 de gripe num SDC de H5 / Indo ou para gripe tipo H5 em um SDC de H1. Para a quimera gripe B, a junção foi estabelecida a conexão na primeira cisteína compreendendo a ponte dissulfeto # 4 (localizada a 4 aminoácidos do subdomínio F'2 e conservada entre os HAs). As quimeras resultantes não exibem qualquer alteração nos padrões de ponte dissulfureto - os HAs H1 / H3 / H5 híbridos contêm 6 pontes de dissulfido e o híbrido B tem 7 deles.

Exemplo 2: Substituição dos subdomínios de ligação ao recetor (RB) ou de ligação ao recetor e esterase (E1-RB-E2) de H5 A / Indonésia / 5/05 pela de H1 A / Brisbane / 59/2007: Comparação da expressão para quimérico e formas nativas.

Para combinar o alto nível de acumulação de VLPs de H5 A / Indonésia / 5/05 com as características de antigenicidade de H1 A / Brisbane / 59/2007, hemaglutininas quiméricas foram projetadas compreendendo domínios de H1 A / Brisbane / 59/2007 fundidos a um H5 A / Indonésia

/ 5/05 cluster de domínio de tronco. As cassetes de expressão para a expressão das fusões de hemaglutinina H5 / H1 estão representadas na Figura 1 e as sequências de aminoácidos das proteínas de fusão maduras produzidas ilustradas na Figura 2.

Comparar o nível de acumulação de hemaglutininas quiméricas H5 / H1 com as suas formas nativas, *Nicotiana benthamiana* plantas foram infiltradas com AGL1 / 774, AGL1 / 691 e AGL1 / 690, e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulação de cada forma de HA nas folhas agro-infiltradas, as proteínas foram extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos monoclonais anti-HA. Uma banda única de aproximadamente 75 kDa (Figura 3), correspondendo em tamanho à forma HA0 não-clivada de hemaglutinina da gripe, foi detetada em extratos de folhas infiltradas com AGL1 / 690 mas não em AGL1 / 774 ou AGL1 / 691, indicando que a hemaglutinina quimérica compreendendo a região de ligação ao recetor de H1 A / Brisbane / 59/2007 fusionada à estrutura H5 A / Indonésia / 5/05 acumulou um nível mais elevado do que a forma nativa de H1 A / Brisbane / 59/2007 (AGL1 / 774) e a hemaglutinina quimérica combinando as regiões de ligação esterase e recetor de H1 A / Brisbane / 59/2007 com a estrutura H5 A / Indonésia / 5/05. O vírus inteiro inativado (WIV) (H1 A / Brisbane / 59/2007) usado como controle positivo foi detetado como múltiplas bandas com uma banda principal a aproximadamente 80 kDa, correspondendo ao peso molecular do precursor HA0 de H1 A / Brisbane / 59/2007. Estes resultados demonstraram que a substituição da região de ligação ao recetor de H5 A / Indonésia / 5/05 pela de H1 A / Brisbane / 59/2007 gerou uma hemaglutinina quimérica que apresentou a região antigénica de H1 e que se acumulou em um nível mais alto que o nativo H1 A / Brisbane / 59/2007 em plantas. No entanto, a hemaglutinina quimérica em que as regiões de ligação da esterase e do recetor de H5 A / Indonésia / 5/05 foram substituídas pelas de H1 A / Brisbane / 59/2007 não se acumulou a um nível detetável nas plantas.

A fusão da região de ligação ao recetor de H1 A / Brisbane / 59/2007 à estrutura H5 A / Indonésia / 5/05 como um método para aumentar a acumulação de VLP apresentadoras de antigénio H1 em plantas foi reavaliada sob o controlo de uma cassette de expressão forte baseada em CPMV-HT. Esta estratégia de fusão também foi comparada com a substituição do peptídeo sinal como meio de aumentar o nível de acumulação. As cassetes de expressão para a expressão das fusões de hemaglutinina H5 / H1 sob CPMV-HT estão representadas na Figura 8 e a sequência de aminoácidos da proteína de fusão madura produzida é apresentada na Figura 2.

As plantas de *Nicotiana benthamiana* foram infiltradas com AGL1 / 732, AGL1 / 733 ou AGL1 / 734, e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulação de cada forma de HA nas folhas agro-infiltradas, as proteínas foram primeiro extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos policlonais anti-H1 (Brisbane). Uma banda única de aproximadamente 75 kDa (Figura 6), correspondendo em tamanho à forma HA0 não clivada de hemaglutinina de gripe, foi detetada em extratos de folhas infiltradas com AGL1 / 732, AGL1 / 733 e AGL1 / 734. Entretanto, embora a hemaglutinina tenha sido detetada em todos os extratos analisados, diferenças importantes no acúmulo puderam ser notadas. Enquanto a expressão de H1A / Brisbane / 59/2007 foi dificilmente detetável sob estas condições quando se utiliza o seu peptídeo sinal natural (732), a substituição do peptídeo sinal pelo de PDI resultou em maior acúmulo de H1 A maduro / Brisbane / 59/2007 (733), e a hemaglutinina H5 / H1 quimérica (734) acumulou-se no nível mais alto. Em conjunto, estes resultados mostram que a fusão do domínio de ligação ao recetor de H1 em uma estrutura H5 leva a um alto acúmulo de hemaglutinina apresentadora de antígenos H1 e que o nível de acumulação obtido para essa fusão em plantas é maior do que o obtido com o nativo. forma com ou sem substituição do péptido sinal.

Exemplo 3: Substituição do subdomínio de ligação ao recetor (RB) de H1 A / Nova Caledónia / 20/99 com o de H5 A / Indonésia / 5/05. Comparação de expressão para formas quiméricas e nativas.

O uso de uma estrutura H1 (de A / New Caledonia / 20/99) para a apresentação da região antigénica de H5 também foi avaliado. As cassetes de expressão para a expressão da fusão de hemaglutinina H1 / H5 estão representadas na Figura 1 e a sequência de aminoácidos da proteína de fusão madura produzida apresentada na Figura 2.

Para comparar o nível de acumulação de hemaglutinina H1 / H5 quimérica com a sua forma nativa, plantas de *Nicotiana benthamiana* foram infiltradas com AGL1 / 660 e AGL1 / 696, e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulação de cada forma de HA nas folhas agro-infiltradas, as proteínas foram extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos policlonais anti-H5 (Indica). Uma faixa única de aproximadamente 75 kDa (Figura 7), correspondendo em tamanho à forma HA0 não-clivada de hemaglutinina da gripe, foi detetada em extratos de folhas infiltradas com AGL1 / 660 e AGL1 / 696, indicando que tanto o H5 A nativo / Indonésia / 5/05 e a hemaglutinina quimérica H1 / H5 se acumulam em alto nível nas plantas.

Exemplo 4: Substituição do ectodomínio de H5 A / Indonésia / 5/05 com o de H3 ou B. Comparação da expressão para formas quiméricas e nativas.

A fusão da ectodamaina de H3 A / Brisbane / 10/2007 ou B Florida / 4/2006 aos subdomínios transmembrana e citoplasmática de H5 A / Indonésia / 5/05 foi avaliada como uma estratégia para apresentar regiões antigénicas de hemaglutinina de H3 e B estirpes, aumentando o seu nível de acumulação nas plantas. As cassetes de expressão para a expressão das fusões de hemaglutinina H5 / H3 e H5 / B estão representadas na Figura 10 e os aminoácidos na margem das fusões são apresentados na Figura 11.

O nível de acumulação de hemaglutinina H5 / B quimérica (745) foi comparado com o de HA B nativa (739) em plantas de *Nicotiana*

benthamiana. As plantas foram infiltradas com AGL1 / 739 e AGL1 / 745, e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulaçãoo de cada forma de HA nas folhas agro-infiltradas, as proteínas foram primeiro extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos policlonais anti-B (Florida). Uma faixa única de aproximadamente 75 kDa (Figura 14), correspondendo em tamanho à forma HA0 não-clivada de hemaglutinina da gripe, foi detetada em extratos de folhas de uma planta infiltrada com AGL1 / 739 enquanto as 3 plantas infiltradas com AGL1 / 745 mostraram sinal positivo correspondendo à hemaglutinina, indicando que a forma quimérica H5 / B da hemaglutinina se acumulou mais regularmente em níveis elevados do que a forma nativa da hemaglutinina B.

Similarmente, o nível de acumulação de hemaglutinina H5 / H3 quimérica (737) foi comparado com o de sua forma nativa (736) em plantas de *Nicotiana benthamiana*. As plantas foram infiltradas com AGL1 / 736 e AGL1 / 737 e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulação de cada forma de HA nas folhas agro-infiltradas, as proteínas foram extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos policlonais anti-H3 (Brisbane). Uma banda única de aproximadamente 75 kDa (Figura 15), correspondendo em tamanho à forma HA0 não-clivada da hemaglutinina da gripe, foi detetada em extratos de folhas infiltradas com AGL1 / 736 e AGL 1/737. Este resultado indica que a fusão de subdomínios transmembrana e citoplasmática de H5 A / Indonésia / 5/05 ao ectodomínio de H3 A / Brisbane / 10/2007 cria uma hemaglutinina quimérica que se acumula a um nível semelhante ao H3A nativo / Brisbane / 10/2007

A produção de VLPs a partir da expressão da hemaglutinina quimérica H5 / B (construção n° 745) foi avaliada utilizando cromatografia de exclusão por tamanhos. Extratos de proteína concentrada de plantas infiltradas com AGL1 / 745 (1,5 mL) foram fracionados por cromatografia de exclusão por tamanho (SEC) em colunas Sephacryl™ S-

500 HR (GE Healthcare Bio-Science Corp., Piscataway, NJ, EUA). Como mostrado na Figura 16, a eluição de Dextran Azul (2 MDa) atingiu um pico na fração 8. Quando as proteínas de 200 µl de cada fração de eluição SEC foram concentradas (5 vezes) por precipitação com acetona e analisadas por Western blotting usando anti-B Flórida), anticorpos policlonais (Figura 16), a hemaglutinina quimérica foi encontrada principalmente na fração 7, indicando a incorporação de HA em estruturas de alto peso molecular. Sem querer ser limitado pela teoria, isso sugere que a proteína HA quimérica foi montada em uma superestrutura grande ou que foi fixada a uma estrutura de alto peso molecular. Os resultados obtidos indicam que o HA quimérico que consiste no ectodomínio de HA B / Florida / 4/2006 fundido aos subdomínios transmembrana e citosólico de H5 A / Indonésia / 5/05 se agrupa em partículas de alto peso molecular e que o perfil de eluição destas partículas de alto peso molecular é indistinguível àquelas das VLPs da gripe.

Exemplo 5: Co-expressão de hemaglutinina quimérica H5 / B (número de construção 747; compreendendo B / Flo HDC e SDC fundida com um TDC H5 / Indo) com Hsp70 e Hsp40 em combinação com modificação do péptido sinal.

A expressão de Hsp40 e Hsp70 em plantas e a co-expressão com hemaglutininas de gripe são descritas no pedido co-pendente PCT / CA2009 / 000032. A Hsp70 citosólica e a Hsp40 (número de construção R870) de origem vegetal podem também ser co-expressas com hemaglutininas quiméricas, para aumentar o seu nível de acumulação nas plantas. A co-expressão pode ser realizada por agro-filtração de plantas de *N. benthamiana* com uma suspensão bacteriana contendo uma mistura (proporção de 1:1:1) de AGL1 contendo a cassete para a expressão do HA quimérico desejado com AGL1 / R870 e AGL1 / 35SHcPro.

O nível de acumulação de hemaglutinina H5 / B quimérica (B / Flo HDC e SDC fundida com um H5 / Indo TDC) foi avaliado em co-expressão com HSP40 e HSP70 em plantas de *Nicotiana benthamiana*. As plantas foram infiltradas com AGL1 / 747, AGL1 / 747 + AGL1 / 443 (vetor vazio) ou

AGL1 / 747 + AGL1 / R870 (HSP40 / HSP70), e as folhas foram colhidas após um período de incubação de seis dias. Para determinar o nível de acumulação de HA quimérico H5 / B nas folhas agro-infiltradas, as proteínas foram primeiro extraídas do tecido foliar infiltrado e analisadas por Western blotting utilizando anticorpos policlonais anti-B (Florida). Uma banda única de aproximadamente 75 kDa (Figura 50), correspondente em tamanho à forma HA0 não-clivada da hemaglutinina da gripe, foi detectada em extratos de folhas de 3 plantas infiltradas com AGL1 / 747 + AGL1 / R870 enquanto as 3 plantas infiltradas com AGL1 / 747 + vetor de controle (443) não mostraram sinal (sob a condição de exposição usada) indicando que o H5 / B quimérico forma de hemaglutinina acumulada em alto nível quando co-expressa com as HSP40 e HSP 70 chaperones.

Na descrição, vários termos são usados extensivamente e são fornecidas definições para facilitar a compreensão de vários aspetos da invenção. Os intervalos numéricos incluem os números que definem o intervalo. Na descrição, a palavra "compreendendo" é usada como um termo aberto, substancialmente equivalente à frase "incluindo, mas não limitado a", e a palavra "compreende" tem um significado correspondente.

LISTAGEM DE SEQUÊNCIAS

<110> Medicago Inc. COUTURE, MANON DARGIS, MICHELE LAVOIE, PIERRE-OLIVIER VEZINA, LOUIS- PHILIPPE
DAOUST, MARC-ANDRE

<120> Chimeric Influenza Virus-Like Particles Comprising Hemagglutinin

45 <130> V82197WO

<140> not assigned

<141> not assigned

50 <150> 61/220,161

<151> 2009-06-24

<160> 113

55 <170> PatentIn version 3.4

<210> 1

<211> 32

<212> DNA
 <213> Artificial sequence

5 <220>
 <223> Synthesized Xmal-pPlas.c

<400> 1
 agttccccgg gctggtatat ttatatgtg tc 32

10 <210>2
 <211> 46
 <212> DNA
 <213> Artificial Sequence

15 <220>
 <223> Synthesized SacI-ATG-pPlas.r

<400> 2
 aatagagctc cattttctct caagatgatt aattaattaa ttagtc 46

20 <210> 3
 <211> 46
 <212> DNA
 <213> Artificial sequence

25 <220>
 <223> Synthesized SacI-PlasTer.c

30 <400> 3
 aatagagctc gtaaaatgc ttcttctct cctatttata atatgg 46

<210> 4
 <211> 48
 <212> DNA
 35 <213> Artificial sequence

<220>
 <223> Synthesized EcoRI-PlasTer.r

40 <400> 4
 ttacgaattc tccttctaa ttggtgtact atcatttacc aaagggga 48

<210> 5
 <211> 25
 45 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthesized Plasto-443C

50 <400> 5
 gtattagtaa ttagaatttg gtgtc 25

<210> 6
 55 <211> 44
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> SpHa(Ind)-Plasto-r

5 <400> 6
 gcaagaagaa gcactatfff ctccatttfc tctcaagatg atta 44

<210> 7
 <211> 45
 <212> DNA
 10 <213> Artificial sequence

<220>
 <223> Plasto-SpHa.c

15 <400> 7
 ttaatcatct tgagagaaaa tggagaaaat agtgcttctt ctgac 45

<210> 8
 <211> 38
 20 <212> DNA
 <213> Artificial sequence

<220>
 <223> HA(Ind)-Sac.r

25 <400> 8
 actttgagct cttaaatgca aattctgcat tgtaacga 38

30 <210> 9
 <211> 24
 <212> DNA
 <213> Artificial sequence

35 <220>
 <223> Synthesized pBinPlus.2613c

<400> 9
 aggaagggaa gaaagcgaaa ggag 24

40 <210> 10
 <211> 56
 <212> DNA
 <213> Artificial sequence

45 <220>
 <223> Synthesized Mut-ATG115.r

50 <400> 10
 gtgccgaagc acgatctgac aacgttgaag atcgctcacg caagaaagac aagaga 56

<210> 11
 <211> 52
 <212> DNA
 55 <213> Artificial Sequence

<220>
 <223> Synthesized Mut-ATG161.c

gtgtcagat cgtgctcgg caccagtaca acgttttctt tcactgaagc ga 52

5 <210> 12
<211> 25
<212> DNA
<213> Artificial sequence

10 <220>
<223> Synthesized LC-C5-1.11Or

<400> 12
tctcctggag tcacagacag ggtgg 25

15 <210> 13
<211> 52
<212> DNA
<213> Artificial Sequence

20 <220>
<223> Synthesized E1 H1B-E1 H51.r

<400> 13
tcatagtcag cgaaatgccc tgggtaacag aggtcattgg ttggattggc ct 52

25 <210> 14
<211> 48
<212> DNA
<213> Artificial Sequence

30 <220>
<223> Synthesized E1 H5N-E1 H1B.c

<400> 14
35 atgacctctg ttaccaggg catttcgctg actatgagga actgaggg 48

<210> 15
<211> 48
<212> DNA
40 <213> Artificial Sequence

<220>
<223> Synthesized E2 H5I-RB H1B.r

45 <400> 15
ccaattcact ttccataatt cctgatccaa agcctctact cagtgcga 48

<210> 16
<211> 51
50 <212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized RB H1B-E2 H51.c

55 <400> 16
ggctttggat caggaattat gaaaagtgaa ttggaatatg gtaactgcaa c 51

<211> 48
 <212> DNA
 <213> Artificial Sequence
 5
 <220>
 <223> Synthesized E1 H1B-F(prime)1 H51.r
 <400> 17
 10 ggctattcct ttaataggc agagctccc gttgtgtgc tttccag 48
 <210> 18
 <211> 48
 <212> DNA
 15 <213> Artificial Sequence
 <220>
 <223> Synthesized F'1 H5N-E1 H1B.c
 20 <400> 18
 aacgggaagc tctgcctatt aaaaggaata gccccactac aattgggt 48
 <210> 19
 <211> 48
 25 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized F(prime)2 H5I-E2 H1B.r
 30 <400> 19
 ggagtttgac acttggtgtt gcatttatcc attggtgcat ttgagttg 48
 <210> 20
 35 <211> 48
 <212> DNA
 <213> Artificial Sequence
 <220>
 40 <223> Synthesized E2 H1B-F(prime)2 H51.c
 <400> 20
 aatgcaccaa tggataaatg caacaccaag tgcaaaactc caatgggg 48
 45 <210> 21
 <211> 48
 <212> DNA
 <213> Artificial Sequence
 50 <220>
 <223> Synthesized E1 H5I-E1 H1NC.r
 <400> 21
 55 tcttcatagt cggtgaaact cctgggtaa catgttccat tctcagga 48
 <210> 22
 <211> 48
 <212> DNA

5 <220>
 <223> Synthesized E1 H1NC-E1 H51.c

 <400> 22
 ctgagaatgg aacatgttac ccaggagatt tcaacgacta tgaagaac 48

 10 <210> 23
 <211>51
 <212> DNA
 <213> Artificial Sequence

 15 <220>
 <223> Synthesized E2 H1NC-RB H51.r

 <400> 23
 atttgagggtg atgattgctg agtccccctt cttgacaatt ttgtatgcat a 51

 20 <210> 24
 <211> 48
 <212> DNA
 <213> Artificial Sequence

 25 <220>
 <223> Synthesized RB H5I-E2 H1NC.c

 <400> 24
 gtcaagaaag gggactcagc aatcatcacc tcaaatgcac caatggat 48
 30
 <210> 25
 <211> 43
 <212> DNA
 <213> Artificial Sequence
 35
 <220>
 <223> Synthesized HA-Sacl.r

 <400> 25
 40 ttaacttaga gctcttagat gcatattcta cactgcaaag ace 43

 <210> 26
 <211> 45
 <212> DNA
 45 <213> Artificial sequence

 <220>
 <223> Synthesized Apal-H1B.c

 50 <400> 26
 tgtcgggccc atgaaagtaa aactactggt cctgttatgc acatt 45

 <210> 27
 <211> 46
 55 <212> DNA
 <213> Artificial Sequence

<220>

<223> Synthesized Stul-H1B.r

 <400> 27
 aaataggcct ttagatgcat attctacact gtaaagaccc attgga 46
 5

 <210> 28
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 10

 <220>
 <223> Synthesized SpPDI-H18.c

 <400> 28
 15 ttctcagatc ttctgctgaca caatatgtat aggctacat gctaacaac 49

 <210> 29
 <211> 47
 <212> DNA
 20 <213> Artificial Sequence

 <220>
 <223> Synthesized SacI-H1B.r
 25
 <400> 29
 ctagagctc ttagatgcat attctacact gtaaagaccc attggaa 47

 <210> 30
 <211> 45
 30 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Synthesized Apal-SpPDI.c
 35
 <400> 30
 ttgtcgggcc catggcgaaa aacgttgcga ttttcggcct attgt 45

 <210> 31
 40 <211> 39
 <212> DNA
 <213> Artificial Sequence

 <220>
 45 <223> Synthesized Apal-H5 (A-Indo).1c

 <400> 31
 tgtcgggccc atggagaaaa tagtgcttct tcttgcaat 39
 50
 <210> 32
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 55
 <220>
 <223> Synthesized H5 (A-Indo)-Stul.1707r

<400> 32

aaataggcct ttaaagcaa attctgcatt gtaacga 37
 <210> 33
 <211> 45
 5 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthesized H3B-SpPDI.r
 10 <400> 33
 tgtcatttc gggaagtttt tgagcgaaga tctgagaagg aacca 45
 <210> 34
 15 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220>
 20 <223> Synthesized SpPDI-H3B.c
 <400> 34
 tctcagatct tcgctcaaaa acttcccga aatgacaaca gcacg 45
 25 <210> 35
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 30 <220>
 <223> Synthesized Stul-H3B.r
 <400> 35
 35 aaaataggcc tcaaatgca aatgtgcac ctaatgtgc cttt 44
 <210> 36
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 40 <220>
 <223> Synthesized TmD H5I-H3B.r
 <400> 36
 45 atttggaag ttctattga cttcagctca acgccctga tctgg 45
 <210> 37
 <211> 48
 <212> DNA
 50 <213> Artificial Sequence
 <220>
 <223> Synthesized H3B-TmD H51.c
 55 <400> 37
 tgagctgaag tcaataggaa ctaccaaact actgtcaatt tattcaac 48
 <210> 38

<211> 50
 <212> DNA
 <213> Artificial sequence

5 <220>
 <223> Synthesized HBF-SpPDI.r

<400> 38
 gttattccag tgcagattcg atcagogaag atctgagaag gaaccaacac 50

10 <210> 39
 <211> 50
 <212> DNA
 <213> Artificial Sequence

15 <220>
 <223> Synthesized SpPDI-HBF.c

<400> 39
 20 cagatctcg ctgatcgaat ctgcactgga ataacatctt caaactcacc 50

<210> 40
 <211> 46
 <212> DNA
 25 <213> Artificial Sequence

<220>
 <223> Synthesized Stul-HBF.r

30 <400> 40
 aaaataggcc ttatagaca gatggagcat gaaacgttgt ctctgg 46

<210> 41
 <211> 45
 35 <212> DNA
 <213> Artificial sequence

<220>
 <223> Synthesized TmD H5I-B Flo.r

40 <400> 41
 tgacagtatt tggtagttat ccaatccatc atcatttaaa gatgc 45

<210> 42
 45 <211> 50
 <212> DNA
 <213> Artificial Sequence

<220>
 50 <223> Synthesized B Flo-TmD H51.c

<400> 42
 ggattggata actaccaaact actgtcaatt tattcaacag tggcgagttc 50

55 <210> 43
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthesized Hsp-40Luz.1c

5 <400> 43
 atgtttgggc gcggaccaac 20

<210> 44
 <211> 31
 <212> DNA
 10 <213> Artificial Sequence

<220>
 <223> Synthesized Hsp40Luz-Sacl.1272r

15 <400> 44
 agctgagctc ctactgtga gcgcattgca c 31

<210> 45
 <211> 36
 20 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthesized Hsp40Luz-Plasto.r

25 <400> 45
 gttggtccgc gcccaaacat ttctctcaa gatgat 36

<210> 46
 30 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 35 <223> Synthesized Hsp70Ara.1c

<400> 46
 atgtcgggta aaggagaagg a 21

40 <210> 47
 <211> 33
 <212> DNA
 <213> Artificial Sequence

45 <220>
 <223> Synthesized Hsp70Ara-Sacl.1956r

<400> 47
 50 agctgagctc ttagtcgacc tcctcgatct tag 33

<210> 48
 <211> 37
 <212> DNA
 55 <213> Artificial Sequence

<220>
 <223> Synthesized Hsp70Ara-Plasto.r

<400> 48
 tcctctcct ttacccgaca tttctctca agatgat 37

5
 <210> 49
 <211> 34
 <212> DNA
 <213> Artificial Sequence

10
 <220>
 <223> Synthesized supP19-plasto.r

<400> 49
 ccttgtag ctggtccat tttctctcaa gatg 34

15
 <210> 50
 <211> 20
 <212> DNA
 <213> Artificial Sequence

20
 <220>
 <223> Synthesized supP19-1c

25
 <400> 50
 atggaacgag ctatacaagg 20

30
 <210> 51
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthesized SupP19-Sacl.r

35
 <400> 51
 agtcgagctc ttactcgctt tcttttcga ag 32

40
 <210> 52
 <211> 1719
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthesized B-Plasto-Native SP-H5 A/Indonesia/5/05

45
 <400> 52

50

55

	aagcttatgg agaaaatagt gcttcttctt gcaatagtca gtcttggttaa aagtgatcag	60
	atttgcattg gttaccatgc aaacaattca acagagcagg ttgacacaat catggaaaag	120
5	aacgttactg ttacacatgc ccaagacata ctggaaaaga cacacaacgg gaagctctgc	180
	gatctagatg gagtgaagcc tctaatttta agagattgta gtgtagctgg atggctcctc	240
	gggaacccaa tgtgtgacga attcatcaat gtaccggaat ggtcttacat agtggagaag	300
10		
	gccaatccaa ccaatgacct ctgttaccba gggagtttca acgactatga agaactgaaa	360
	cacctattga gcagaataaa ccattttgag aaaattcaaa tcaccccaaa aagttcttgg	420
15	tccgatcatg aagcctcatc aggagttagc tcagcatgtc catacctggg aagtcctctc	480
	ttttttagaa atgtggtatg gcttatcaaa aagaacagta catacccaac aataaagaaa	540
	agctacaata ataccaacca agaggatctt ttggtactgt ggggaattca ccacctaata	600
20		
	gatgcggcag agcagacaag gctatatcaa aaccaacca cctatatattc cattgggaca	660
	tcaaacactaa accagagatt ggtacaaaa atagctacta gatccaaagt aaacgggcaa	720
25	agtgaagga tggagttctt ctggacaatt ttaaaccta atgatgcaat caacttcgag	780
	agtaatggaa atttcattgc tccagaatat gcatacaaaa ttgtcaagaa aggggactca	840
	gcaattatga aaagtgaatt ggaatatggt aactgcaaca ccaagtgtca aactccaatg	900
30		
	ggggcgataa actctagtat gccattccac aacatacacc ctctcaccat cggggaatgc	960
	cccaaatatg tgaaatcaaa cagattagtc cttgcaacag ggctcagaaa tagccctcaa	1020
35	agagagagca gaagaaaaaa gagaggacta ttggagcta tagcaggttt tatagaggga	1080
	ggatggcagg gaatggtaga tggttggtat gggtagcacc atagcaatga gcaggggagt	1140
	gggtacgctg cagacaaaaga atccactcaa aaggcaatag atggagtcac caataaggtc	1200
40		
	aactcaatca ttgacaaaat gaacactcag tttgaggccg ttggaaggga atttaataac	1260
	ttagaaagga gaatagagaa tttaaacaag aagatggaag acgggtttct agatgtctgg	1320
	acttataatg ccgaacttct ggttctcatg gaaaatgaga gaactctaga ctttcatgac	1380
45		
	tcaaatgtta agaacctcta cgacaaggtc cgactacagc ttagggataa tgcaaaggag	1440
	ctgggtaacg gttgtttcga gttctatcac aaatgtgata atgaatgtat ggaaagtata	1500
50		
	agaaacggaa cgtacaacta tccgcagtat tcagaagaag caagattaaa aagagaggaa	1560
	ataagtgggg taaaattgga atcaatagga acttaccaaa tactgtcaat ttattcaaca	1620

55

gtggcgagtt ccctagcact ggcaatcatg atggetgggc tatctttatg gatgtgctcc 1680

aatggatcgt tacaatgcag aatttgatt taagagctc 1719

<210> 53

<211> 3194

<212> DNA

<213> Artificial Sequence

5 <220>

<223> Synthesized 660 expression cassette from HindIII to Eco-R1

<400> 53

10 aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagttagca agtgtgtaca 120
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
15 taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa 240
catttgagaa aattttgttg ttctctcttt tcattggtca aaaacaatag agagagaaaa 300
20 aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360

25

30

35

40

45

50

	acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt	420
	aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg	480
5	caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta	540
	dacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag	600
	aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc	660
10	agttaactca tttttatatt tcatagatca aataagagaa ataacgggtat attaatccct	720
	ccaaaaaaaa aaaacgggtat atttactaaa aaatctaagc cacgtaggag gataacagga	780
15	tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataaccca	840
	ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt ctccacaca	900
	tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt	960
20	tgtgagtcta cactttgatt cccttcaaac acatacaaaag agaagagact aattaattaa	1020
	ttaatcatct tgagagaaaa tggagaaaat agtgcttctt cttgcaatag tcagtcttgt	1080
	taaaagtgat cagatttgca ttggttacca tgcaaacaat tcaacagagc aggttgacac	1140
25	aatcatggaa aagaacgta ctgttacaca tgcccaagac atactggaaa agacacacaa	1200
	cgggaagctc tgcgatctag atggagtga gctcttaatt ttaagagatt gtagttagc	1260
30	tggatggctc ctcggaacc caatgtgtga cgaattcatc aatgtaccgg aatggcttta	1320
	catagtggag aaggccaatc caaccaatga cctctgttac ccaggaggatt tcaacgacta	1380
	tgaagaactg aaacacctat tgagcagaat aaaccatttt gagaaaattc aatcatccc	1440
35	caaaagttct tggccgatc atgaagcctc atcaggagtt agctcagcat gtccatacct	1500
	gggaagtccc tccttttt:ta gaaatgtgggt atggcttata aaaaagaaca gtacataccc	1560
	aacaataaag aaaagctaca ataataccaa ccaagaggat cttttggtac tgtggggaat	1620
40	tcaccatcct aatgatgcgg cagagcagac aaggctatat caaaacccaa ccacctatat	1680
	ttccattggg acatcaacac taaaccagag attggtacca aaaatagcta ctagatccaa	1740
45	agtaaacggg caaagtggaa ggatggagtt cttctggaca attttaaaac ctaatgatgc	1800
	aatcaacttc gagagtaatg gaaatttcat tgctccagaa tatgcataca aaattgtcaa	1860
	gaaaggggac tcagcaatta tgaaaagtga attggaatat ggtaactgca acaccaagtg	1920
50		

tcaaactcca atgggggcga taaactctag tatgccattc cacaacatac accctctcac	1980
catcggggaa tgccccaat atgtgaaatc aaacagatta gtccttgcaa cagggctcag	2040
aaatagccct caaagagaga gcagaagaaa aaagagagga ctatttggag ctatagcagg	2100
ttttatagag ggaggatggc agggaatggt agatggttgg tatgggtacc accatagcaa	2160

55

tgagcagggg agtgggtacg ctgcagacaa agaatccact caaaaggcaa tagatggagt	2220
caccaataag gtcaactcaa tcattgacaa aatgaacact cagtttgagg ccgttggaag	2280
5 ggaatttaat aacttagaaa ggagaataga gaatttaaac aagaagatgg aagacggggtt	2340

	tctagatgtc tggacttata atgccgaact tctggttctc atggaaaatg agagaactct	2400
	agactttcat gactcaaag ttaagaacct ctacgacaag gtccgactac agcttaggga	2460
10	taatgcaaag gagctgggta acggttgttt cgagttctat cacaaatgtg ataatgaatg	2520
	tatggaaagt ataagaaacg gaacgtacaa ctatccgcag tattcagaag aagcaagatt	2580
15	aaaaagagag gaaataagtg gggtaaaatt ggaatcaata ggaacttacc aaatactgtc	2640
	aattttattca acagtggcga gttccctagc actggcaatc atgatggctg gtctatcttt	2700
	atggatgtgc tccaatggat cgttacaatg cagaatttgc atttaagagc tctaagttaa	2760
20	aatgcttctt cgtctcctat ttataaatatg gtttggttatt gtttaattttg ttcttgtaga	2820
	agagcttaat taatcgttgt tgttatgaaa tactatttgt atgagatgaa ctggtgtaat	2880
25	gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca	2940
	tgaagacctg ccgcgtacaa ttgtcttata tttgaacaac taaaattgaa catcttttgc	3000
	cacaacttta taagtgggta atatagctca aatatatggt caagttcaat agattaataa	3060
30	tggaaatatc agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt	3120
	tatatcatcc cctttgataa atgatatgac accaattagg aaggagcatg ctcgaggcct	3180
	ggctggccga attc	3194

35

<210> 54

<211>1711

<212> DNA

40

<213> Artificial Sequence

<220>

<223> Synthesized Wild-type H1/NC coding sequence lacking TmD and Ctail

45

<400> 54

	atgaaagcaa aactactggt cctgttatgt acatttacag ctacatatgc agacacaata	60
	tgtataggct accatgccaa caactcaacc gacactgttg acacagtact tgagaagaat	120
5	gtgacagtga cacactctgt caacctactt gaggacagtc acaatggaaa actatgtcta	180
	ctaaaaggaa tagccccact acaattgggt aattgcagcg ttgccggatg gatcttagga	240

	aaccagaat ggaattact gattccaag gaatcatggt cctacattgt agaaacacca	300
10		
	aatcctgaga atggaacatg ttaccacaggg tatttcgccg actatgagga actgagggag	360
	caattgagtt cagtatcttc atttgagaga ttcgaaatat tccccaaaga aagctcatgg	420
15	cccaaccaca ccgtaaccgg agtatcagca tcatgctccc ataatgggaa aagcagtttt	480
	tacagaaatt tgctatggct gacggggaag aatggtttgt acccaaacct gagcaagtcc	540
20	tatgtaaaca acaaagagaa agaagtcctt gtactatggg gtgttcatca cccgcctaac	600
	atagggaacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca	660
	cattatagca gaagattcac ccagaaata gccaaaagac ccaaagtaag agatcaggaa	720
25		
	ggaagaatca actactactg gactctgctg gaacctgggg atacaataat atttgaggca	780
	aatggaaatc taatagcgcc atggtatgct tttgcactga gtagaggctt tggatcagga	840
	atcatcacct caaatgcacc aatggatgaa tgtgatgcga agtgtcaaac acctcaggga	900
30		
	gctataaaca gcagtcttcc tttccagaat gtacaccag tcacaatagg agagtgtcca	960
	aagtatgtca ggagtgcaaa attaaggatg gttacaggac taaggaaacat cccatccatt	1020
35	caatccagag gtttgtttgg agccattgcc ggtttcattg aaggggggtg gactggaatg	1080
	gtagatgggt ggtatggtta tcatcatcag aatgagcaag gatctggcta tgctgcagat	1140
	caaaaaagta cacaaaatgc cattaacggg attacaaaca aggtgaattc tgtaattgag	1200
40		
	aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaattgga aagaaggatg	1260
	gaaaacttaa ataaaaaagt tgatgatggg tttctagaca tttggacata taatgcagaa	1320
45	ttgttggttc tactggaaaa tgaaaggact ttggatttcc atgactccaa tgtgaagaat	1380
	ctgtatgaga aagtaaaaag ccaattaaag aataatgcc aagaaatagg aaacgggtgt	1440
	tttgaattct atcacaagtg taacaatgaa tgcattgaga gtgtgaaaaa tggaacttat	1500
50		
	gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtgaaa	1560
	ttggaatcaa tgggagtcta tcagattctg gcgatctact caactgtcgc cagttccctg	1620
55	gttcttttgg tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttgcag	1680
	tgtagaatat gcattctgaga ccagaatttc a	1711

<210> 55

<211> 1556

<212> DNA

<213> Artificial Sequence

5

<220>

<223> Synthesized BgIII-H1 A/NewCaledonia/20/99-SacI/Stul.

<400> 55

10

agatcttcgc tgacacaata tgtataggct accatgccaa caactcaacc gacactgttg 60

acacagtact tgagaagaat gtgacagtga cacactctgt caacctactt gaggacagtc 120

15

acaatggaaa actatgtcta ctaaaaggaa tagccccact acaattgggt aattgcagcg 180

ttgccggatg gatcttagga aaccacagaat gcgaattact gatttccaag gaatcatggt 240

cctacattgt agaaacacca aatcctgaga atggaacatg ttaccacaggg tatttcgccg 300

20

25

30

35

40

45

50

	actatgagga actgagggag caattgagtt cagtatcttc atttgagaga ttogaaatat	360
	tccccaaaga aagctcatgg cccaaccaca ccgtaaccgg agtatcagca tcatgctccc	420
5	ataatgggaa aagcagtttt tacagaaatt tgctatggct gacggggaag aatggtttgt	480
	acccaaacct gagcaagtcc tatgtaaaca acaaagagaa agaagtcctt gtactatggg	540
	gtgttcatca cccgcctaac atagggaacc aaagggcact ctatcataca gaaaatgctt	600
10		
	atgtctctgt agtgtcttca cattatagca gaagattcac ccagaaata gccaaaagac	660
	ccaaagtaag agatcaggaa ggaagaatca actactactg gactctgctg gaacctgggg	720
15	atacaataat atttgaggca aatggaaatc taatagcgcc atggtatgct tttgcaactga	780
	gtagaggctt tggatcagga atcatcacct caaatgcacc aatggatgaa tgtgatgcga	840
	agtgtaaac acctcaggga gctataaaca gcagtcttcc tttccagaat gtacaccag	900
20		
	tcacaatagg agagtgtcca aagtatgtca ggagtgcaaa attaaggatg gttacaggac	960
	taaggaacat cccatccatt caatccagag gtttgtttgg agccattgcc ggtttcattg	1020
25	aaggggggtg gactggaatg gtagatgggt ggtatggta tcatcatcag aatgagcaag	1080
	gatctggcta tgctgcagat caaaaaagta caaaaaatgc cattaacggg attacaaaca	1140
	aggtcaattc tgtaattgag aaaatgaaca ctcaattcac agctgtgggc aaagagttca	1200
30	acaaattgga aagaaggatg gaaaacttaa ataaaaaagt tgatgatggg tttctagaca	1260
	tttggacata taatgcagaa ttgttggttc tactggaaaa tgaaaggact ttggatttcc	1320
	atgactccaa tgtgaagaat ctgtatgaga aagtaaaaag ccaattaaag aataatgcca	1380
35		
	aagaaatagg aaacgggtgt tttgagttct atcacaagtg taacaatgaa tgcattggaga	1440
	gtgtgaaaaa tggtagctat gactatccaa aatattccga agaatacaag ttaaacaggg	1500
40	agaaaattga tggagtgaat ttggaatcaa tgggagtata ctaagagctc aggcct	1556

<210> 56

<211> 219

<212> DNA

45 <213> Artificial Sequence

<220>

<223> Synthesized KpnI-H1 A/NewCaledonia/20/99 TmD+Ctail-SacI/Stul.

50 <400> 56

55

ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat	60
ggagtgaaat tggaatcaat gggagtatac cagattctgg cgatctactc aactgtcgcc	120
agttccctgg ttcttttggg ctccctgggg gcaatcagct tctggatgtg ttccaatggg	180
tctttgcagt gtagaatatg catctaagag ctcaggcct	219

<210> 57
<211> 1781
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Synthesized Signal peptide of alfalfa protein disulfide isomerase (POI) gene

<400> 57

10

15

20

25

30

35

40

45

50

	ccaaatcctt aacattcttt caacaccaac aatggcgaaa aacgttgcca ttttcggttt	60
	attgttttct cttcttctgt tggttccttc tcagatcttc gctgaggaat catcaactga	120
5		
	cgctaaggaa tttgttctta cattggataa cactaatttc catgacactg ttaagaagca	180
	cgatttcata gtcgttgaat tctacgcacc ttggtgtgga cactgtaaga agctagcccc	240
	agagtatgag aaggctgctt ctatcttgag cactcacgag ccaccagttg ttttggctaa	300
10		
	agttgatgcc aatgaggagc acaacaaaga cctcgcatcg gaaaatgatg ttaagggatt	360
	cccaaccatt aagattttta ggaatggtgg aaagaacatt caagaatata aagggtccccg	420
15		
	tgaagctgaa ggtattgttg agtatttgaa aaaacaaagt ggccctgcat ccacagaaat	480
	taaatctgct gatgatgcca ccgcttttgt tggtgacaac aaagttgtta ttgtcggagt	540
	tttcctaaa ttttctggtg aggagtacga taacttcatt gcattagcag agaagttgcg	600
20		
	ttctgactat gactttgctc acactttgaa tgccaaacac cttccaaagg gagactcatc	660
	agtgtctggg cctgtggtta ggtatttta gccatttgac gagctctttg ttgactcaaa	720
25		
	ggatttcaat gtagaagctc tagagaaatt cattgaagaa tccagtacct caattgtgac	780
	tgtcttcaac aatgagccta gcaatcacct ttttgttgct aaattcttta actctcccaa	840
	cgcaaaggct atgttggtca tcaactttac taccgaaggc gctgaatctt tcaaaacaaa	900
30		
	ataccatgaa gtggctgagc aatacaaaac acagggagtt agctttcttg ttggagatgt	960
	tgagtctagt caaggtgcct tccagtattt tggactgaag gaagaacaag tacctctaata	1020
	tattattcag cataatgatg gcaagaagtt tttcaaacc aatttggaac ttgatcaact	1080
35		
	cccaacttgg ttgaaggcat acaaggatgg caagggtgaa ccatttgtca agtctgaacc	1140
	tattcctgaa actaacaacg agcctgttaa agtgggtggt gggcaaactc ttgaggacgt	1200
40		
	tgttttcaag tctgggaaga atgttttgat agagttttat gtccttgggt gtggtcactg	1260
	caagcagttg gtcctaatct tggatgaagt tgctgtctca ttccaaagcg atgctgatgt	1320
	tgttattgca aaactggatg caactgccaa cgatatccca accgacacct ttgatgtcca	1380
45		
	aggctatcca accttgact tcaggtcagc aagtggaaaa ctatcacaat acgacgggtgg	1440
	taggacaaaag gaagacatca tagaattcat tgaaaagaac aaggataaaa ctggtgctgc	1500
50		
	tcatcaagaa gtagaacaac caaaagctgc tgctcagcca gaagcagaac aacaaaaaga	1560
	tgagctttga aaagttccgc ttggaggata tcggcacaca gtcattctgcg ggctttacaa	1620

5

ctcttttgta tctcagaatc agaagttagg aaatottagt gccaatctat ctatctttgc	1680
gtttcatttt atctttttgg tttactctaa tgtattactg aataatgtga gttttggcgg	1740
agtttagtac tggaactttt gtttctgtaa aaaaaaaaaa a	1781

<210> 58

<211> 1457

10

<212> DNA

<213> Artificial Sequence

<220>

15

<223> Synthesized PromPlasto-PDISP-Plasto 3(prime)UTR plasmid sequence

<400> 58

	ctggtatatt	tatatgttgt	caaataactc	aaaaaccata	aaagtttaag	ttagcaagtg	60
	tgtacatttt	tacttgaaca	aaaatattca	cctactactg	ttataaatca	ttattaaaca	120
5	ttagagtaaa	gaaatatgga	tgataagaac	aagagtagtg	atattttgac	aacaattttg	180
	ttgcaacatt	tgagaaaatt	ttgttgttct	ctcttttcat	tggtcaaaaa	caatagagag	240
	agaaaaagga	agagggagaa	taaaaacata	atgtgagtat	gagagagaaa	gttgtacaaa	300
10	agttgtacca	aaatagttgt	acaaatatca	ttgaggaatt	tgacaaaagc	tacacaaata	360
	agggttaatt	gctgtaaata	aataaggatg	acgcattaga	gagatgtacc	attagagaat	420
15	ttttggcaag	tcattaaaaa	gaaagaataa	attattttta	aaattaaaag	ttgagtcatt	480
	tgattaaaca	tgtgattatt	taatgaattg	atgaaagagt	tggtattaaag	ttgtattagt	540
	aattagaatt	tggtgtcaaa	tttaatttga	catttgatct	tttcctatat	attgccccat	600
20	agagtcagtt	aactcatttt	tatatttcat	agatcaaata	agagaaataa	cggtatatta	660
	atccctccaa	aaaaaaaaaa	cggtatat	actaaaaaat	ctaagccacg	taggaggata	720
	acaggatccc	cgtaggagga	taacatccaa	tccaaccaat	cacaacaatc	ctgatgagat	780
25	aaccctcttt	aagcccacgc	atctgtggca	catctacatt	atctaaatca	cacattcttc	840
	cacacatctg	agccacacaa	aaaccaatcc	acatctttat	caccatttct	ataaaaaatc	900
30	acactttgtg	agtctacact	ttgattccct	tcaaacacat	acaaagagaa	gagactaatt	960
	aattaattaa	tcatcttgag	agaaaatggc	gaaaaacggt	gcgattttcg	gcttattggt	1020
	ttctcttctt	gtgttggttc	cttctcagat	ctgagctcta	agttaaaatg	cttcttcgtc	1080
35	tcctatttat	aatatgggtt	gttattgtta	attttgttct	tgtagaagag	cttaattaat	1140
	cgttgttggt	atgaaatact	atltgtatga	gatgaactgg	tgtaatgtaa	ttcatttaca	1200
40	taagtggagt	cagaatcaga	atgtttcctc	cataactaac	tagacatgaa	gacctgccgc	1260

	gtacaattgt cttatatttg aacaactaaa attgaacatc ttttgccaca actttataag	1320
	tggttaatat agctcaaata tatggtcaag ttcaatagat taataatgga aatatcagtt	1380
45		
	atcgaaattc attaacaatc aacttaacgt tattaactac taattttata tcatcccctt	1440
50	tgataaatga tagtaca	1457

<210> 59

<211>3206

<212> DNA

55 **<213> Artificial Sequence**

<220>

<223> Synthesized 540 expression cassette from Hindi II

<400> 59

	aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta	60
5	tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagttagca agtgtgtaca	120
	tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag	180
	taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa	240
10	catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa	300
	aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt	360
15	acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggtt	420
	aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg	480
	caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta	540
20	aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag	600
	aatttgggtgt caaatttaat ttgacatttg atcttttctt atatattgcc ccatagagtc	660
25	agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatocct	720
	ccaaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga	780
	tccccgtagg aggataacat ccaatccaac caatcacac aatcctgatg agataaccca	840
30	ctttaagccc aogcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca	900
	tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt	960
	tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa	1020
35	ttaatcatct tgagagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct	1080
	tcttgtgttg gttccttctc agatcttcgc tgacacaata tgtataggct accatgccaa	1140
40	caactcaacc gacactgttg acacagtact tgagaagaat gtgacagtga cacactctgt	1200
	caacctactt gaggacagtc acaatggaaa actatgtcta ctaaaaggaa tagccccact	1260
	acaattgggt aattgcagcg ttgccggatg gatcttagga aaccagaat gcgaattact	1320
45	gatttccaag gaatcatggt cctacattgt agaaacacca aatcctgaga atggaacatg	1380
	ttaccagggg tatttcgccg actatgagga actgagggag caattgagtt cagtatcttc	1440
50		

	atttgagaga ttcgaaatat tccccaaaga aagctcatgg cccaaccaca ccgtaaccgg	1500
	agtatcagca tcatgtctccc ataatgggaa aagcagtttt tacagaaatt tgctatggct	1560
5	gacggggaag aatggtttgt acccaaacct gagcaagtcc tatgtaaaca acaaagagaa	1620
	agaagtcctt gtactatggg gtgttcatca cccgcctaac ataggaacc aaagggcact	1680
	ctatcataca gaaaatgctt atgtctctgt agtgtcttca cattatagca gaagattcac	1740
10	cccagaaata gccaaaagac ccaaagtaag agatcaggaa ggaagaatca actactactg	1800
	gactctgctg gaacctgggg atacaataat atttgaggca aatggaaatc taatagcgcc	1860
	atggtatgct tttgcactga gtagaggctt tggatcagga atcatcacct caaatgcacc	1920
15	aatggatgaa tgtgatgcga agtgtcaaac acctcaggga gctataaaca gcagtcttcc	1980
	tttccagaat gtacaccag tcacaatagg agagtgtcca aagtatgtca ggagtgcaaa	2040
20	attaaggatg gttacaggac taaggaacat cccatccatt caatccagag gtttgtttgg	2100
	agccattgcc ggtttcattq aaggggggtg gactggaatq gtagatgggt ggtatggtta	2160
	tcatcatcag aatgagcaag gatctggcta tgctgcagat caaaaaagta caaaaatgc	2220
25	cattaacggg attacaaaca aggtcaattc tgtaattgag aaaatgaaca ctcaattcac	2280
	agctgtgggc aaagagttca acaaattgga aagaaggatg gaaaacttaa ataaaaaagt	2340
	tgatgatggg tttctagaca tttggacata taatgcagaa ttgttggttc tactggaaaa	2400
30	tgaaaggact ttggatttcc atgactccaa tgtgaagaat ctgtatgaga aagtaaaaag	2460
	ccaattaaag aataatgcc aagaaatagg aaacgggtgt tttgagttct atcacaagtg	2520
35	taacaatgaa tgcattggaga gtgtgaaaaa tggtaacctat gactatccaa aatattccga	2580
	agaatcaaag ttaaacaggg agaaaattga tggagtgaat ttggaatcaa tgggagtata	2640
	ccagattctg gcgatcta:::t caactgtcgc cagttccctg gttcttttgg tctccctggg	2700
40	ggcaatcagc ttctggatgt gttccaatgg gtctttgcag tgtagaatat gcatctaaga	2760
	gctctaagtt aaaatgcttc ttctgtctcct atttataata tgggtttgta ttgttaattt	2820
	tgttcttgta gaagagctta attaatcggt gttgttatga aatactatct gtatgagatg	2880
45	aactgggtga atgtaattca ttacataag tggagtcaga atcagaatgt ttctccata	2940
	actaactag2 catgaagacc tgccgcgtac aattgtctta tatttgaaca actaaaattg	3000

```

50      aacatctttt gccacaactt tataagtggg taatatagct caaatatatg gtcaagttca 3060
      atagattaat aatggaaata tcagttatcg aaattcatta acaatcaact taacgttatt 3120
      aactactaat tttatatcat cccctttgat aaatgatagt acaccaatta ggaaggagca 3180

55      tgctcgaggc ctggctggcc gaattc                                     3206

```

55

<210> 60
<211> 1788
<212> DNA
<213> Artificial Sequence

5

<220>
<223> Synthesized fragment OralIII-Plasto(-84+ 1)-H1 A/Brisbane/59/07-Sacl.

<400> 60

10

15

20

25

30

35

40

45

50

	cactttgtga gtctacactt tgattccctt caaacacata caaagagaag agactaatta	60
	attaattaat catcttgaga gaaaatgaaa gtaaaactac tggtcctggt atgcacattt	120
5	acagctacat atgcagacac aatatgtata ggctaccatg ctaacaactc gaccgacact	180

	gttgacacag tacttgaaaa gaatgtgaca gtgacacact ctgtcaacct gcttgagaac	240
	agtcacaatg gaaaactatg totattaaaa ggaatagccc cactacaatt gggtaattgc	300
10	agcgttgccg ggtggatctt aggaaaccca gaatgcgaat tactgatttc caaggagtca	360
	tggtcctaca ttgtagaaaa accaaatcct gagaatggaa catgttaccg agggcatttc	420
15	gctgactatg aggaactgag ggagcaattg agttcagtat cttcatttga gaggttcgaa	480
	atattcccca aagaaagctc atggcccaac cacaccgtaa ccggagtgtc agcatcatgc	540
	tcccataatg gggaaagcag tttttacaga aatttgctat ggctgacggg gaagaatggg	600
20	ttgtacccaa acctgagcaa gtcctatgca aacaacaaag aaaaagaagt ccttgacta	660
	tggggtgttc atcacccgcc aaacataggt gacaaaaagg ccctctatca tacagaaaat	720
25	gcttatgtct ctgtagtgct ttcacattat agcagaaaaat tcaccccaga aatagccaaa	780
	agacccaaaag taagagatca agaaggaaga atcaattact actggactct gcttgaaccc	840
	ggggatacaa taatatttga ggcaaatgga aatctaatag cgccaagata tgctttcgca	900
30	ctgagtagag gctttggatc aggaatcatc aactcaaatg caccaatgga taaatgtgat	960
	gcgaagtgcc aaacacctca gggagctata aacagcagtc ttcctttcca gaacgtacac	1020
	ccagtcacaa taggagagtg tccaaagtat gtcaggagtg caaaattaag gatgggttaca	1080
35	ggactaagga acatcccatc cattcaatcc agaggtttgt ttggagccat tgccggtttc	1140
	attgaagggg ggtggactgg aatggtagat gggttggtatg gttatcatca tcagaatgag	1200
40	caaggatctg gctatgctgc agatcaaaaa agcacacaaa atgccattaa tgggattaca	1260
	aacaaggtca attctgtaat tgagaaaatg aacactcaat tcacagcagt gggcaaagag	1320
	ttcaacaaat tggaaagaag gatggaaaac ttgaataaaa aagttgatga tgggtttata	1380
45	gacatttggg catataatgc agaactgttg gttctactgg aaaatgaaag gactttggat	1440
	ttccatgact ccaatgtgaa gaatctgtat gagaaagtaa aaagccagtt aaagaataat	1500
50	gctaaagaaa taggaaatgg gtgttttgag ttctatcaca agtgtaacga tgaatgcatg	1560

agagtgtaa agaatggaac ttatgactat ccaaaatatt ccgaagaatc aaagttaaac 1620
 agggagaaaa ttgatggagt gaaattggaa tcaatgggag tctatcagat tctggcgatc 1680
 tactcaacag tcgccagttc tctggttctt ttggtctccc tgggggcaat cagcttctgg 1740
 atgtgttcca atgggtcttt acagtgtaga atatgcatct aagagctc 1788

<210>61
<211> 3185
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized 774 expression cassette from Hindi II
<400> 61

20

25

30

35

40

45

50

	aagcttgcta ggggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta	60
	tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagttagca agtgtgtaca	120
5	tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag	180
	taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa	240
	catttgagaa aattttgttg ttctctcttt tcattggtca aaaacaatag agagagaaaa	300
10	aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt	360
	acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataagggtt	420
15	aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg	480
	caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta	540
	aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag	600
20	aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc	660
	agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatecct	720
25	ccccgtagg aggataacat ccaatccaac caatcacac aatcctgatg agataaccca	780
	ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca	840
30	tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt	900
		960

	tgtgagtcta cacttttgatt cccttcaaac acatacaaag agaagagact aattaattaa	1020
	ttaatcatct tgagagaaaa tgaaagtaaa actactggtc ctggttatgca catttacagc	1080
35	tacatatgca gacacaatat gtataggcta ccatgctaac aactcgaccg aactgttga	1140
	cacagtactt gaaaagaatg tgacagtgc acactctgtc aacctgcttg agaacagtca	1200
40	caatggaaaa ctatgtctat taaaaggaat agccccacta caattgggta attgcagcgt	1260
	tgccgggtgg atcttaggaa acccagaatg cgaattactg atttccaagg agtcatggtc	1320
	ctacattgta gaaaaaccaa atcctgagaa tggaacatgt taccagggc atttcgctga	1380
	ctatgaggaa ctgagggagc aattgagttc agtatcttca tttgagaggt tcgaaatatt	1440
5	cccaaagaa agctcatggc ccaaccacac cgtaaccgga gtgtcagcat catgetceca	1500
	taatggggaa agcagttttt acagaaatct gctatggctg acggggaaga atggtttgta	1560
	cccaaacctg agcaagtcct atgcaaacia caaagaaaaa gaagtccttg tactatgggg	1620
10	tgttcatcac ccgcaaaca taggtgacca aaaggccctc tatcatacag aaaatgctta	1680
	tgtctctgta gtgtcttcac attatagcag aaaattcacc ccagaaatag ccaaagacc	1740
15	caaagtaaga gatcaagaag gaagaatcaa ttactactgg actctgcttg aaccgggga	1800
	tacaataata tttgaggcaa atggaaatct aatagcgcca agatatgctt tcgcactgag	1860
	tagaggcttt ggatcaggaa tcatcaactc aaatgcacca atggataaat gtgatgcgaa	1920
20	gtgccaaaca cctcaggag ctataaacag cagtcttctt tccagaacg tacaccagc	1980
	cacaatagga gagtgtccaa agtatgtcag gagtgcacaaa ttaaggatgg ttacaggact	2040
	aaggaaatc ccatccattc aatccagagg tttgtttgga gccattgccg gtttcattga	2100
25	aggggggtgg actggaatgg tagatggttg gtatggttat catcatcaga atgagcaagg	2160
	atctggctat gctgcagatc aaaaaagcac acaaatgcc attaatggga ttacaaacia	2220
30	ggtcaattct gtaattgaga aaatgaacac tcaattcaca gcagtgggca aagagttcaa	2280
	caaattggaa agaaggatgg aaaacttgaa taaaaaagtt gatgatgggt ttatagacat	2340
	ttggacatat aatgcagAAC tgttggttct actggaaaat gaaaggactt tggatttcca	2400
35	tgactccaat gtgaagaatc tgtatgagaa agtaaaaagc cagttaaaga ataatgctaa	2460
	agaaatagga aatgggtgtt ttgagttcta tcacaagtgt aacgatgaat gcatggagag	2520

	tgtaaagaat ggaacttatg actatccaaa atattccgaa gaatcaaagt taaacaggga	2580
40		
	gaaaattgat ggagtgaat tggaatcaat gggagtctat cagattctgg cgatctactc	2640
	aacagtcgcc agttctctgg ttcttttggg ctccctgggg gcaatcagct tctggatgtg	2700
45		
	ttccaatggg tctttacagt gtagaatatg catctaagag ctctaagta aaatgcttct	2760
	tcgtctccta ttataatat gggtttgttat tgtaattttt gttctttag aagagcttaa	2820
	ttaatcgttg ttgttatgaa atactatttg tatgagatga actgggtgtaa tgtaattcat	2880
50		
	ttacataagt ggagtcagaa tcagaatggt tcctccataa ctaactagac atgaagacct	2940
	gccgcgtaca attgtcttat atttgaacaa ctaaaattga acatcttttg ccacaacttt	3000
	ataagtgggt aatatagctc aaatatatgg tcaagttcaa tagattaata atggaaatat	3060
55		
	<u>cagttatcga aattcattaa caatcaactt aacgttatta actactaatt ttatatcatc</u>	<u>3120</u>

ccctttgata aatgatagta caccaattag gaaggagcat gctcgaggcc tggctggccg

aattc

5

55

<210> 62

<211> 2065

<212> DNA

<213> Artificial Sequence

10

<220>

<223> Synthesized Expression cassette number 828, from PacI to AseI

<400> 62

15

20

25

30

35

40

45

50

3180

3185

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc	180
	caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
10	acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
	tttgataaaa ggaacgtgg ggaacccga accaaacctt cttctaaact ctctctcatc	420
15	tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga	480
	tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt	600
20	cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
	tctgctgact ttcggcggtt gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
25	tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
	acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatatctg	840
	cccaaatttg togggcccac ggttttcaca cctcagatac ttggacttat gcttttttgg	900
30	atttcagcct ccagaggtga tattgtgcta actcagtctc cagccaccct gtctgtgact	960
	ccaggagata gtgtcagtct ttctgcagg gccagccaaa gtattagcaa caacctacac	1020
	tggtttcaac aaaaatcgca tgagtctcca aggtcttctc tcaagtatgc ttcccagtc	1080
35	atatctggga tcccctccag gttcagtggc agtggatctg ggacagattt cactctcagt	1140
	atcaacagtg tgaagactga agattttgga atgtttttct gtcaacagag taacagctgg	1200
40	cctctcacgt tcggtgatgg gacaaagctg gagctgaaac gggctgatgc tgcaccaact	1260
	gtatccatct toccaccatc cagtgcagcag ttaacatctg gaggtgcctc agtcgtgtgc	1320
	ttcttgaaca acttctaccc caaagacatc aatgtcaagt ggaagattga tggcagtgaa	1380
45	cgacaaaatg gcgtcctgaa cagttggact gatcaggaca gcaaagacag cacctacagc	1440

atgagcagca ccctcacgtt gaccaaggac gagtatgaac gacataacag ctatacctgt 1500
 gaggccactc acaagacatc aacttcaccc attgtcaaga gcttcaacag gaatgagtgt 1560
 5 tagaggccta ttttcttttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg 1620
 tgagcgggtt tctgtgctca gagtgtgttt attttatgta atttaatttc tttgtgagct 1680
 cctgttttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaataaa 1740
 10 aaaaaaaaaa aagaccggga attcgatatac aagcttatcg acctgcagat cgttcaaaca 1800
 tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcatg attatcatat 1860
 15 aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgatg acgttatatta 1920
 tgagatgggt ttttatgatt agagtcccg aattatacat ttaatacgcg atagaaaaca 1980
 aaatatagcg cgcaaacatg gataaattat cgcgcgcggg gtcattctatg ttactagatt 2040
 20 ctagagtctc aagcttcggc gcgcc 2065

<210> 63

<211>3194

25 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct number 690, from HindIII to EcoRI

30

<400> 63

35

40

45

aagcttgcta ggggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca 120
5
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa 240
10
catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
accaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt 420
15
aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg 480
caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag 600
20
aatttgggtg caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct 720
25
ccaaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataaccca 840
ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt ctccacaca 900
30

35

40

45

50

	tctgagccac	acaaaaacca	atccacatct	ttatcaccca	ttctataaaa	aatcacactt	960
	tgtgagtcta	cactttgatt	cccttcaaac	acatacaaag	agaagagact	aattaattaa	1020
5	ttaatcatct	tgagagaaaa	tgagaaaaat	agtgtcttct	cttgcaatag	tcagtcttgt	1080
	taaaagtgat	cagatttgca	ttggttacca	tgcaaaacaat	tcaacagagc	aggttgacac	1140
	aatcatggaa	aagaacgtta	ctgttacaca	tgcccaagac	atactggaaa	agacacacaa	1200
10	cggaagctc	tgcatctag	atggagttaa	gcctctaatt	ttaagagatt	gtagtgtagc	1260
	tggttggtc	ctcggaacc	caatgtgtga	cgaattcatc	aatgtaccgg	aatggtctta	1320
	catagtggag	aaggccaatc	caaccaatga	cctctgttac	ccagggcatt	tcgtgacta	1380
15	tgaggaactg	aggagcaat	tgagttcagt	atctt:cattt	gagaggttcg	aaatattccc	1440
	caaagaaagc	tcattggccca	accacaccgt	aaccggagtg	tcagcatcat	gtcccataa	1500
20	tggaagaaagc	agtttttaca	gaaatttgct	atggctgacg	gggaagaatg	gtttgtaccc	1560
	aaacctgagc	aagtcctatg	caaacaacaa	agaaaaagaa	gtccttgtag	tatgggtgt	1620
	tcattaccgg	ccaaacatag	gtgacaaaa	ggccctctat	catacagaaa	atgcttatgt	1680
25	ctctgtagt	tcttcacatt	atagcagaaa	attcacccca	gaaatagcca	aaagacccaa	1740
	agtaagagat	caagaaggaa	gaatcaatta	ctactggact	ctgcttgaac	ccggggatac	1800
	aataatattt	gaggcaaagc	gaaatctaag	agcgccaaga	tatgctttcg	cactgagtag	1860
30	aggctttgga	tcaggaatta	tgaaaagtga	attggaatat	ggtaactgca	acaccaagt	1920
	tcaaactcca	atgggggcca	taaactctag	tatgccattc	cacaacatac	accctctcac	1980
	catcggggaa	tgccccaaat	atgtgaaatc	aaacagatta	gtccttgcaa	cagggctcag	2040
35	aaatagccct	caaagagaga	gcagaagaaa	aaagagagga	ctatttggag	ctatagcagg	2100
	ttttatagag	ggaggatggc	agggaatggt	agatggttgg	tatgggtacc	accatagcaa	2160
40	tgagcagggg	agtgggtacg	ctgcagacaa	agaatccact	caaaaggcaa	tagatggagt	2220
	caccaataag	gtcaactcaa	tcattgacaa	aatgaacact	cagtttgagg	ccgttggaag	2280
	ggaatttaaat	aacttagaaa	ggagaataga	gaatttaaac	aagaagatgg	aagacgggtt	2340
45	tctagatgtc	tggacttata	atgccgaact	tctggttctc	atggaaaatg	agagaactct	2400
	agactttcat	gactcaaagc	ttaagaacct	ctacgacaag	gtccgactac	agcttaggga	2460
	taatgcaaag	gagctgggta	acgggttgtt	cgagttctat	cacaaatgtg	ataatgaatg	2520

50

tatggaaagt a.taagaaacg gaacgtacaa ctatccgcag tattcagaag aagcaagatt	2580
aaaaa.gagag gaaataagtg gggtaaaatt ggaatcaata ggaacttacc aaatactgtc	2640
aattttattca acagtggcga gttccctagc actggcaatc atgatggctg gtctatcttt	2700
atggatgtgc tccaatggat cgttacaatg cagaatttgc atttaagagc tctaagttaa	2760

55

	aatgcttcctt cgtctcctat ttataatatg gtttggttatt gtttaattttg ttcttgtaga	2820
	agagcttaat taatcggtgt tgttatgaaa tactatttgt atgagatgaa ctggtgtaat	2880
5	gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca	2940
	tgaagacctg ccgcgtacaa ttgtcttata tttgaacaac taaaattgaa catcttttgc	3000
	cacaacttta taagtgggta atatagctca aatatatggt caagttcaat agattaataa	3060
10	tggaaatata agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt	3120
	tatatcatcc cctttgataa atgatagtag accaattagg aaggagcatg ctcgaggcct	3180
15	ggctggccga attc	3194

<210> 64

<211>3194

<212> DNA

20 <213> Artificial Sequence

<220>

<223> Synthesized Construct number 691, from HindIII to EcoRI

25 <400> 64

30

35

40

45

	aagcttgcta ggggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta	60
	tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca	120
5	tttttacttg aacaaaaata ttcacctact actgtttataa atcattatta aacattagag	180
	taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa	240
	catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa	300
10		
	aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt	360
	accaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataagggtt	420
15	aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg	480
	caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta	540
	aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag	600
20		
	aatttggtgt caaatttaat ttgacatttg atcttttctt atatattgcc ccatagagtc	660
	agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct	720
	ccaaaaaaaaaaa aaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga	780
25		
	tccccgtagg aggataacat ccaatccaac caatcacac aatcctgatg agataaccca	840
	ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca	900
30	tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt	960
	tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa	1020
	ttaatcatct tgagagaaaa tggagaaaat agtgcttctt cttgcaatag tcagtcttgt	1080
35		
40		
45		
50		

	taaaagtgat cagatttgca ttggttacca tgcaaacaat tcaacagagc aggttgacac	1140
	aatcatggaa aagaacgtta ctgttacaca tgcccaagac atactggaaa agacacacaa	1200
5	cggaagci:c tgcctattaa aaggaatagc ccactacaa ttgggtaatt gcagcgttg	1260
	cggttgatc ttaggaaacc cagaatgcga attactgatt tocaaggagt catggtccta	1320
	cattgtagaa aaaccaaato ctgagaatgg aacatgttac ccagggcatt tcgctgacta	1380
10	tgaggaactg agggagcaat tgagttcagt atcttcattt gagagggttcg aaatattccc	1440
	caaagaaagc tcatggccca accacaccgt aaccggagtg tcagcatcat gctcccataa	1500
15	tggggaaagc agtttttaca gaaatttgct atggctgacg gggaagaatg gtttgtaccc	1560
	aaacctgagc aagtcctatg caaacaacaa agaaaaagaa gtccttgtag tatggggtgt	1620
	tcatcacccg ccaaacatag gtgacaaaa ggccctctat catacagaaa atgcttatgt	1680
20	ctctgtagtg tcttcacatt atagcagaaa attcacccca gaaatagcca aaagacccaa	1740
	agtaagagat caagaaggaa gaatcaatta ctactggact ctgcttgaac ccggggatac	1800
	aataatattt gaggcaaagc gaaatctaag agcgccaaga tatgcttttcg cactgagtag	1860
25	aggctttgga tcaggaatca tcaactcaaa tgcaccaatg gataaatgca acaccaagt	1920
	tcaaactcca atgggggcca taaactctag tatgccattc cacaacatac accctctcac	1980
30	cu.tcggggaa tgccccaaat atgtgaaatc aaacagatta gtccttgcaa cagggtcag	2040
	aaatagccct caaagagaga gcagaagaaa aaagagagga ctatttggag ctatagcagg	2100
	ttttatagag ggaggtggc aggggaatgg ataggttgg tatgggtacc accatagcaa	2160
35	tgagcagggg agtgggtacg ctgcagacaa agaattccact caaaaggcaa tagatggagt	2220
	caccaataag gtcaactcaa tcattgacaa aatgaacact cagtttgagg ccgttggaag	2280
	ggaatttaat aacttagaaa ggagaataga gaatttaaac aagaagatgg aagacgggtt	2340
40	tctagatgtc tggacttata atgccgaact tctggttctc atggaaaatg agagaactct	2400
	agactttcat gactcaaatg ttaagaacct ctacgacaag gtccgactac agcttaggga	2460
45	taatgcaaag gagctgggta acggttggtt cgagttctat cacaaatgtg ataataatg	2520
	tatggaaagt ataagaaacg gaacgtacaa ctatccgcag tattcagaag aagcaagatt	2580
	aaaaagagag gaaataagtg gggtaaaatt ggaatcaata ggaacttacc aaatactgtc	2640

50

aatttattca acagtggcga gttccctagc actggcaatc atgatggctg gtctatcttt 2700

atggatgtgc tccaatggat cgttacaatg cagaatttgc atttaagagc tctaagttaa 2760

aatgcttctt cgtctcctat ttataatatg gtttgttatt gttaattttg ttcttgtaga 2820

55

agagcttaat taatcgttgt tgttatgaaa tactatttgt atgagatgaa ctgggtgtaat 2880

	gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca	2940
	tgaagacctg ccgcgtacaa ttgtcttata ttgaacaac taaaattgaa catcttttgc	3000
5	cacaacttta taagtgggta atatagctca aatatatggt caagttcaat agattaataa	3060
	tggaaatata agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt	3120
	tatatcatcc cctttgataa atgatagtag accaattagg aaggagcatg ctcgaggcct	3180
10	ggctggccga attc	3194

<210> 65

<211>3206

15 **<212> DNA**

<213> Artificial Sequence

<220>

<223> Synthesized Construct number 696, from HindIII to EcoRI

20

<400> 65

25

30

35

40

45

	aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta	60
	tatttatatg ttgtcaaata actcaaaaac cataaaagt tttaagtttagca agtgtgtaca	120
5	tttttacttg aacaaaaata ttcacctact actgtttataa atcattatta aacattagag	180
	taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa	240
	catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa	300
10		
	aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt	360
	acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataagggtt	420
15	aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg	480
	caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta	540
	aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag	600
20		
	aatttgggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc	660
	agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct	720
25	ccaaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga	780
	tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataaccca	840
	ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca	900
30	tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt	960

	tgtgagtcta cacttttgatt cccttcaaac acatacaaag agaagagact aattaattaa	1020
	ttaatcatct tgagagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgtttttctct	1080
35	tcttgtgttg gttccttctc agatcttcgc tgacacaata tgtataggct accatgccaa	1140
	caactcaacc gacactgttg acacagtact tgagaagaat gtgacagtga cacactctgt	1200
40	caacctactt gaggacagtc acaatggaaa actatgtcta ctaaaaggaa tagccccact	1260
	acaattgggt aattgcagcg ttgccggatg gatcttagga aaccagaat gogaattact	1320
5	gatttccaag gaatcatggt cctacattgt agaaacacca aatcctgaga atggaacatg	1380
	ttaccagggt agtttcaacg actatgaaga actgaaacac ctattgagca gaataaacca	1440
	ttttgagaaa attcaaatca tccccaaaag ttcttgggtc gatcatgaag cctcatcagg	1500
10	agtttagctca gcatgtccat acctgggaag tcctctcttt tttagaaatg tggtaggct	1560
	tatcaaaaag aacagtacat acccaacaat aaagaaaagc tacaataata ccaaccaaga	1620
	ggatcttttg gtactgtgqg gaattcacca tcctaattgat gcggcaqagc agacaaggct	1680
15	atatcaaaac ccaaccacct atatttccat tgggacatca aactaaacc agagattggt	1740
	acaaaaaata gctactagat ccaaagtaaa cgggcaaagt ggaaggatgg agttcttctg	1800
20	gacaatttta aaacctaatg atgcaatcaa cttcgagagt aatggaaatt tcattgctcc	1860
	agaatatgca taaaaattg tcaagaaagg ggactcagca atcatcacct caaatgcacc	1920
	aatggatgaa tgtgatgcga agtgtcaaac acctcaggga gctataaaca gcagtcttcc	1980
25	tttccagaat gtacaccag tcacaatagg agagtgtcca aagtatgtca ggagtgcaaa	2040
	attaaggatg gttacaggac taaggaacat cccatccatt caatccagag gtttgtttgg	2100
	agccattgcc ggtttcattg aaggggggtg gactggaatg gtagatgggt ggtatggta	2160
30	tcatcatcag aatgagcaag gatctggcta tgctgcagat caaaaaagta caaaaaatgc	2220
	cattaacggg attacaaaca aggtcaattc tgtaattgag aaaatgaaca ctcaattcac	2280
	agctgtgggc aaagagttca acaaattgga aagaaggatg gaaaacttaa ataaaaaagt	2340
35	tgatgatggg tttctagaca tttggacata taatgcagaa ttgttggttc tactggaaaa	2400
	tgaaaggact ttggatttcc atgactccaa tgtgaagaat ctgtatgaga aagtaaaaag	2460

40	ccaattaaag aataatgcca aagaaatagg aaacgggtgt tttgagttct atcacaagtg	2520
	taacaatgaa tgcattggaga gtgtgaaaaa tggtagctat gactatccaa aatattccga	2580
	agaatcaaag ttaaacaggg agaaaattga tggagtgaag ttggaatcaa tgggagtata	2640
45	ccagattctg gcgattctact caactgtcgc cagttccctg gttcttttgg tctccctggg	2700
	ggcaatcagc ttctggatgt gttccaatgg gtctttgcag tgtagaatat gcatctaaga	2760
	gctctaagtt aaaatgcttc ttctgtctct atttataata tggtttgta ttgttaattt	2820
50	tgttcttgta gaagagctta attaatcgtt gttgttatga aatactatct gtatgagatg	2880
	aactgggtga atgtaattca ttacataag tggagtcaga atcagaatgt ttccctccata	2940
	actaactaga catgaagacc tgccgcgtac aattgtctta tatttgaaca actaaaattg	3000
55	aacatctttt gccacaactt tataagtggg taatatagct caaatatatg gtcaagttca	3060

atagattaat aatggaaata tcagttatcg aaattcatta acaatcaact taacgttatt	3120
aactactaat tttatatcat cccctttgat aaatgatagt acaccaatta ggaaggagca	3180
tgctcgaggc ctggctggcc gaattc	3206

5

<210> 66

<211> 3058

10 <212> DNA

<213> Artificial Sequence

<220>

15 <223> Synthesized Construct 732, from Paci to AseI

<400> 66

20

25

30

35

40

45

aattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60
gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg 120
5 ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggcccaaa gatggacccc 180
caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240
attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300
10 acccttcctc tatataagga agttcatttc atttgagag gtattaaaat cttaataggt 360
tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc 420
15 tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480
tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg 540
tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600
20 cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat 660
tctgctgact ttccggcggg gcaatatctc tacttctgct tgacgaggta ttgttgccg 720
25 tacttcttct ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780
acagagtttt ccggtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840
cccaaatttg tcgggcccac gaaagtaaaa ctactgggtc tgttatgcac atttacagct 900
30 acatatgcag acacaatatg tataggctac catgctaaca actcgaccga cactgttgac 960
acagtacttg aaaagaatgt gacagtgaca cactctgtca acctgcttga gaacagtcac 1020
aatggaaaac tatgtctatt aaaaggaata gcccactac aattgggtaa ttgcagcggt 1080
35 gccgggtgga tcttaggaaa ccagaaatgc gaattactga tttccaagga gtcattgggtc 1140
tacattgtag aaaaaccaa tctgagaat ggaacatgtt acccagggca tttcgctgac 1200
40 tatgaggaac tgaggagca attgagttca gtatcttcat ttgagagggt cgaaatattc 1260
cccaaagaaa gtcattggcc caaccacacc gtaaccggag tgtcagcatc atgtcccat 1320
aatggggaaa gcagttttta cagaaatttg ctatggctga cggggaagaa tggtttgtac 1380
45

50

55

	ccaaacctga gcaagtccta tgcaaacac aaagaaaaag aagtccttgt actatggggt	1440
	gttcatcacc cgccaaacat aggtgaccaa aaggccctct atcatacaga aaatgcttat	1500
5	gtctctgtag tgtcttcaca ttatagcaga aaattcaccc cagaaatagc caaaagaccc	1560
	aaagtaagag atcaagaagg aagaatcaat tactactgga ctctgcttga acccggggat	1620
	acaataatat ttgaggcaaa tggaatcta atagcgccaa gatatgcttt cgcactgagt	1680
10	agaggctttg gatcaggaat catcaactca aatgcaccaa tggataaatg tgatgcgaag	1740
	tgccaaacac ctcaggagc tataaacagc agtcttcctt tccagaacgt acaccagtc	1800
15	acaataggag agtgtccaaa gtatgtcagg agtgcaaaat taaggatggt tacaggacta	1860
	aggaacatcc catccattca atccagaggt ttgtttggag ccattgccgg tttcattgaa	1920
	ggggggtgga ctggaatggt agatggttgg tatggttatt atcatcagaa tgagcaagga	1980
20	tctggctatg ctgcagatca aaaaagcaca caaatgcc a ttaatgggat tacaacaag	2040
	gtcaattctg taattgagaa aatgaacact caattcacag cagtgggcaa agagttcaac	2100
25	aaattggaaa gaaggatgga aaacttgaat aaaaaagttg atgatgggtt tatagacatt	2160
	tggacatata atgcagaact gttggttcta ctggaaaatg aaaggacttt ggatttccat	2220
	gactccaatg tgaagaatct gtatgagaaa gtaaaaagcc agttaagaa taatgctaaa	2280
30	gaaataggaa atgggtgttt tgagttctat cacaagtgt aacgatgaatg catggagagt	2340
	gtaaagaatg gaacttatga ctatccaaaa tattccgaag aatcaaagtt aaacaggag	2400
	aaaattgatg gagtgaaatt ggaatcaatg ggagtctatc agattctggc gatctactca	2460
35	acagtcgcca gttctctggt tcttttggtc tccctggggg caatcagctt ctggatgtgt	2520
	tccaatgggt ctttacagtg tagaatatgc atctaaaggc ctattttctt tagtttgaat	2580
40	ttactgttat tcggtgtgca tttctatggt tggtgagcgg ttttctgtgc tcagagtgtg	2640
	tttattttat gtaatttaaat ttctttgtga gtcctgttt agcaggtcgt cccttcagca	2700
	aggacacaaa aagattttta tttatttaaa aaaaaaaaaa aaaaagaccg ggaattcgat	2760

45

	atcaagctta	tgcacctgca	gatcggtcaa	acatttggca	ataaagtttc	ttaagattga	2820
	atcctgttgc	cgggtcttgcg	atgattatca	tataatttct	gttgaattac	gttaagcatg	2880
50	taataattaa	catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	2940
	cgcaattata	catttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	3000
	tatcgcgcg	ggtgtcatct	atgttactag	attctagagt	ctcaagcttc	ggcgcgcc	3058

55

<210> 67

<211> 1719

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized construct number 787

5

<400> 67

10

15

20

25

30

35

40

45

50

55

5

10

15

20

25

30

35

40

45

50

atggcgaaaa	acgttgcgat	tttcggctta	ttgttttctc	ttcttgtgtt	ggttccttct	60
cagatcttcg	ctgacacaat	atgtataggc	taccatgcta	acaactcgac	cgacactggt	120
gacacagtac	ttgaaaagaa	tgtgacagtg	acacactctg	tcaacctgct	tgagaacagt	180
cacaatggaa	aactatgtct	attaaaagga	atagccccac	tacaattggg	taattgcagc	240
gttgccgggt	ggatcttagg	aaaccagaa	tgcaattac	tgatttccaa	ggagtcatgg	300
tcttacattg	tagaaaaacc	aatcctgag	aatggaacat	gttaccagag	gcatttcgct	360
gactatgagg	aactgagggg	gcaattgagt	tcagtatctt	catttgagag	gttcgaaata	420
ttccccaaag	aaagctcatg	gcccaaccac	accgtaaccg	gagtgtcagc	atcatgctcc	480
cataatgggg	aaagcagttt	ttacagaaat	ttgctatggc	tgacggggaa	gaatggtttg	540
tacccaaacc	tgagcaagtc	ctatgcaaac	aacaaagaaa	aagaagtcct	tgtactatgg	600
ggtgttcac	acccgccaaa	cataggtgac	caaaaggccc	tctatcatac	agaaaatgct	660
tatgtctctg	tagtgtcttc	acattatagc	agaaaattca	cccagaaat	agccaaaaga	720
cccaaagtaa	gagatcaaga	aggaagaatc	aattactact	ggactctgct	tgaaccggg	780
gataacaata	tatttgaggc	aatggaaat	ctaatagcgc	caagatatgc	tttcgcactg	840
agtagaggct	ttggatcagg	aatcatcaac	tcaaatgcac	caatggataa	atgtgatgcg	900
aagtgccaaa	cacctcaggg	agctataaac	agcagtcttc	ctttccagaa	cgtacacca	960
gtcacaatag	gagagtgtcc	aaagtatgtc	aggagtgcaa	aattaaggat	ggttacagga	1020
ctaaggaaca	tcccatccat	tcaatccaga	ggtttgtttg	gagccattgc	cggtttcatt	1080
gaaggggggt	ggactggaat	ggtagatggt	tggtatggtt	atcatcatca	gaatgagcaa	1140
ggatctggct	atgctgcaga	tcaaaaagc	acacaaaatg	ccattaatgg	gattacaaac	1200
aaggtcaatt	ctgtaattga	gaaaatgaac	actcaattca	cagcagtggg	caaagagttc	1260
aacaaattgg	aaagaaggat	ggaaaacttg	aataaaaaag	ttgatgatgg	gtttatagac	1320
atttgacat	ataatgcaga	actgttggtt	ctactggaaa	atgaaaggac	tttgatttc	1380
catgactcca	atgtgaagaa	tctgtatgag	aaagtaaaaa	gccagttaaa	gaataatgct	1440
aaagaaatag	gaaatgggtg	ttttgagttc	tatcacaagt	gtaacgatga	atgcatggag	1500
agtgtaaaga	atggaactta	tgactatcca	aaatattccg	aagaatcaaa	gttaaacagg	1560

gagaaaattg atggagtgaa attggaatca atgggagtct atcagattct ggcgatctac 1620
tcaacagtcg ccagttctct ggttcttttg gtctccctgg gggcaatcag cttctggatg 1680
tgttccaatg ggtctttaca gtgtagaata tgcattctaa 1719

55

<210> 68

<211> 3079

<212> DNA

<213> Artificial Sequence

5

<220>

<223> Synthesized Construct number 733, from PacI to AseI

<400> 68

10

ttaattaaga attcgagctc caccgoggaa acctcctcgg attccattgc ccagctatct 60

gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg 120

15

ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc 180

caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240

attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300

20

acccttcctc tatataagga agttcatttc atttggagag gtattaaaaat cttaataggt 360

tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc 420

25

tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480

tcgtgcttog gcaccagtac aacgttttct ttactgaag cgaaatcaaa gatctctttg 540

tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600

30

cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat 660

tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgcttg 720

35

tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780

acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840

cccaaatttg tcgggcccac ggcgaaaaac gttgcgattt tcggcttatt gttttctctt 900

40

cttgtgttgg ttcttctca gatcttcgct gacacaatat gtataggcta ccatgctaac 960

aactcgaccg acactgttga cacagtactt gaaaagaatg tgacagtgac acactctgtc 1020

aacctgcttg agaacagtca caatggaaaa ctatgtctat taaaaggaaat agccccacta 1080

45

caattgggta attgcagcgt tgccgggtgg atcttaggaa acccagaatg cgaattactg 1140

	atttccaagg agtcatgggc ctacattgta gaaaaaccaa atcctgagaa tggaacatgt	1200
50	taccagggc atttcgctga ctatgaggaa ctgagggagc aattgagttc agtatcttca	1260
	tttgagaggt tcgaaatatt ccccaaagaa agctcatggc ccaaccacac cgtaaccgga	1320
	gtgtcagcat catgetceca taatggggaa agcagttttt acagaaattt gctatggctg	1380
55	acggggaaga atggtttgta cccaaacctg agcaagtcct atgcaaacaa caaagaaaaa	1440
	gaagtccttg tactatgggg tgttcatcac ccgcaaaca taggtgacca aaaggccctc	1500

	tatcatacag aaaatgctta tgtctctgta gtgtcttcac attatagcag aaaattcacc	1560
	ccagaaatag ccaaaagacc caaagtaaga gatcaagaag gaagaatcaa ttactactgg	1620
5	actctgcttg aaccgggga tacaataata tttgaggcaa atggaaatct aatagcgcca	1680
	agatatgctt tcgcactgag tagaggcttt ggatcaggaa tcatcaactc aaatgcacca	1740
	atggataaat gtgatgcaa gtgccaaaca cctcaggag ctataaacag cagtcttcct	1800
10	ttccagaacg tacaccagc cacaatagga gagtgtccaa agtatgtcag gagtgcaaaa	1860
	ttaaggatgg ttacaggact aaggaacatc ccatccattc aatccagagg tttgtttgga	1920
15	gccattgccg gtttcattga aggggggtgg actggaatgg tagatggttg gtatggttat	1980
	catcatcaga atgagcaagg atctggctat gctgcagatc aaaaaagcac acaaatgcc	2040
	attaatggga ttacaaacaa ggtcaattct gtaattgaga aaatgaacac tcaattcaca	2100
20	gcagtgggca aagagttcaa caaattggaa agaaggatgg aaaacttgaa taaaaagtt	2160
	gatgatgggt ttatagacat ttggacatat aatgcagaac tgttggttct actggaaaat	2220
25	gaaaggactt tggatttcca tgactccaat gtgaagaatc tgtatgagaa agtaaaaagc	2280
	cagttaaaga ataatgctaa agaaaatagga aatgggtgtt ttgagttcta tcacaagtgt	2340
	aacgatgaat gcatggagag tgtaaagaat ggaacttatg actatccaaa atattccgaa	2400
30	gaatcaaagt taaacaggga gaaaattgat ggagtgaat tggaatcaat gggagtctat	2460
	cagattctgg cgatctactc aacagtcgcc agttctctgg ttcttttgggt ctccctgggg	2520
	gcaatcagct tctggatgtg ttccaatggg tctttacagt gtagaatatg catctaaagg	2580
35	cctattttct ttagtttgaa tttactgtta ttcgggtgtgc atttctatgt ttggtgagcg	2640
	gttttctgtg ctacagagtgt gtttatttta tgtaatttaa tttctttgtg agctcctgtt	2700
40	tagcaggctg tccttcagc aaggacacaa aaagatttta attttattaa aaaaaaaaaa	2760
	aaaaaagacc gggaattcga tatcaagctt atcgacctgc agatcgttca aacatttggc	2820
	aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc atataatttc	2880
45		

	tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta tttatgagat	2940
	gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa aacaaaatat	3000
50	agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcacc tatgttacta gattotagag	3060
	tctcaagctt cggcgcgcc	3079

<210> 69

55 **<211>3067**

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct 734, from Paci to Asel

<400> 69

5

10

15

20

25

30

35

40

45

50

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc	180
	caccacgag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
10		
	acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
	tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc	420
15	tctcttaaag caaacttctc tcttgtcttt ctgctgtgag cgatcttcaa cgttgtcaga	480
	togtgcctcg gcaccagtao aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt	600
20	cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
	tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
	tacttcttct tctctcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
25		
	acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
	cccaaatttg tcgggcccac ggagaaaata gtgcttcttc ttgcaatagt cagtcttggt	900
30	aaaagtgatc agatttgcac tggttaccat gcaacaatt caacagagca ggttgacaca	960
	atcatggaaa agaacgttac tgttacacat gcccaagaca tactggaaaa gacacacaac	1020
	gggaagctct gcgatctaga tggagtgaag cctctaattt taagagattg tagtgtagct	1080
35	ggatggctcc tcgggaaccc aatgtgtgac gaattcatca atgtaccgga atggtcttac	1140
	atagtggaga aggccaatcc aaccaatgac ctctgttacc cagggcattt cgctgactat	1200
	gaggaactga gggagcaatt gaggttcagta tcttcatttg agaggttcga aatattcccc	1260
40		
	aaagaaagct catggcccaa ccacaccgta accggagtgt cagcatcatg ctcccataat	1320
	ggggaaagca gtttttacag aaatttgcta tggctgacgg ggaagaatgg tttgtacca	1380
45	aacctgagca agtcctatgc aaacaacaaa gaaaaagaag tccttgact atggggtgtt	1440
	catcacccgc caaacatagg tgacaaaaag gcctctatc atacagaaaa tgettatgtc	1500
	tctgtagtgt cttcacatta tagcagaaaa ttcaccccag aaatagccaa aagacccaaa	1560
50		
	gtaagaqatc aagaaggaag aatcaattac tactggactc tgcttgaacc cggggataca	1620

55 ataatatttg aggcaaatgg aaatctaata gcgccaagat atgctttcgc actgagtaga 1680
ggctttggat caggaattat gaaaagtga ttggaatatg gtaactgcaa caccaagtgt 1740
caaactccaa tgggggcat aaactctagt atgccattcc acaacataca ccctctcacc 1800

	atcggggaat gccccaaata tgtgaaatca aacagattag tccttgcaac agggctcaga	1860
	aatagccctc aaagagagag cagaagaaaa aagagaggac tatttgagc tatagcaggt	1920
5	tttatagagg gaggatggca gggaatggta gatggttggt atgggtacca ccatagcaat	1980
	gagcagggga gtgggtacgc tgcagacaaa gaatccactc aaaaggcaat agatggagtc	2040
	accaataagg tcaactcaat cattgacaaa atgaacactc agtttgaggc cgttggaagg	2100
10	gaatttaata acttagaaag gagaatagag aatttaaaca agaagatgga agacgggttt	2160
	ctagatgtct ggacttataa tgccgaactt ctggttctca tggaaaatga gagaactcta	2220
15	gactttcatg actcaaagt taagaacctc tacgacaagg tccgactaca gcttagggat	2280
	aatgcaaagg agctgggtaa cggttgtttc gagttctatc acaaagtga taatgaatgt	2340
	atggaaagta taagaaacgg aacgtacaac tatccgcagt attcagaaga agcaagatta	2400
20	aaaagagagg aaataagtgg ggtaaaattg gaatcaatag gaacttacca aatactgtca	2460
	atttattcaa cagtggcgag ttccctagca ctggcaatca tgatggctgg tctatcttta	2520
25	tggatgtgct ccaatggatc gttacaatgc agaatttgca tttaaaggcc tattttcttt	2580
	agtttgaaat tactgttatt cgggtgtgcat ttctatgttt ggtgagcggg tttctgtgct	2640
	cagagtgtgt ttattttatg taatttaatt tctttgtgag ctctgttta gcaggctgtc	2700
30	ccttcagcaa ggacacaaaa agattttaat tttattaaaa aaaaaaaaaa aaaagaccgg	2760
	gaattcgata tcaagcttat cgacctgcag atcgttcaaa catttggaac taaagtttct	2820
	taagattgaa tcctgttgcc ggtcttgcca tgattatcat ataatttctg ttgaattacg	2880
35	ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga	2940
	ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact	3000
40	aggataaatt atcgcgcgcg gtgtcatcta tgttactaga ttctagagtc tcaagcttcg	3060
	gcgcgcc	3067

<210> 70

45 <211> 1791

<212> DNA

<213> Artificial Sequence

<220>

50 <223> Synthesized fragment Oral II-Plasto(-B4+ 1)-H3A/Brisbane/10/07-SacI

<400> 70

55 cactttgtga gtctacactt tgattccctt caaacacata caagagaag agactaatta 60
attaattaat catcttgaga gaaaatgaag actatcattg ctttgagcta cattctatgt 120
ctggttttca ctcaaaaact tcccggaaat gacaacagca cggcaacgct gtgccttggg 180

	cacccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt	240
	actaatgcta ctgagctggt tcagagttcc tcaacagggtg aaatatgcga cagtcctcat	300
5		
	cagatccttg atggagaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt	360
	gatggcttcc aaaataagaa atgggacctt tttgttgaac gcagcaaagc ctacagcaac	420
10		
	tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagtgtc ctcatccggc	480
	acactggagt ttaacaatga aagtttcaat tggactggag tcactcaaaa cggaacaagc	540
	tctgcttgca taaggagatc taataacagt ttcttttagta gattgaattg gttgaccac	600
15		
	ttaaaattca aatacccagc attgaacgtg actatgccaa acaatgaaaa atttgacaaa	660
	ttgtacattt ggggggttca ccaccogggg acggacaatg accaaatctt cctgtatgct	720
	caagcatcag gaagaatcac agtctctacc aaaagaagcc aacaaactgt aatcccgaat	780
20		
	atcggatcta gaccagagt aaggaatatc ccacagagaa taagcatcta ttggacaata	840
	gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt	900
25		
	tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaa	960
	tgcaattctg aatgcatcac tccaaacgga agcattccca atgacaaacc attccaaaat	1020
	gtaaacagga tcacatacgg ggctgtccc agatatgtta agcaaaacac tctgaaattg	1080
30		
	gcaacaggga tgcgaaatgt accagagaaa caaactagag gcatatttgg cgcaatcgcg	1140
	ggtttcatag aaaatggttg ggagggaatg gtggatgggt ggtatggttt caggcatcaa	1200
35		
	aattctgagg gaataggaca agcagcagat ctcaaaagca ctcaagcagc aatcgatcaa	1260
	atcaatggga agctgaatag gttgatcggg aaaaccaacg agaaattcca tcagattgaa	1320
	aaagagttct cagaagtcga agggagaatc caggaccttg agaaatatgt tgaggacacc	1380
40		
	aaaatagatc tctggtcata caacgcggag cttcttggtg ccctggagaa ccaacataca	1440
	attgatctaa ctgactcaga aatgaacaaa ctgtttgaaa aaacaaagaa gcaactgagg	1500
45		
	gaaaatgctg aggatatggg caatggttgt ttcaaaatat accacaaatg tgacaatgcc	1560
	tgcataggat caatcagaaa tggaacttat gaccacgatg tatacagaga tgaagcatta	1620
	aacaaccggg tccagatcaa gggcgttgag ctgaagtcag gatacaaaga ttggatacta	1680
50		
	tggaatttcct ttgcatatc atgttttttg ctttgtgttg ctttgttggg gttcatcatg	1740

tgggcctgcc aaaaaggcaa cattaggtgc aacatttgca tttgagagct c

1791

55

<210> 71

<211> 3085

<212> DNA

<213> Artificial Sequence

<220>
<223> Synthesized Construct 736, from Paci to Asel

<400> 71

5

10

15

20

25

30

35

40

45

50

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcgtt gaagatgctt ctgccgacag tgggcccaaa gatggacccc	180
	caccacagag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
10		
	acctttctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
	tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc	420
	tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga	480
15		
	tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggg	600
20	cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
	tctgctgact ttccggcggg gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
	tacttcttct tcttcttctc tgctgattgg ttctataaga aatctagtat tttctttgaa	780
25		
	acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
	cccaaatttg tcggggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt	900
	cttgtgttgg ttcccttctca gatcttcgct caaaaacttc ccggaaatga caacagcacg	960
30		
	gcaacgctgt gccttgggca ccatgcagta ccaaacggaa cgatagtga aacaatcacg	1020
	aatgaccaaa ttgaagttac taatgctact gagctggttc agagttcctc aacagggtgaa	1080
35	atatgcgaca gtctcatca gatccttgat ggagaaaact gcacactaat agatgctcta	1140
	ttgggagacc ctcatgtgta tggcttccaa aataagaaat gggacctttt tgttgaacgc	1200
	agcaaagcct acagcaactg ttacccttat gatgtgccgg attatgcctc ccttaggtca	1260
40		
	ctagttgcct catccggcac actggagttt aacaatgaaa gtttcaattg gactggagtc	1320
	actcaaaacg gaacaagctc tgcttgcata aggagatcta ataacagttt ctttagtaga	1380
	ttgaattggg tgaccacatt aaaattcaaa taccagcat tgaacgtgac tatgccaaac	1440
45		
	aatgaaaaat ttgacaaatt gtacatttgg ggggttcacc acccgggtac ggacaatgac	1500
	caaactcttc tgtatgctca agcatcagga agaatacacag tctctaccaa aagaagccaa	1560

50	caaactgtaa tcccgaatat cggatctaga cccagagtaa ggaatatccc cagcagaata	1620
	agcatctatt ggacaatagt aaaaccggga gacatacttt tgattaacag cacaggggaat	1680
	ctaattgctc ctaggggtta cttcaaaata cgaagtggga aaagctcaat aatgagatca	1740
55	gatgcacca ttggcaaatg caattctgaa tgcatcactc caaacggaag cattccaat	1800

	gacaaacccat tccaaaatgt aaacaggatc acatacgggg cctgtcccag atatgttaag	1860
	caaaacactc tgaaattggc aacagggatg cgaaatgtac cagagaaaca aactagaggc	1920
5	atatttggcg caatcgcggg tttcatagaa aatggttggg agggaatggt ggatggttgg	1980
	tatggtttca ggcatacaaa ttctgaggga ataggacaag cagcagatct caaaagcact	2040
	caagcagcaa tcgatcaaat caatgggaag ctgaataggt tgatcgggaa aaccaacgag	2100
10	aaattccatc agattgaaaa agagttctca gaagtcgaag ggagaatcca ggaccttgag	2160
	aaatatgttg aggacaccaa aatagatctc tggtcataca acgcggagct tcttgttgcc	2220
15	ctggagaacc aacatacaat tgatctaact gactcagaaa tgaacaaact gtttgaaaaa	2280
	acaaagaagc aactgaggga aaatgctgag gatatgggca atggttgttt caaaatatac	2340
	cacaaatgtg acaatgcctg cataggatca atcagaaatg gaacttatga ccacgatgta	2400
20	tacagagatg aagcattaaa caaccggttc cagatcaagg gcgttgagct gaagtcagga	2460
	tacaaagatt ggatactatg gatttccttt gccatatcat gttttttgct ttgtgttgct	2520
25	ttgttggggg tcatcatgtg ggctgcca aaaggcaaca ttaggtgcaa catttgcatt	2580
	tgaaggccta ttttctttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg	2640
	tgagcggttt tctgtgctca gagtgtgttt attttatgta atttaatttc tttgtgagct	2700
30	cctgttttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa	2760
	aaaaaaaaaa aagaccggga attcgatatc aagcttatcg acctgcagat cgttcaaaaa	2820
35	tttggaata aagtttctta agattgaatc ctgttgccgg tcttgcatg attatcatat	2880
	aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta	2940
	tgagatgggt ttttatgatt agagtcctgc aattatacat ttaatacgcg atagaaaaca	3000
40	aaatatagcg cgcaaaactag gataaattat cgcgcgcggg gtcacttatg ttactagatt	3060
	ctagagtctc aagcttcggc gcgcc	3085

45 <210> 72

<211> 3088

<212> DNA

<213> Artificial Sequence

50

<220>

<223> Synthesized Construct 737, from Pac1 to AseI

<400> 72

55

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggcccaaa gatggacccc	180
	caccacgag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240

10

15

20

25

30

35

40

45

50

	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
	acccttcctc tatataagga agttcatttc atttgagag gtattaaaat cttaataggt	360
5	tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc	420
	tctcttaaag caaacttctc tcttgctttt cttgcgtgag cgatcttcaa cgttgtcaga	480
	tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
10		
	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggg	600
	cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
15	tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgccctg	720
	tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
	acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
20	cccaaatttg tcgggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt	900
	cttgtgttgg ttcttctca gatcttcgct caaaaacttc ccggaaatga caacagcacg	960
	gcaacgctgt gccttgggca ccatgcagta ccaaacggaa cgatagtga aacaatcacg	1020
25		
	aatgacaaaa ttgaagttac taatgctact gagctggttc agagttcctc aacagggtgaa	1080
	atatgcgaca gtctcatca gatccttgat ggagaaaact gcacactaat agatgctcta	1140
30	ttgggagacc ctcaagtgtga tggcttccaa aataagaaat gggacctttt tgttgaacgc	1200
	agcaaagcct acagcaactg ttacccttat gatgtgccg attatgcctc ccttaggtca	1260
	ctagttgcct catccggcac actggagttt aacaatgaaa gtttcaattg gactggagtc	1320
35		
	actcaaaacg gaacaagctc tgcttgcata aggagatcta ataacagttt ctttagtaga	1380
	ttgaattggg tgaccactt aaaattcaaa taccagcat tgaacgtgac tatgccaaac	1440
	aatgaaaaat ttgacaaatt gtacatttgg ggggttcacc acccggtac ggacaatgac	1500
40		
	caaatcttcc tgtatgctca agcatcagga agaatacag tctctacaa aagaagccaa	1560
	caaactgtaa tcccgatat cggatctaga ccagagtaa ggaatatccc cagcagaata	1620
45	agcatctatt ggacaatagt aaaaccggga gacatacttt tgattaacag cacagggaat	1680
	ctaattgctc ctaggggtta cttcaaaata cgaagtggga aaagctcaat aatgagatca	1740
	gatgcacca ttggcaaatg caattctgaa tgcactctc caaacggaag cattcccaat	1800

50

gacaaaccat tccaaaatgt aaacaggatc acatacgggg cctgtcccag atatgttaag 1860

caaaacactc tgaaattggc aacagggatg cgaaatgtac cagagaaaca aactagagqc 1920

atatattggc caatcgcggt tttcatagaa aatgggtggg agggaatggt ggatgggttg 1980

55

tatggtttca ggcatcaaaa ttctgaggga ataggacaag cagcagatct caaaagcact 2040

	caagcagcaa tcgatcaaat caatgggaag ctgaataggt tgatcgggaa aaccaacgag	2100
	aaattccatc agattgaaaa agagttctca gaagtcgaag ggagaatcca ggaccttgag	2160
5	aaatatgttg aggacaccaa aatagatctc tggtcataca acgcggagct tcttggtgcc	2220
	ctggagaacc aacatacaat tgatctaact gactcagaaa tgaacaaact gtttgaaaaa	2280
	acaaagaagc aactgagga aaatgctgag gatatgggca atggttggtt caaaatatac	2340
10	cacaaatgtg acaatgcctg cataggatca atcagaaatg gaacttatga ccacgatgta	2400
	tacagagatg aagcattaaa caaccgggtc cagatcaagg gcgttgagct gaagtcaata	2460
15	ggaacttacc aaatactgtc aatttattca acagtggcga gttccctagc actggcaatc	2520
	atgatggctg gtctatcttt atggatgtgc tccaatggat cgttacaatg cagaatttgc	2580
	atttaaaggc ctatcttctt tagtttgaat ttactgttat tcggtgtgca tttctatggt	2640
20	tggtgagcgg ttttctgtgc tcagagtgtg tttattttat gtaatttaat ttctttgtga	2700
	gctcctgttt agcaggtcgt cccttcagca aggacacaaa aagattttta ttttattaaa	2760
25	aaaaaaaaa aaaaagaccg ggaattcgat atcaagctta tcgacctgca gatcgttcaa	2820
	acatttggca ataaagtctt ttaagattga atcctgttgc cggctcttgcg atgattatca	2880
	tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat	2940
30	ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa	3000
	acaaaatata gcgcgcaaac taggataaat tatcgcgcg cgtgtcatct atgttactag	3060
35	attctagagt ctcaagcttc ggcgcgcc	3088

<210> 73

<211> 1845

<212> DNA

40 **<213> Artificial Sequence**

<220>

<223> Synthesized fragment Oral II-Plasto(-84+ 1)-HA B/Florida/4/05-SacI

45 **<400> 73**

	cactttgtga gtctacactt tgattccctt caaacacata caaagagaag agactaatta	60
	attaattaat catcttgaga gaaaatgaag gcaataattg tactactcat ggtagtaaca	120
5	tccaatgcag atcgaatctg cactggaata acatcttcaa actcacctca tgtgggtcaaa	180
	acagccactc aaggggaggt caatgtgact ggtgtgatac cactaacaac aacaccaaca	240
	aaatcttatt ttgcaaactc caaaggaaca aggaccagag ggaaactatg cccagactgt	300
10	ctcaactgca cagatctgga tgtggctttg ggcagaccaa tgtgtgtggg gaccacacct	360
	tcggcgaagg cttcaatact ccacgaagtc aaacctgtta catccgggtg ctttcctata	420
15	atgcacgaca gaacaaaaat caggcaacta cccaatcttc tcagaggata tgaaaatatc	480

20

25

30

35

40

45

50

	aggctatcaa cccaaaacgt catcgatgcg gaaaaggcac caggaggacc ctacagactt	540
	ggaacctcag gatcttgccc taacgctacc agtaagagcg gatttttcgc aacaatggct	600
5		
	tgggctgtcc caaaggacaa caacaaaaat gcaacgaacc cactaacagt agaagtacca	660
	tacatttgta cagaagggga agaccaaatac actgtttggg ggttccattc agataacaaa	720
10		
	acccaaatga agaacctcta tggagactca aatcctcaaa agttcacctc atctgctaata	780
	ggagtaacca cacactatgt ttctcagatt ggcagcttcc cagatcaaac agaagacgga	840
	ggactaccac aaagcggcag gattgtttgtt gattacatga tgcaaaaacc tgggaaaaca	900
15		
	ggaacaattg tctaccaaag aggtgttttg ttgcctcaaa aggtgtggtg cgcgagtggc	960
	aggagcaaag taataaaagg gtccttgccct ttaattggtg aagcagattg ccttcatgaa	1020
	aaatacgggtg gattaaacaa aagcaagcct tactacacag gagaacatgc aaaagccata	1080
20		
	ggaaattgcc caatatgggt gaaaacacct ttgaagctcg ccaatggaac caaatataga	1140
	cctcctgcaa aactattaaa ggaaaggggt ttcttcggag ctattgctgg tttcctagaa	1200
25		
	ggaggatggg aaggaatgat tgcaggctgg cacggataca catctcacgg agcacatgga	1260
	gtggcagtgg cggcggacct taagagtacg caagaagcta taaacaagat aacaaaaaat	1320
	ctcaattctt tgagtgagct agaagtaaag aatcttcaaa gactaagtgg tgccatggat	1380
30		
	gaactccaca acgaaatact cgagctggat gagaaagtgg atgatctcag agctgacact	1440
	ataagctcgc aaatagaact tgcagtcttg ctttccaacg aaggaataat aaacagtga	1500
35		
	gatgagcatc tattggcact tgagagaaaa ctaaagaaaa tgctgggtcc ctctgctgta	1560
	gagataggaa atggatgctt cgaaacccaaa cacaagtgca accagacctg cttagacagg	1620
	atagctgctg gcacctttaa tgcaggagaa ttttctctcc ccacttttga ttactgaac	1680
40		
	attactgctg catcttttaa tgatgatgga ttggataacc atactatact gctctattac	1740
	tcaactgctg cttctagttt ggctgtaaca ttgatgctag ctatttttat tgtttatatg	1800
45		
	gtctccagag acaacgtttc atgctccatc tgtctataag agctc	1845

<210> 74

<211> 3142

<212> DNA

50 <213> Artificial Sequence

 <220>

 <223> Synthesized Construct 739, from Paci to AseI

55 <400> 74

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggtcctaaa gatggacccc	180

10

15

20

25

30

35

40

45

50

	cacccacgag gagcatcgc:g gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
5	acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
	tttgataaaa gcgaacgtgg gg2aaccoga accaaaacct cttctaaact ctctctcatc	420
	tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga	480
10	tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcgggtgtgt	600
	cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
15		
	tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgcttg	720
	tacttcttct tcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
20	acagagtttt ccggtggttt tcgaacttgg agaaagattg tt2agcttct gtatattctg	840
	cccaaatttg tcgggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt	900
	cttgtgttgg ttccttctca gatcttcgct gatcgaatct gcaactggaat aacatcttca	960
25	aactcacctc atgttgtcaa aacagccact caaggggagg tcaatgtgac tgggtgtgata	1020
	ccactaacia caacaccaac aaaatcttat tttgcaaatc tcaaaggaac aaggaccaga	1080
	gggaaactat gccagactg tctcaactgc acagatctgg atgtggcttt gggcagacca	1140
30		
	atgtgtgtgg ggaccacacc ttcggcgaag gcttcaatac tccacgaagt caaacctgtt	1200
	acatccgggt gctttcctat aatgcacgac agaacaaaaa tcaggcaact acccaatctt	1260
	ctcagaggat atgaaaatat caggctatca acccaaaacg tcatcgatgc ggaaaaggca	1320
35		
	ccaggaggac cctacagact tggaacctca ggatcttgcc ctaacgctac cag-::aagagc	1380
	ggatttttct caacaatggc ttgggctgtc ccaaaggaca acaacaaaaa tgcaacgaac	1440
40	ccactaacag tagaagtacc atacatttgt acagaagggg aagaccaaat cactgttttg	1500
	gggttccatt cagataacia aacccaaatg aagaacctct atggagactc aaatcctcaa	1560
	aagttcacct catctgctaa tggagtaacc acacactatg tttctcagat tggcagcttc	1620
45	ccagatcaaa cagaagacgg aggactacca caaagcggca ggattgttgt tgattacatg	1680
	atgcaaaaac ctgggaaaac aggaacaatt gtctacaaa gaggtgtttt gttgcctcaa	1740
	aagggtgtgt gcgcgagtgg caggagcaaa gtaataaaag ggtccttgcc tttaatttgt	1800
50		
	gaagcagatt gccttcatga aaaatacggg ggattaaaca aaagcaagcc ttactacaca	1860

```
ggagaacatg caaaagccat aggaaattgc ccaatatggg tgaaaacacc tttgaagctc 1920
qccaatgqaa ccaaatatag acctcctgca aaactattaa aggaaagggg tttcttcgga 1980
gctattgctg gtttcc-::aga aggaggatgg gaaggaatga ttgcaggctg gcacggatac 2040
```

	acatctcacg gagcacatgg agtggcagtg gcggcggacc ttaagagtac gcaagaagct	2100
	ataaacaaga taacaaaaaa tctcaattct ttgagtgagc tagaagtaaa gaatcttcaa	2160
5	agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga tgagaaagtg	2220
	gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt gctttccaac	2280
10	gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa actaaagaaa	2340
	atgctggggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa acacaagtgc	2400
	aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga attttctctc	2460
15	cccacttttg attcactgaa cattactgct gcactcttaa atgatgatgg attggataac	2520
	catactatac tgctctatta ctcaactgct gcttctagtt tggctgtaac attgatgcta	2580
	gctattttta ttgtttatat ggtctccaga gacaacgttt catgctccat ctgtctataa	2640
20	aggcctatth tctttagtht gaatttactg ttattcgggtg tgcatttcta tgtttggtga	2700
	gcggthttct gtgtcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct	2760
25	gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaaa	2820
	aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt	2880
	ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt atcatataat	2940
30	ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttattttatga	3000
	gatgggtttt tatgattaga gtcccgaat tatacattta atacgcgata gaaaacaaaa	3060
35	tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta ctagattcta	3120
	gagtctcaag cttcggcgcg cc	3142

<210> 75

40 <211>3142

<212> DNA

<213> Artificial Sequence

<220>

45 <223> Construct 745, from Pacl to AseI

<400> 75

	ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
	gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg	120
5	ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtcccaa gatggacccc	180
	caccacgag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
	attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
10		
	acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
	tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcctc	420
15		
20		
25		
30		
35		
40		
45		
50		

	tctcttaaag caaactttct tcttgtcttt cttgogtgag cgatcttcaa cgttgtcaga	480
	tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
5	tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt	600
	cttgggaaaa gaaagcttgc tggaggtgc tgttcagccc catacattac ttgttacgat	660
	tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
10		
	tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
	acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
15	cccaaatttg tcgggcccac ggcgaaaaac gttgcgattt tcggcttatt gttttctctt	900
	cttgtgttgg ttcttctca gatcttcgct gatcgaatct gcactggaat aacatcttca	960
	aactcacctc atgtggtcaa aacagccact caaggggagg tcaatgtgac tgggtgtgata	1020
20	ccactaacia caacaccaac aaaatcttat ttgcaaata tcaaaggaac aaggaccaga	1080
	gggaaactat gccagactg tctcaactgc acagatctgg atgtggcttt gggcagacca	1140
	atgtgtgtgg ggaccacacc ttcggcgaag gcttcaatac tccacgaagt caaacctgtt	1200
25		
	acatccgggt gctttcctat aatgcacgac agaacaaaaa tcaggcaact acccaatctt	1260
	ctcagaggat atgaaaatat caggctatca acccaaaacg tcatcgatgc ggaaaaggca	1320
30	ccaggaggac cctacagact tggaacctca ggatcttgcc ctaacgctac cagtaagagc	1380
	ggatttttctg caacaatggc ttgggctgct ccaaaggaca acaacaaaaa tgcaacgaac	1440
	ccactaacag tagaagtacc atacatttgt acagaagggg aagaccaaata cactgtttgg	1500
35	gggttccatt cagataacia aacccaaatg aagaacctct atggagactc aaatcctcaa	1560
	aagttcacct catctgctaa tggagtaacc acacactatg tttctcagat tggcagcttc	1620
	ccagatcaaa cagaagacgg aggactacca caaagcggca ggattgttgt tgattacatg	1680
40		
	atgcaaaaaa ctgggaaaaa aggaacaatt gtctacaaaa gaggtgtttt gttgcctcaa	1740
	aaggtgtggt gcgcgagtgg caggagcaaa gtaataaaaag ggtccttgcc ttttaattggt	1800
45	gaagcagatt gccttcatga aaaatacggg ggattaaaca aaagcaagcc ttactacaca	1860
	ggagaacatg caaaagccat aggaaattgc ccaatatggg tgaaaacacc tttgaagctc	1920
	gccaatggaa ccaaatatag acctctgca aaactattaa aggaaagggg tttcttcgga	1980

50

gctattgctg gtttcctaga aggaggatgg gaaggaatga ttgcaggctg gcacggatac 2040

acatctcacg gagcacatgg agtggcagtg gcggcggacc ttaagagtac gcaagaagct 2100

ataaacaaga taacaaaaaa tctcaattct ttgagtgagc tagaagtaaa gaatcttcaa 2160

55

agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga tgagaaagtg 2220

	gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt gctttccaac	2280
	gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa actaaagaaa	2340
5	atgctggggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa acacaagtgc	2400
	aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga attttctctc	2460
	cccacttttg attcactgaa cattactgct gcactcttaa atgatgatgg attggataac	2520
10	taccaaatac tgtcaattta ttcaacagtg gcgagttccc tagcactggc aatcatgatg	2580
	gctggtctat ctttatggat gtgctccaat ggatcggtac aatgcagaat ttgcatttaa	2640
15	aggcctattt tctttagttt gaatttactg ttattcgggtg tgcatttcta tgtttggtga	2700
	gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct	2760
	gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaaa	2820
20	aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt	2880
	ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt atcatataat	2940
25	ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga	3000
	gatgggtttt tatgattaga gtcccgaat tatacattta atacgcgata gaaaacaaaa	3060
	tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta ctagattcta	3120
30	gagtctcaag cttcggcgcg cc	3142

<210> 76

<211>1272

35 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Msjl coding sequence

40

<400> 76

45

	atgtttgggc gcggaaccaac aaggaagagt gataacacca aatattacga tattcttgggt	60
	gtttcaaaaa gtgctagtga agatgaaatc aagaaagcct atagaaaggc agcgatgaag	120
5	aaccatccag ataaggggtgg ggatcctgag aagttcaagg agttgggcca agcatatgaa	180
	gtgttgagcg atcctgaaaa gaaagaactg tatgatcaat atggtgaaga tgcccttaaa	240
	gaaggaatgg ggggagggcg aggaagctca ttccataatc cgtttgatat ttccgaatca	300
10	ttttttggtg caggcttttg tgggtggtgg ccttcacgcg caagaagaca gaagcaagga	360
	gaagatgtgg tgcattctat aaagggttcc ttggaggatg tgtataacgg cactacaaag	420
15	aagctatcac ttctaggaa tgactgtgc tcaaatgta aagggaagg ttcaaaaagt	480
	ggaactgctg gaagggtgtt tggatgccag ggcacaggta tgaagattac cagaaggcaa	540
	attggactgg gcatgattca acaaatgcaa cacgtctgtc ctgactgcaa aggaacaggc	600
20		
	gaggtcatta gtgagagaga tagatgccct caatgcaagg gaaacaagat tactcaagaa	660
	aagaaggtgc tggaggtgca tgtggaaaag gggatgcagc agggtcacaa gattgtattc	720
25	gaaggacaag ctgatgaagc tcctgatata atcacaggag acatagtttt tgtcttgcaa	780
	gtaaagggac atccgaagtt tcggagggag cgtgatgacc tccacattga acacaatttg	840
	agcttaactg aggctctctg tggcttccag tttaatgtca cacatcttga tggaaggcaa	900
30		
	ctattggtca aatcgaaccc cggcgaagtc atcaagccag gtcaacataa agctataaat	960
	gatgagggaa tgccacaaca tggtaggccg ttcatgaagg gacgcctata catcaagttt	1020
35	agtgttgatt tcccggattc gggttttctt tcccgaagcc aaagcctgga attagaaaag	1080
	atattacctc aaaagacaag caagaacttg tcccgaagg aggtagatga ttgtgaggag	1140
	accaccctgc atgatgtcaa tattgcagag gagatgagtc gaaagaagca acaataccgt	1200
40		
	gaggcatatg atgacgatga tgatgaagat gatgagcact cgcagcctcg ggtgcaatgc	1260
	gctcaacagt ag	1272
45	<210> 77 <211> 4402 <212> DNA <213> Artificial Sequence	
50	<220>	

<223> Synthesized Construct number R850, from HindIII

<400> 77

55

	aagcttgc	atgcctgc	aggtcgact	ctctagagg	atccccgat	ctgtggtct	gtacatttc	60
	tgccgcct	ttgcattc	acttggcc	caaaagag	tagagaga	aggaagaga	gagcccaga	120
5								
	ttcaaga	agcgacct	tggtgcaa	gtgcactc	gaggtcaga	aaactgtat	catatgtg	180
	agagaaa	ggggaatt	tgaatgat	ggtaggtc	ctattact	ttaatttt	gat	240
	caataaa	tttctgt	atacttct	tatttaga	tcataaatt	gttcatta	agc	300
10								
	ataaatt	acgatca	atcaaat	atgcaaat	gttagcag	tacttggt	taattct	360
	cttggtt	ttcttga	catggcca	taagcatg	aaataaat	agtatcat	tgtgtttg	420
15								
	tcagactt	caagatgc	actctgtt	tggtgata	aaacagc	tttaatta	agcgcttat	480
	tatcatat	gattgtgt	tttgctc	agacttca	gagcatct	ctcttggt	ttcatata	540
	tgcccta	agaaata	aatagtc	atcatatg	attgtgtt	gttcagac		600
20								
	cctgttt	gggttaa	acagctt	aatgaagt	gcctataga	ataagtc		660
	ttttgt	acagtt	ttatttct	atgaaagt	agaaatag	tcaattgt	ttataag	720
25								
	tcctgg	agagctt	gtggaaa	taaccaga	aaagattg	gacacgtc	atgagctg	780
	cataaga	tctcccta	acagtc	ctcaaaa	agtggtt	atgtagat	aaattcaa	840
30								
35								
40								
45								
50								

	aatctaaaca gaccctaaat ccattatggt acctatcatt ttagcttatt ccattcttat	900
	taagaatgtc atgagataac ataatagataa cacattatct tgacacaaat gggcagatct	960
5	agcaatttaa ctctggagtc cttcaagact gctgttctta cgaagttcac gtccctgaat	1020
	catgttcctg tatggaagcc tgaaagacct caaattctaa aaggtggcga taaattgaag	1080
	gtttacaaaa tataccctgc gggcttgaca cagaggcaag ctctttatac cttccagttc	1140
10	aacggggatg ttgatttcag aagtcacttg gagagcaatc cttgtgccaa gtttgaagta	1200
	atcttttgtgt agcatatgtt gagctaccta caattttacat gatcacctag cattagctct	1260
	ttcacttaac tgagagaatg aagttttagg aatgagtatg accatggagt cggcatggct	1320
15	ttgtaatgcc taccctactt tggccaactc atcgggggatt tacattcaga aaatatacat	1380
	gacttcaacc atacttaaac ccctttttgt aagataactg aatgttcata tttaatgttg	1440
20	ggttgtagt tttttacttg attatatcca gacagttaca agttggacaa caagattgtg	1500
	ggctctgtact gttatttatt tatttttttt ttagcagaaa caccttatct tttgtttcgt	1560
	ttgaatgtag aatgaaaata aaagaaagaa aatataacat catcgggcgc gcttgtctaa	1620
25	tttcgggcag ttaggatcct ctccggtcac cggaaagttt cagtagaaga aacaaaacac	1680
	cgtgactaaa atgatactat tattttatct attgtgtttt tcttttttct accggaactt	1740
	tttagaacgg atcccaactc gttccggggc cgctacaact gaaacaaaag aagatatttt	1800
30	ctctctcttc agaaatgtaa gttttccttt acagataccc attcaccatt tgattcagat	1860
	gtggtgacta gagataaagc atactaattt gactcttgga aaccataaa gtttatgtta	1920
	tccgtgttct ggaccaatcc acttgggggc ataacctgtg tctatgtgtg gtttggtttc	1980
35	::::attctgatt tatg::::ggcga cttgtaattt aaaatctagg aggggcagac attgaacaat	2040
	ccaatattt taataactta tgcaagattt tttttattaa tgagatgatg tgtttgtgac	2100
40	tgagattgag tcatacatct cactaagaaa tggttccaag taccaaaacta tcatgaccca	2160
	gttgcaaaca tgacgttcgg gagtggtcac tttgatagtt caatttcac ttggcttctt	2220
	attcctttta taattctaatt tcttcttggtg taaactattt catgtattat tttctttta	2280
45	aatttacatg tcattttatt tgcctcacta actcaattt gcatataaca atgataagtg	2340
	atattttgac tcacaaaatt tacatcaaat ttcgacatcg tttattatgt tcattggatg	2400
	attaacaaat ataacaaact ttgcaactaa ttaaccacca actgaatata attaactata	2460
50	actgtgaaag tagttaactc atttttatat ttcatagatc aaataagaga aataacggta	2520

tattaatccc	tccaaaaaaa	aaaaacggta	tatttactaa	aaaatctaag	ccacgtagga	2580
ggataacagg	atccccgtag	gaggataaca	tccaatccaa	ccaatcacia	caatcctgat	2640
gagataaccc	actttaagcc	cacgcacatctg	tggcacatct	acattatcta	aatcacacat	2700

	tcttccacac atctgagcca cacaaaaacc aatccacatc tttatcaccc attctataaa	2760
	aatcacact ttgtgagtct acactttgat tcccttcaaa cacatacaaa gagaagagac	2820
5	taattaatta attaatcatc ttgagagaaa atgtttgggc gcggaccaac aaggaagagt	2880
	gataacacca aatattacga tattcttggg gtttcaaaaa gtgctagtga agatgaaatc	2940
	aagaaagcct atagaaaggc agcgatgaag aaccatccag ataaggggtg ggatcctgag	3000
10	aagttcaagg agttgggcca agcatatgaa gtgttgagcg atcctgaaaa gaaagaactg	3060
	tatgatcaat atgggtgaaga tgcccttaaa gaaggaatgg ggggaggcgc aggaagctca	3120
15	tttcataatc cgtttgatat tttcgaatca ttttttgggt caggctttgg tgggtggtggt	3180
	ccttcacgcg caagaagaca gaagcaagga gaagatgtgg tgcattctat aaaggtttcc	3240
	ttggaggatg tgtataacgg cactacaaag aagctatcac tttctaggaa tgcactgtgc	3300
20	tcaaaatgta aagggaaagg ttcaaaaagt ggaactgctg gaagggtgtt tggatgccag	3360
	ggcacaggta tgaagattac cagaaggcaa attggactgg gcatgattca acaaagcaa	3420
25	cacgtctgtc ctgactgcaa aggaacaggc gaggtcatta gtgagagaga tagatgccct	3480
	caatgcaagg gaaacaagat tactcaagaa aagaagggtg tggagggtgca tgtggaaaag	3540
	gggatgcagc aggggtcaca gattgtattc gaaggacaag ctgatgaagc tctgataca	3600
30	atcacaggag acatagtttt tgtcttgcaa gtaaaggac atccgaagtt tcggaggagg	3660
	cgtgatgacc tccacattga acacaatttg agcttaactg aggtctcttg tggcttccag	3720
35	tttaatgtca cacatcttga tggaaggcaa ctattgggtca aatcgaaccc cggcgaagtc	3780
	atcaagccag gtcaacataa agctataaat gatgagggaa tgccacaaca tggtaggccg	3840
	ttcatgaagg gacgcctata catcaagttt agtgttgatt tcccggattc gggttttctt	3900
40	tccccaagcc aaagcctgga attagaaaag atattacctc aaaagacaag caagaacttg	3960
	tcccaaaagg aggtagatga ttgtgaggag accaccctgc atgatgtcaa tattgcagag	4020
	gagatgagtc gaaagaagca acaataccgt gaggcataatg atgacgatga tgatgaagat	4080
45	gatgagcact cgcagcctcg ggtgcaatgc gctcaacagt aggagctcag ctcgaatttc	4140
	cccgatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt	4200

50	gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa	4260
	tgcatgacgt tatttatgag atgggttttt atgattagag tcccgaatt atacatttaa	4320
	tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca	4380
55	tctatgttac tagatogaat tc	4402

<210> 78

<211> 5086

<212> DNA

<213> Artificial Sequence

5

<220>

<223> Synthesized Construct number R860, from HindIII to EcoRI

<400> 78

10

15

20

25

30

35

40

45

50

	aagcttgc	at	gcctgcaggt	cgactctaga	ggatccccgg	gctgggtctgt	acattcatct	60
	tgccgccttt	gcattcactt	ggccacaaag	agtagagaga	aggaagagaa	gagcccagac	120	
5	ttcaagaagc	gaccttgcaa	gtgcactcga	gggtcagaaa	ctgtatatca	tatctatgtg	180	
	agagaaaggg	gaacatttga	gatggagtcc	atttacttga	ggtatactta	ttattttgat	240	
	caataaattt	gtatacttct	tatttagatc	aataaatttg	tcattaagct	ataatccaaa	300	
10								
	ataaattacg	atcaaatatg	caaagttag	ccagtacttg	tgttaaactt	gatggcatct	360	
	cttggtttct	ttggcaatca	catgcctaag	aaataaatag	tatcatatga	ttgtgtttgg	420	
15	tcagacttca	gagtcagatg	actctgtttg	gataaacagc	ttaattaagc	gcttatagaa	480	
	tatcatatga	ttgtgtttgg	tcagacttca	gagcatctct	tggtttctct	ggcaatcata	540	
	tgccaaagaa	ataaatagta	tcatatgatt	gtgtttggtc	agacttcaga	gtcagatgac	600	
20								
	cctgtttggg	taaacagctt	aattaagtgc	ttatagaata	agcgcttatc	atataagtgc	660	
	ttttgtacag	ttatttctat	gaaagtagaa	gaaatagtca	tattgtttta	atataagcta	720	
25	tcctggagag	cttgtggaaa	taaccagaaa	agaacttatg	gacacgtcat	gagctgttta	780	
	cataagatct	ccctaacagt	ctcaaaagtg	tttatgccag	tagataaatt	caaataagtc	840	
	aatctaaaca	gaccctaaat	ccattatggg	acctatcatt	ttagcttatt	ccatctttat	900	
30								
	taagaatgtc	atgagataac	ataatgataa	cacattattt	tgacacaaat	gggcagatct	960	
	agcaatttaa	ctctggagtc	cttcaagact	gctgttctta	cgaagttcac	gtccctgaat	1020	
	catgttcctg	tatggaagcc	tgaaagacct	caaattctaa	aaggtggcga	taaattgaag	1080	
35								
	gtttacaaaa	tataccctgc	gggcttgaca	cagaggcaag	ctctttatac	cttccagttc	1140	
	aacgggggatg	ttgatttcag	aagtcacttg	gagagcaatc	cttgtgccaa	gtttgaagta	1200	
40	atttttgtgt	agcatatgtt	gagctaccta	caattttacat	gatcacctag	cattagctct	1260	
	ttcacttaac	tgagagaatg	aagttttagg	aatgagtatg	accatggagt	cggcatggct	1320	
	ttgtaatgcc	taccctactt	tgGCCAACTC	atcggggatt	tacattcaga	aaatatacat	1380	
45								
	gacttcaacc	atacttaaac	ccctttttgt	aagataactg	aatgttcata	tttaatgttg	1440	
	ggttgtagtg	tttttacttg	attatatcca	gacagttaca	agttggacaa	caagattgtg	1500	
50	ggctctgtact	gttattttatt	tatttttttt	ttagcagaaa	caccttatct	tttgtttcgt	1560	
	ttgaatgtag	aatgaaaata	aaagaaagaa	aatataacat	catcggccgc	gcttgtctaa	1620	

	tttcgggcag ttaggatcct ctccgggtcac cggaaagttt cagtagaaga aacaaaacac	1680
	cgtgactaaa atgatactat tattttatatt attgtgtttt tottttttct accggaactt	1740
5	tttagaacgg atcccaactc gttccggggc cgctacaact gaaacaaaag aagatatattt	1800
	ctctctcttc agaaatgtaa gttttccttt acagataccc attcaccatt tgattcagat	1860
	gtggtgacta gagataaagc atactaatat gactcttgga aaccataaaa gtttatgtta	1920
10	tccgtgttct ggaccaatcc acttgggggc ataacctgtg tctatgtgtg gtttggttct	1980
	cattctgatt tatgcggcga ctgtgaattt aaaatctagg aggggcagac attgaacaat	2040
	cccaatatat taataactta tgcaagattt tttttattaa tgagatgatg tgtttgtgac	2100
15	tgagattgag tcatacatat cactaagaaa tgggtccaag taccaaaacta tcatgaccca	2160
	gttgcaaaca tgacgttcgg gagtggtcac tttgatagtt caatttcac tgggttctt	2220
20	attcctttta taattctaata tcttcttggtg taaactattt catgtattat ttttctttaa	2280
	aatttacatg tcatttatat tgctcacta actcaattt gcataataca atgataagtg	2340
	atattttgac tcacaaaatt tacatcaaat ttcgacatog tttattatgt toattggatg	2400
25	attaacaaat ataacaaact ttgcaactaa ttaaccacca actgaatata attaaactata	2460
	actgtgaaag tagttaactc atttttatat ttcatagatc aaataagaga aataacggta	2520
	tattaatccc tccaaaaaaa aaaaacggta tatttactaa aaaatctaag ccacgtagga	2580
30	ggataacagg atccccgtag gaggataaca tccaatccaa ccaatcaca caatcctgat	2640
	gagataaccc actttaagcc cagcatctg tggcacatct acattatcta aatcacacat	2700
	tcttccacac atctgagcca cacaaaaacc aatccacatc tttatcacc attctataaa	2760
35	aaatcacact ttgtgagtct acactttgat tcccttcaaa cacatacaaa gagaagagac	2820
	taattaatta attaatcatc ttgagagaaa atgtcgggta aaggagaagg accagctatc	2880
40	ggtatcgatc ttggtaccac ttactcttgc gtcggagtat ggcaacacga ccgtgttgag	2940
	atcattgcta atgatcaagg aaacagaacc acgccatctt acgttgcttt caccgactcc	3000
	gagagggtga tcggtgacgc agctaagaat caggtcgcca tgaaccccg taacaccgtt	3060
45	ttcgacgcta agagggtgat cggtcgtcgt ttctctgaca gctctgttca gagtgacatg	3120
	aaattgtggc cattcaagat tcaagccgga cctgccgata agccaatgat ctacgtcgaa	3180
	tacaaggggtg aagagaaaga gttcgcagct gaggc.gattt cttccatggt tcttattaag	3240

50

atgcgtgaga ttgctgaggc ttaccttggg gtcacaaatca agaacgccgt tgttacccgtt 3300

ccagcttact tcaacgactc tcagcgtcag gctacaaagg atgctgggtgt catcgctggg 3360

ttgaacgtta tgcgaatcat caacgagcct acagccgccg ctattgccta cgggtottgac 3420

55

aaaaaggcta ccagcgttgg agagaagaat gttcttatct tcgatcttgg tgggtggcact 3480

	tttgatgtct ctcttcttac cattgaagag ggtatctttg aggtgaaggc aactgctggt	3540
	gacacccatc ttggtgggga agattttgac aacagaatgg ttaaccactt tgtccaagag	3600
5	ttcaagagga agagtaagaa ggatatcacc ggtaacccaa gagctcttag gaggttgaga	3660
	acttcctgtg agagagcgaa gaggactctt tcttccactg ctcagaccac catcgagatt	3720
	gactctctat acgagggtat cgacttctac tccaccatca cccgtgctag atttgaggag	3780
10	ctcaacatgg atctcttcag gaagtgtatg gagccagttg agaagtgtct tcgtgatgct	3840
	aagatggaca agagcactgt tcatgatgtt gtccttgttg gtggttctac ccgtatccct	3900
15	aaggttcagc aattgctcca ggacttcttc aacggcaaag agctttgcaa gtctattaac	3960
	cctgatgagg ctggtgccta cgggtgtgct gtccaggag ctattctcag cggatgaagga	4020
	aacgagaagg ttcaagatct tctattgctc gatgtcactc ctctctccct tggtttgaa	4080
20	actgccggtg gtgtcatgac cactttgatc ccaaggaaca caaccatccc aaccaagaag	4140
	gaacaagtct tctccaccta ctcagacaac caaccgggtg tgttgatcca ggtgtacgaa	4200
25	ggagagagag ccagaaccaa ggacaacaac cttcttggtg aatttgagct ctccggaatt	4260
	cctccagctc ctctggtgt cccccagatc acagtctgct ttgacattga tgccaatggt	4320
	atcctcaatg tctctgctga ggacaagacc accggacaga agaacaagat caccatcacc	4380
30	aatgacaagg gtcgtctctc caaggatgag attgagaaga tggttcaaga ggctgagaag	4440
	tacaagtccg aagacgagga gcacaagaag aaggttgaag ccaagaacgc tctcgagaac	4500
	tacgcttaca acatgaggaa caccatccaa gacgagaaga ttggtgagaa gctcccggtc	4560
35	gcagacaaga agaagatcga ggattctatt gagcaggcga ttcaatggct cgagggtaac	4620
	cagttggctg aggtgatga gttcgaagac aagatgaagg aattggagag catctgcaac	4680
40	ccaatcattg ccaagatgta ccaaggagct ggtggtgaag ccggtggtcc aggtgcctct	4740
	ggtatggacg atgatgctcc ccctgcttca ggcggtgctg gacctaagat cgaggaggtc	4800
	gactaagagc tcagctcgaa tttccccgat cgttcaaaca tttggcaata aagtttctta	4860
45	agattgaatc ctggtgccgg tcttgcatg attatcatat aatttctgtt gaattaogtt	4920
	aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt	4980

50 agagtcccg c aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag 5040
 gataaattat cgcgcgcggt gtcattctatg ttactagatc gaattc 5086

<210> 79

55 <211> 9493

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized Construct number R870, from HindIII to EcoRI

<400> 79

5

10

15

20

25

30

35

40

45

50

	aagcttgcat	gcctgcaggt	cgactctaga	ggatccccgg	gctggctctgt	acattcatct	60
	tgccgccttt	gcattcactt	ggccacaaag	agtagagaga	aggaagagaa	gagcccagac	120
5							
	ttcaagaagc	gaccttgcaa	gtgcactcga	gggtcagaaa	ctgtatatca	tatctatgtg	180
	agagaaaggg	gaacatttga	gatggagtcc	atttacttga	ggtataactta	ttattttgat	240
	caataaattt	gtataacttct	tatttagatc	aataaatttg	tcattaagct	ataatccaaa	300
10							
	ataaattacg	atcaaatacg	caaagttag	ccagtacttg	tgttaaactt	gatggcatct	360
	cttggtttct	ttggcaatca	catgcctaag	aaataaatag	tatcatatga	ttgtgtttgg	420
15							
	tcagacttca	gagtcagatg	actctgtttg	gataaacagc	ttaattaagc	gcttatagaa	480
	tatcatatga	ttgtgtttgg	tcagacttca	gagcatctct	tggtttctct	ggcaatcata	540
	tgccaaagaa	ataaatagta	tcatatgatt	gtgtttggtc	agacttcaga	gtcagatgac	600
20							
	cctgtttggg	taaacagctt	aattaagtgc	ttatagaata	agcgcttatc	atataagtgc	660
	ttttgtacag	ttatttctat	gaaagtagaa	gaaatagtca	tattgtttta	atataagcta	720
25							
	tcctggagag	cttgtggaaa	taaccagaaa	agaacttatg	gacacgtcat	gagctgttta	780
	cataagatct	ccctaacagt	ctcaaaagtg	tttatgccag	tagataaatt	caaataagtc	840
	aatctaaaca	gaccctaaat	ccattatggg	acctatcatt	ttagcttatt	ccatctttat	900
30							
	taagaatgtc	atgagataac	ataatgataa	cacattatct	tgacacaaat	gggcagatct	960
	agcaatttaa	ctctggagtc	cttcaagact	gctgttctta	cgaagttcac	gtccctgaat	1020
35							
	catgttcctg	tatggaagcc	tgaaagacct	caaattctaa	aagggtggcg	taaattgaag	1080
	gtttacaaaa	tataccctgc	gggcttgaca	cagaggcaag	ctctttatac	cttccagttc	1140
	aacggggatg	ttgatttcag	aagtcacttg	gagagcaatc	cttgtgccaa	gtttgaagta	1200
40							
	atttttgtgt	agcatatgtt	gagctaccta	caatttacat	gatcacctag	cattagctct	1260
	ttcacttaac	tgagagaatg	aagttttagg	aatgagtatg	accatggagt	cgccatggct	1320
	ttgtaatgcc	taccctactt	tgcccaactc	atcggggatt	tacattcaga	aaatatacat	1380
45							
	gacttcaacc	atacttaaac	ccctttttgt	aagataactg	aatgttcata	tttaatgttg	1440
	ggttgtagtg	tttttacttg	attatatcca	gacagttaca	agttggacaa	caagattgtg	1500
50							
	ggctctgtact	gttatattat	tatttttttt	ttagcagaaa	caccttatct	tttgtttcgt	1560
	ttgaatgtag	aatgaaaata	aaagaaagaa	aatataacat	catcggccgc	gcttgtctaa	1620

tttcgggcag ttaggacccct ctccgggtcac cggaaagttt cagtagaaga aacaaaacac 1680

55

cgtgactaaa atgatactat tattttatatt attgtgtttt tcttttttct accggaactt 1740

	tttaqaacgg atcccaactc gttccggggc cgctacaact gaaacaaaag aagatatttt	1800
	ctctctcttc agaaatgtaa gttttccttt acagataccc attcaccatt tgattcagat	1860
5	gtggtgacta gagataaagc atactaattt gactcttgga aaccataaa gtttatgtta	1920
	tccgtgttct ggaccaatcc acttgggggc ataacctgtg tctatgtgtg gtttggtttc	1980
	cattctgatt tatgcggcga cttgtaattt aaaatctagg aggggcagac attgaacaat	2040
10	cccaatattt taataactta tgcaagattt tttttattaa tgagatgatg tgtttgtgac	2100
	tgagattgag tcatacattt cactaagaaa tggttccaag taccaaacta tcatgaccca	2160
	gttgcaaaca tgacgttcgg gagtggtcac tttgatagtt caatttcac ttggcttctt	2220
15	attcctttta taattctaata tcttcttgtg taaactattt catgtattat ttttctttaa	2280
	aatttacatg tcatttattt tgcctcacta actcaatttt gcatataaca atgataagt	2340
20	atattttgac tcacaaaatt tacatcaaat ttcgacatcg tttattatgt tcattggatg	2400
	attaacaaat ataacaaact ttgcaactaa ttaaccacca actgaatata attaactata	2460
	actgtgaaag tagttaactc atttttatat ttcatagatc aaataagaga aataacggta	2520
25	tattaatccc tccaaaaaaa aaaaacggta tatttactaa aaaatctaag ccacgtagga	2580
	ggataacagg atccccgtag gaggataaca tccaatccaa ccaatcacia caatcctgat	2640
	gagataaccc actttaagcc cagcatctg tggcacatct acattatcta aatcacacat	2700
30	tcttccacac atctgagcca cacaaaaacc aatccacatc tttatcacc attctataaa	2760
	aatcacact ttgtgagtct acactttgat tcccttcaaa cacatacaaa gagaagagac	2820
	taattaatta attaatactc ttgagagaaa atgtcgggta aaggagaagg accagctatc	2880
35	ggtatcgatc ttggtaccac ttactcttgc gtcggagtat ggcaacacga ccgtgttgag	2940
	atcattgcta atgatcaagg aaacagaacc acgccatctt acgttgcttt caccgactcc	3000
40	gagaggttga tcggtgacgc agctaagaat caggtcgcca tgaaccccg taacaccgtt	3060
	ttcgacgcta agaggttgat cggtcgtcgt ttctctgaca gctctgttca gagtgacatg	3120
	aaattgtggc cattcaagat tcaagccgga cctgccgata agccaatgat ctacgtcgaa	3180
45	tacaaggggtg aagagaaaaga gttcgcagct gaggagattt cttccatggt tcttattaag	3240

	atgcgtgaga ttgctgaggc ttaccttggg gtcacaatca agaacgccgt tgttaccgtt	3300
	ccagcttact tcaacgactc tcagcgtcag gctacaaagg atgctgggtg catcgctggg	3360
50	ttgaacgtta tgcgaaatcat caacgagcct acagccgccg ctattgccta cgggtottgac	3420
	aaaaaggcta ccagcgttgg agagaagaat gttcttatct tcgatcttgg tgggtggcact	3480
	tttgatgtct ctcttcttac cattgaagag ggtatctttg aggtgaaggc aactgctggg	3540
55	gacacccatc ttggtgggga agattttgac aacagaatgg ttaaccactt tgtccaagag	3600

	ttcaagagga agagtaagaa ggatatcacc ggtaacccaa gagctcttag gaggttgaga	3660
	acttcctgtg agagagcgaa gaggactctt tcttccactg ctcagaccac catcgagatt	3720
5	gactctctat acgaggggtat cgacttctac tccaccatca cccgtgctag atttgaggag	3780
	ctcaacatgg atctcttcag gaagtgtatg gagccaqttg agaagtgtct tcgtgatgct	3840
	aagatggaca agagcactgt tcatgatgtt gtccttggtg gtggttctac ccgtatccct	3900
10	aaggttcagc aattgctcca ggacttcttc aacggcaaag agctttgcaa gtctattaac	3960
	cctgatgagg ctgttgcccta cggtgctgct gtccaggagg ctattctcag cggngaagga	4020
15	aacgagaagg ttcaagatct tctattgctc gatgtcactc ctctctccct tggtttgga	4080
	actgccggtg gtgtcatgac cactttgatc ccaaggaaca caaccatccc aaccaagaag	4140
	gaacaagtct tctccaccta ctcagacaac caaccoggtg tgttgatcca ggtgtacgaa	4200
20	ggagagagag ccagaaccaa ggacaacaac cttcttggtg aatttgagct ctccggaatt	4260
	cctccagctc ctggtggtgt ccccagatc acagtctgct ttgacattga tgccaatggt	4320
	atcctcaatg tctctgctga ggacaagacc accggacaga agaacaagat caccatcacc	4380
25	aatgacaagg gtcgtctctc caaggatgag attgagaaga tggttcaaga ggctgagaag	4440
	tacaagtccg aagacgagga gcacaagaag aaggttgaag ccaagaacgc tctcgagaac	4500
30	tacgcttaca acatgaggaa caccatccaa gacgagaaga ttggtgagaa gctcccggt	4560
	gcagacaaga agaagatcga ggattctatt gagcaggcga ttcaatggct cgagggtaac	4620
	cagttggctg aggctgatga gttcgaagac aagatgaagg aattggagag catctgcaac	4680
35	ccaatcattg ccaagatgta ccaaggagct ggtggtgaag ccggtggtcc aggtgcctct	4740
	ggtatggacg atgatgctcc ccctgcttca ggcggtgctg gacctaagat cgaggaggtc	4800
	gactaagagc tcagctcgaa tttccccgat cgttcaaaca tttggcaata aagtttctta	4860
40	agattgaatc ctgttgccgg tcttgcgatg attatcatat aatttctggt gaattacgtt	4920
	aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt	4980
45	agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag	5040
	gataaattat cgcgcgcggt gtcattctatg ttactagatc gaattcgtaa tcatggatc	5100
	agctgtttcc tgtgtgaaat tggtatccgg ggctggtctg tacattcatc ttgccgcctt	5160

50

tgcattcact tggccacaaa gagtagagag aaggaagaga agagcccaga cttcaagaag 5220

cgaccttgca agtgcaactcg agggtcagaa actgtatatc atatctatgt gagagaaagg 5280

ggaacatttg agatggagtc catttacttg aggtatactt attattttga tcaataaatt 5340

55

tgtatacttc ttatttagat caataaattt gtcattaagc tataatccaa aataaattac 5400

	gatcaaatat gcaaatgtta gccagtactt gtgttaaact tgatggcatc tcttggtttc	5460
	tttggcaatc acatgcctaa gaaataaata gtatcatatg atttgtgttg gtcagacttc	5520
5	agagtcagat gactctgttt ggataaacag ctttaattaag cgcttataga atatcatatg	5580
	atttgtgttg gtcagacttc agagcatctc ttggtttctc tggcaatcat atgcctaaga	5640
	aataaatagt atcatatgat tgtgttttgt cagacttcag agtcagatga ccctgtttgg	5700
10	gtaaacagct taattaagtg cttatagaat aagcgcttat catataagtg cttttgtaca	5760
	gttattttcta tgaaagtaga agaaatagtc atattgtttt aatataagct atcctggaga	5820
	gcttgtggaa ataaccagaa agaacttat ggacacgtca tgagctgttt acataagatc	5380
15	tccctaacag tctcaaaagt gtttatgcc a gtagataaat tcaaataagt caatctaaac	5940
	agaccctaaa tccattatgg tacctatcat tttagcttat tccatcttta ttaagaatgt	6000
20	catgagataa cataatgata acacattatt ttgacacaaa tgggcagatc tagcaattta	6060
	actctg9agt ccttcaagac t9ctgttctt acgaagttca cgtccctgaa tcatgttcct	6120
	gtatggaagc ctgaaagacc tcaaattcta aaaggtggcg ataaattgaa gggtttacaaa	6180
25	atataccct9 cgggcttgac acagaggcaa gctctttata ccttcagtt caacgg9gat	6240
	gttgatttca gaagtcactt ggagagcaat ccttgtgcc a gtttgaagt aatttttgtg	6300
	tagcatatgt tgagctacct acaatttaca tgatcaccta gcattagctc tttcacttaa	6360
30	ctgagagaat gaagtttttag gaatgagtat gaccatggag tcggcatggc tttgtaatgc	6420
	ctaccctact ttggccaact catcggggat ttacattcag aaaatataca tgacttcaac	6480
	cataacttaaa cccctttttg taagataact gaatgttcat atttaatgtt gggttgtagt	6540
35	gtttttactt gattatatcc agacagttac aagttggaca acaagattgt gggctctgtac	6600
	tgttattttat ttattttttt tttagcagaa acaccttacc ttttgtttcg tttgaatgta	6660
40	gaatgaaaat aaaagaaaga aaatataaca tcatcggccg cgcttgtcta atttcgggca	6720
	gttaggatcc tctccggtca ccggaaagt ttcagtagaag aaacaaaaca ccgtgactaa	6780
	aatgatacta ttattttatt tatttgtgtt ttcttttttc taccggaact ttttagaacg	6840
45	ratcccaact cgttccgggg ccgctacaac tgaaacaaaa gaagatattt tctctctctt	6900
	cagaaatgta agttttcctt tacagatacc cattcaccat ttgattcaga tgtggtgact	6960
	agagataaag cataactaatt tgactcttgg aaacccataa agtttatgtt atccgtgttc	7020

50

tggaccaatc cacttggggg cataacctgt gtctatgtgt ggtttggttt ccattctgat 7080

ttatgcggcg acttgtaatt taaaatctag gaggggcaga cattgaacaa tcccaatatt 7140

ttaataactt atgcaagatt ttttttatta atgagatgat gtgtttgtga ctgagattga 7200

55

gtcat=tcatt tcactaagaa atggttccaa gtaccaaact atcatgaccc agttgcaaac 7260

	atgacgttcg ggagtgggtca ctttgatagt tcaatttcat cttgggttct tattcctttt	7320
	ataattctaa ttcttcttgt gtaaaactatt tcatgtatta tttttcttta aaatttacat	7380
5	gtcattttatt ttgcctcact aactcaattt tgcatataac aatgataagt gatattttga	7440
	ctcacaaaat ttacatcaaa tttcgacatc gtttattatg ttcattggat gattaacaaa	7500
	tataacaaac tttgcaacta attaaccacc aactgaatat aattaactat aactgtgaaa	7560
10	gtagttaact cttttttata tttcatagat caaataagag aaataacggt atattaatcc	7620
	ctccaaaaaa aaaaaacggt atatttacta aaaaatctaa gccacgtagg aggataacag	7680
15	gatccccgta ggaggataac atccaatcca accaatcaca acaatcctga tgagataacc	7740
	cactttaagc ccacgcatct gtggcacatc tacattatct aaatcacaca ttcttcaca	7800
	catctgagcc acacaaaaac caatccacat ctttatcacc cattctataa aaaatcacac	7860
20	tttgtgagtc tacactttga ttcccttcaa acacatacaa	
	agagaagaga ctaattaatt	7920
	aattaatcat cttgagagaa aatgtttggg cgcggaacca caaggaagag tgataacacc	7980
	aaatattacg atattcttgg tgtttcaaaa agtgctagtg aagatgaaat caagaaagcc	8040
25	tatagaaagg cagcgatgaa gaaccatcca gataagggtg gggatcctga gaagttcaag	8100
	gagttggggc aagcatatga agtggtgagc gatcctgaaa agaaagaact gtatgatcaa	8160
30	tatggtgaag atgcccttaa agaaggaatg gggggaggcg caggaaagctc atttcataat	8220
	ccgtttgata ttttcgaatc attttttggg gcaggctttg gtggtggtgg tccttcacgc	8280
	gcaagaagac agaagcaagg agaagatgtg gtgcattcta taaaggtttc cttggaggat	8340
35	gtgtataacg gcactacaaa gaagctatca ctttctagga atgcactgtg ctcaaaatgt	8400
	aaagggaaag gttcaaaaag tggaactgct ggaagggtgtt ttggatgcca gggcacaggt	8460
	atgaagatta ccagaaggca aattggactg ggcatgattc acaaaatgca acacgtctgt	8520
40	cctgactgca aaggaacagg cgaggctatt agtgagagag atagatgcc tcaatgcaag	8580
	ggaaacaaga ttactcaaga aaagaagggtg ctggagggtg atgtggaaaa ggggatgcag	8640
45	cagggtcaca agattgtatt cgaaggacaa gctgatgaag ctcttgatac aatcacagga	8700
	gacataqttt ttgtcttgca agtaaaggga catccgaagt ttcgagggga gcgtgatgac	8760

	ctccacattg aacacaattt gagcttaact gaggctctct gtggcttcca gtttaatgtc	8820
50	acacatcttg atggaaggca actattgggc aaatcgaacc ccggcgaagt catcaagcca	8880
	ggccaacata aagctataaa tgatgagggg atgccacaac atggtaggcc gttcatgaag	8940
	ggacgcctat acatcaagtt tagtggtgat ttcccggatt cgggttttct ttccccaagc	9000
55	caaagcctgg aattagaaaa gatattacct caaaagacaa gcaagaactt gtcccaaaag	9060

	gaggtagatg attgtgagga gaccaccctg catgatgtca atattgcaga ggagatgagt	9120
	cgaaagaagc aacaataccg tgaggcatat gatgacgatg atgatgaaga tgatgagcac	9180
5	tgcgagcctc ggggtgcaatg cgctcaacag taggagctca gctcgaattt ccccgatcgt	9240
	tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt	9300
	atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg	9360
10	ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta ataogcgata	9420
	gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta	9480
15	ctagatcgaa ttc	9493

<210> 80

<211> 568

<212> PRT

20 **<213> Artificial Sequence**

<220>

<223> Synthesized amino acid translation of coding sequence in construct 690 expression cassette

25 **<400> 80**

30

35

40

45

50

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

5

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45

10

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80

15

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly His Phe Ala
100 105 110

20

Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
115 120 125

Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
130 135 140

25

Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
145 150 155 160

30

Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
165 170 175

Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp

35

40

	180	185	190
	Gly Val His His Pro Pro Asn Ile Gly Asp Gln Lys Ala Leu Tyr His		
	195	200	205
5	Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys		
	210	215	220
	Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly		
10	225	230	235
	Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile		
		245	250
	Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu		
15	260	265	270
	Ser Arg Gly Phe Gly Ser Gly Ile Met Lys Ser Glu Leu Glu Tyr Gly		
		275	280
20	Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser		
	290	295	300
	Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys		
	305	310	315
25	Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser		
		325	330
	Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile		
		340	345
30	Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr		
		355	360
	Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys		
		370	375
35	Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser		
	385	390	395
	Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe		
40		405	410
	Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp		
		420	425
45	Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met		
		435	440
	Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu		
		450	455
50	Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly		
	465	470	475
	Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu		
		485	490

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
 515 520 525
 Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
 5 530 535 540
 Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
 545 550 555 560
 10 Ser Leu Gln Cys Arg Ile Cys Ile
 565

<210> 81

<211> 568

15 <212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Amino acid translation of coding sequence in construct 691 expression cassette

20

<400> 81

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
 1 5 10 15
 25 Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 20 25 30
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 30 35 40 45
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
 50 55 60
 35 Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 65 70 75 80
 Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
 85 90 95
 40 Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
 100 105 110
 Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
 45 115 120 125
 Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
 130 135 140
 Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
 50 145 150 155 160

Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
165 170 175

55

Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
180 185 190

Gly Val His His Pro Pro Asn Ile Gly Asp Gln Lys Ala Leu Tyr His

	195	200	205
	Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys		
	210	215	220
5			
	Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly		
	225	230	235 240
	Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile		
10	245	250	255
	Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu		
	260	265	270
	Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp		
15	275	280	285
	Lys Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser		
	290	295	300
20			
	Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys		
	305	310	315 320
	Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser		
	325	330	335
25			
	Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile		
	340	345	350
	Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr		
	355	360	365
30			
	Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys		
	370	375	380
	Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser		
	385	390	395 400
35			
	Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe		
	405	410	415
	Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp		
40	420	425	430
	Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met		
	435	440	445
45			
	Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu		
	450	455	460
	Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly		
	465	470	475 480
50			
	Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu		

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560

Ser Leu Gln Cys Arg Ile Cys Ile
565

10 <210>82

<211> 572

<212> PRT

<213> Artificial Sequence

15 <220>

<223> Synthesized Amino acid translation of coding sequence in construct 696 expression cassette

<400> 82

20 Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
20 25 30

25

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
50 55 60

30

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65 70 75 80

35

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
85 90 95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
100 105 110

40

Thr Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu
115 120 125

Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser
130 135 140

45

Ser Trp Ser Asp His Glu Ala Ser Ser Gly Val Ser Ser Ala Cys Pro
145 150 155 160

	Tyr	Leu	Gly	Ser	Pro	Ser	Phe	Phe	Arg	Asn	Val	Val	Trp	Leu	Ile	Lys
50					165					170					175	
	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile	Lys	Lys	Ser	Tyr	Asn	Asn	Thr	Asn
				180					185					190		
55	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp	Gly	Ile	His	His	Pro	Asn	Asp	Ala
			195					200					205			
	Ala	Glu	Gln	Thr	Arg	Leu	Tyr	Gln	Asn	Pro	Thr	Thr	Tyr	Ile	Ser	Ile

210 215 220
 Gly Thr Ser Thr Leu Asn Gln Arg Leu Val Pro Lys Ile Ala Thr Arg
 225 230 235 240
 5
 Ser Lys Val Asn Gly Gln Ser Gly Arg Met Glu Phe Phe Trp Thr Ile
 245 250 255
 Leu Lys Pro Asn Asp Ala Ile Asn Phe Glu Ser Asn Gly Asn Phe Ile
 260 265 270
 10
 Ala Pro Glu Tyr Ala Tyr Lys Ile Val Lys Lys Gly Asp Ser Ala Ile
 275 280 285
 Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
 15 290 295 300
 Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
 305 310 315 320
 20 Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
 325 330 335
 Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
 340 345 350
 25 Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
 355 360 365
 Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
 370 375 380
 30
 Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
 385 390 395 400
 Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
 405 410 415
 35
 Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg JY:et Glu Asn Leu Asn Lys
 420 425 430
 Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
 40 435 440 445
 Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
 450 455 460
 45 Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
 465 470 475 480
 Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
 485 490 495
 50 Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
 500 505 510
 Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
 515 520 525

Glu Ser Met Gly Val Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala
530 535 540

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Trp Met
 545 550 555 560

Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

5 565 570

<210> 83

<211>575

<212> PRT

10 <213> Artificial Sequence

<220>

<223> Synthesized amino acid translation of coding sequence in construct 737 expression cassette

15 <400> 83

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
 1 5 10 15

20 Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
 20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
 35 40 45

25 Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
 50 55 60

30 Glu Leu Val Gln Ser Ser Ser Thr Gly Glu Ile Cys Asp Ser Pro His
 65 70 75 80

Gln Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
 85 90 95

35 Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
 100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
 115 120 125

40 Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
 130 135 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
 145 150 155 160

45 Ser Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn
 165 170 175

	Trp	Leu	Thr	His	Leu	Lys	Phe	Lys	Tyr	Pro	Ala	Leu	Asn	Val	Thr	Met
50					180				185					190		
	Pro	Asn	Asn	Glu	Lys	Phe	Asp	Lys	Leu	Tyr	Ile	Trp	Gly	Val	His	His
		195						200					205			
55	Pro	Gly	Thr	Asp	Asn	Asp	Gln	Ile	Phe	Leu	Tyr	Ala	Gln	Ala	Ser	Gly
		210					215					220				
	Arg	Ile	Thr	Val	Ser	Thr	Lys	Arg	Ser	Gln	Gln	Thr	Val	Ile	Pro	Asn

Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser
545 550 555 560

Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile

565

570

575

5

<210> 84

<211> 593

<212> PRT

<213> Artificial Sequence

10

<220>

<223> Synthesized amino acid translation of coding sequence in construct 745 expression cassette

<400> 84

15

20

25

30

35

40

45

50

	Met	Ala	Lys	Asn	Val	Ala	Ile	Phe	Gly	Leu	Leu	Phe	Ser	Leu	Leu	Val	
	1			5						10					15		
5	Leu	Val	Pro	Ser	Gln	Ile	Phe	Ala	Asp	Arg	Ile	Cys	Thr	Gly	Ile	Thr	
				20					25					30			
	Ser	Ser	Asn	Ser	Pro	His	Val	Val	Lys	Thr	Ala	Thr	Gln	Gly	Glu	Val	
			35					40					45				
10	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	Thr	Thr	Thr	Pro	Thr	Lys	Ser	Tyr	
		50					55					60					
	Phe	Ala	Asn	Leu	Lys	Gly	Thr	Arg	Thr	Arg	Gly	Lys	Leu	Cys	Pro	Asp	
	65					70					75					80	
15																	
	Cys	Leu	Asn	Cys	Thr	Asp	Leu	Asp	Val	Ala	Leu	Gly	Arg	Pro	Met	Cys	
					85					90					95		
	Val	Gly	Thr	Thr	Pro	Ser	Ala	Lys	Ala	Ser	Ile	Leu	His	Glu	Val	Lys	
20					100				105					110			
	Pro	Val	Thr	Ser	Gly	Cys	Phe	Pro	Ile	Met	His	Asp	Arg	Thr	Lys	Ile	
					115				120				125				
25	Arg	Gln	Leu	Pro	Asn	Leu	Leu	Arg	Gly	Tyr	Glu	Asn	Ile	Arg	Leu	Ser	
		130					135					140					
	Thr	Gln	Asn	Val	Ile	Asp	Ala	Glu	Lys	Ala	Pro	Gly	Gly	Pro	Tyr	Arg	
	145					150					155					160	
30																	
	Leu	Gly	Thr	Ser	Gly	Ser	Cys	Pro	Asn	Ala	Thr	Ser	Lys	Ser	Gly	Phe	
					165					170					175		
	Phe	Ala	Thr	Met	Ala	Trp	Ala	Val	Pro	Lys	Asp	Asn	Asn	Lys	Asn	Ala	
				180						185				190			
35																	
	Thr	Asn	Pro	Leu	Thr	Val	Glu	Val	Pro	Tyr	Ile	Cys	Thr	Glu	Gly	Glu	
			195					200					205				
	Asp	Gln	Ile	Thr	Val	Trp	Gly	Phe	His	Ser	Asp	Asn	Lys	Thr	Gln	Met	
40			210				215					220					
	Lys	Asn	Leu	Tyr	Gly	Asp	Ser	Asn	Pro	Gln	Lys	Phe	Thr	Ser	Ser	Ala	
	225					230					235					240	
45																	
	Asn	Gly	Val	Thr	Thr	His	Tyr	Val	Ser	Gln	Ile	Gly	Ser	Phe	Pro	Asp	

50

55

		245		250		255	
	Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp						
		260		265		270	
5							
	Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg						
		275		280		285	
	Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys						
10		290		295		300	
	Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His						
		305		310		315	320
	Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu						
15		325		330		335	
	His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu						
		340		345		350	
20	Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu Lys						
		355		360		365	
	Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly Trp						
		370		375		380	
25	Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His						
		385		390		395	400
	Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn						
		405		410		415	
30							
	Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn						
		420		425		430	
	Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu						
		435		440		445	
35							
	Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser						
		450		455		460	
	Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser						
40		465		470		475	480
	Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu						
		485		490		495	
45	Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His						
		500		505		510	
	Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn						
		515		520		525	
50	Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala						
		530		535		540	

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ser Ile
545 550 555 560

55

Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly
565 570 575

Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys

580

585

590

Ile

5

<210> 85

<211> 405

<212> DNA

<213> Artificial Sequence

10

<220>

<223> Synthesized SacI-Piast0 3(prime)UTR sequence

<400> 85

15

	gagctctaag ttaaaatgct tcttcgtctc ctatttataa tatggtttgt tattgttaat	60
	tttgttcttg tagaagagct taattaatcg ttgttggtat gaaatactat ttgtatgaga	120
	tgaactggtg taatgtaatt catttacata agtggagtca gaatcagaat gtttcctcca	180
20	taactaacta gacatgaaga cctgccgcgt acaattgtct tatatttgaa caactaaaat	240
	tgaacatctt ttgccacaac ttataagtg gttaatatag ctcaaataa tggatcaagtt	300
25	caatagatta ataattgaaa tatcagttat cgaaattcat taacaatcaa cttaacgtta	360
	ttaactacta attttatatc atcccctttg ataaatgata gtaca	405

30 <210> 86
 <211> 1718
 <212> DNA
 <213> Artificial Sequence

35 <220>
 <223> Synthesized POI sp - A/California/04/09

<400> 86

40

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttggtgtt ggttcctctc 60
 agatcttcgc tgacacatta tgtatagggt atcatgcgaa caattcaaca gacactgtag 120
 5 acacagtact agaaaagaat gtaacagtaa cacactctgt taaccttcta gaagacaagc 180

ataacgggaa actatgcaaa ctaagagggg tagccccatt gcatttgggt aaatgtaaca 240
 ttgctggctg gatcctggga aatccagagt gtgaatcact ctccacagca agctcatggt 300

10 cctacattgt ggaaacacct agttcagaca atggaacgtg ttaccagga gatttcacgc 360
 attatgagga gctaagagag caattaagct cagtgtcatc atttgaaagg tttgagatat 420

15 tccccaagac aagttcatgg cccaatcatg actcgaacaa aggtgtaacg gcagcatgtc 480
 ctcatgctgg agcaaaaagc ttctacaaaa atttaatatg gctagttaaa aaaggaaatt 540
 catacccaaa gctcagcaaa tcctacatta atgataaagg gaaagaagtc ctcggtgctat 600

20 ggggcattca ccatccatct actagtgtctg accaaciaaag tctctatcag aatgcagata 660
 catatgtttt tgtggggtca tcaagataca gcaagaagtt caagccggaa atagcaataa 720

25 gacccaaagt gagggatcaa gaaggagaa tgaactatta ctggacacta gtagagccgg 780
 gagacaaaat aacattcgaa gcaactggaa atctagtgggt accgagatat gcattcgcaa 840

30 tggaaagaaa tgctggatct ggtattatca tttcagatac accagtccac gattgcaata 900
 caacttgtca aacaccaag ggtgctataa acaccagcct cccatttcag aatatacatc 960
 cgatcacaat tggaaaatgt ccaaaatatg taaaaagcac aaaattgaga ctggccacag 1020

35 gattgaggaa tatcccgctc attcaatcta gaggactatt tggggccatt gccggtttca 1080
 ttgaaggggg gtggacagg atggtagatg gatggtacgg ttatcaccat caaaatgagc 1140
 aggggtcagg atatgcagcc gacctgaaga gcacacagaa tgccattgac gagattacta 1200

40

	acaaagtaaa ttctgttatt gaaaagatga atacacagtt cacagcagta ggtaaagagt	1260
	tcaaccacct ggaaaaaaga atagagaatt taaataaaaa agttgatgat ggtttcctgg	1320
45	acatttggac ttacaatgcc gaactgttgg ttctattgga aaatgaaaga actttggact	1380
	accacgattc aaatgtgaag aacttatatg aaaaggtaag aagccagcta aaaaacaatg	1440
	ccaaggaaat tggaaacggc tgctttgaat ttaccacaa atgcgataac acgtgcatgg	1500
50		
	aaagtgtcaa aaatgggact tatgactacc caaaatactc agaggaagca aaattaaaca	1560
	gagaagaaat agatggggta aagctggaat caacaaggat ttaccagatt ttggcgatct	1620
55	attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc agtttctgga	1680
	tgtgctctaa tgggtctcta cagtgtagaa tatgtatt	1718

<210> 87

<211> 573

<212> PRT

<213> Artificial Sequence

5

<220>

<223> Synthesized POI sp - A/California 04/09

<400> 87

10

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Leu Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

20

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Lys His Asn Gly Lys
50 55 60

Leu Cys Lys Leu Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn
65 70 75 80

25

30

35

40

45

50

Ile Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr
85 90 95

Ala Ser Ser Trp Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly

5

100 105 110

Thr Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln
115 120 125

10

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr
130 135 140

Ser Ser Trp Pro Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys
145 150 155 160

15

Pro His Ala Gly Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val
165 170 175

Lys Lys Gly Asn Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp
180 185 190

20

Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr
195 200 205

Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Thr Tyr Val Phe
210 215 220

25

Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile
225 230 235 240

Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr
245 250 255

30

Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu
260 265 270

Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly
275 280 285

35

Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln
290 295 300

Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His
305 310 315 320

40

Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu
325 330 335

45

Arg Leu Ala Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly
340 345 350

Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
355 360 365

50

Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
370 375 380

Tyr	Ala	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Asn	Ala	Ile	Asp	Glu	Ile	Thr
385					390					395					400

55

Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala

405 410 415
 Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
 5 420 425 430
 Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
 435 440 445
 10 Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
 450 455 460
 Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
 465 470 475 480
 15 Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
 485 490 495
 Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
 500 505 510
 20 Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
 515 520 525
 Leu Glu Ser Thr Arg Ile Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val
 25 530 535 540
 Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile Ser Phe Trp
 545 550 555 560
 30 Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
 565 570

<210> 88

<211> 747

35 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized 2X35S promoter sequence

40

<400> 88

45

	gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagttctca	60
	gaagacccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga	120
5	ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
	tacaaatgcc atcattgcga taaaggaaag gccatcggtg aagatgcctc tgccgacagt	240
	ggtcccaaaag atggaccccc acccacgagg agcatcggtg aaaaagaaga cgttccaacc	300
10	acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac	360
	tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa	420
15	agggtaatat ccggaaacct cctcggtatc cattgcccag ctatctgtca ctttattgtg	480
	aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
	atcggtgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
20	atcggtgaaa aagaagacgt tccaaccacg ttttcaaagc aagtggattg atgtgatatc	660
	tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttctctata	720
25	taaggaagtt catttcattt ggagagg	747

<210> 89

<211> 43

30 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthesized PacI-MCS-2X35S.c

35

<400> 89

aattgttaat taatcgaca agcttcatg cctgcagtc aac 43

<210> 90

40 <211> 48

<212> DNA

<213> Artificial Sequence

<220>

45 <223> Synthesized CPMV 5(prime)UTR-2X35S.r

<400> 90

tcaaaaccta ttaagatttt aatacctctc caaatgaaat gaacttc **48**

- 50
- <210> 91**
 - <211> 49**
 - <212> DNA**
 - <213> Artificial Sequence**

- 55
- <220>**
 - <223> Synthesized 2X35S-CPMV 5(prime)UTR.c**
 - <400> 91**

ttggagaggt attaaaatct taatagggtt tgataaaagc gaacgtggg 49

 <210> 92
 <211> 44
 5 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Synthesized Apal-M prot.r
 10

 <400> 92
 tctccatggg cccgacaaat ttgggcagaa tatacagaag ccta 44

 <210> 93
 15 <211>3580
 <212> DNA
 <213> Artificial Sequence

 <220>
 20 <223> Synthesized Construct 747, from Paci to AseI

 <400> 93

 25 ttaattaagt cgacaagctt gcatgcctgc aggtcaacat ggtggagcac gacacacttg 60

 30

 35

 40

 45

 50

 55

	tctactccaa aaatatcaaa gatacagtct cagaagacca aagggcaatt gagacttttc	120
	aacaaaggggt aatatccgga aacctcctcg gattccattg cccagctatc tgtcacttta	180
5	ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccatcattgc gataaaggaa	240
	aggccatcgt tgaagatgcc tctgccgaca gtgggtcccaa agatggaccc ccacccacga	300
	ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg	360
10		
	ataacatgggt ggagcacgac acacttgtct actccaaaaa tatcaaagat acagtctcag	420
	aagaccaaag ggcaattgag acttttcaac aaagggtaat atccggaaac ctccctcgat	480
	tccattgccc agctatctgt cactttattg tgaagatagt ggaaaaggaa ggtggctcct	540
15		
	acaaatgcc a tcattgcgat aaaggaaagg ccatcgttga agatgcctct gccgacagtg	600
	gtoccaaaga tggaccccca cccacgagga gcatcgtgga aaaagaagac gttccaacca	660
20	cgtottcaaa gcaagtggat tgatgtgata tctccactga cgtaagggat gacgcacaat	720
	cccactatcc ttcgcaagac ccttcctcta tataaggaag ttcatttcat ttggagaggt	780
	attaaaatct taataggttt tgataaaagc gaacgtgggg aaacccgaac caaaccttct	840
25	tctaaactct ctctcatctc tcttaaagca aacttctctc ttgtctttct tgcgtgagcg	900
	atcttcaacg ttgtcagatc gtgcttcgqc accagtacaa cgttttcttt cactgaagcg	960
	aatcaaaga tctctttgtg gacacgtagt gcggcgccat taaataacgt gtacttgtcc	1020
30		
	tattcttgtc ggtgtggtct tgggaaaaga aagcttgctg gaggtgctg ttcagcccca	1080
	tacattactt gttacgattc tgctgacttt cggcggtgac aatatctcta cttctgcttg	1140
	acgaggtatt gttgcctgta cttctttctt cttcttcttg ctgattgggt ctataagaaa	1200
35		
	tctagtatth tctttgaaac agagttttcc cgtggttttc gaacttgag aaagattggt	1260
	aagcttctgt atattctgcc caaatttctc gggcccatgg cgaaaaacgt tgcgattttc	1320
40	ggcttattgt tttctcttct tgtgttggtt ccttctcaga tcttcgctga tcgaatctgc	1380
	actggaataa catcttcaaa ctcacctcat gtgggtcaaaa cagccactca aggggaggtc	1440
	aatgtgactg gtgtgatacc actaacaaca acaccaaca aatcttattt tgcaaactct	1500
45	aaaggaacaa ggaccagagg gaaactatgc ccagactgtc tcaactgcac agatctggat	1560
	gtggcttttg gcagaccaat gtgtgtgggg accacacctt cggcgaaggc ttcaatactc	1620
	cacgaagtca aacctgttac atccgggtgc tttcctataa tgcacgacag aacaaaaalc	1680
50		
	aggcaactac ccaatcttct cagaggatat gaaaatatca ggctatcaac ccaaaacgtc	1740
	atcqtatcgg aaaaggcacc aggaggaccc tacagacttg gaacctcagg atcttgcctt	1800

55

aacgctacca gtaagagcgg atttttcgca acaatggctt gggctgtccc aaaggacaac 1860
aacaaaaatg caacgaaccc actaacagta gaagtaccat acatttgtac agaaggggaa 1920

	gaccaaataca ctgtttgggg gttccattca gataacaaaa cccaaatgaa gaacctctat	1980
	ggagactcaa atcctcaaaa gttcacctca tctgctaatag gagtaaccac acactatgtt	2040
5	tctcagattg gcagcttccc agatcaaaca gaagacggag gactaccaca aagcggcagg	2100
	attgttggtg attacatgat gcaaaaacct gggaaaacag gaacaattgt ctaccaaaga	2160
	ggtgttttgt tgcctcaaaa ggtgtggtgc gcgagtggca ggagcaaagt aataaaagg	2220
10	tccttgccctt taattggtga agcagattgc cttcatgaaa aatacgggtg attaaacaaa	2280
	agcaagcctt actacacagg agaacatgca aaagccatag gaaattgccc aatatgggtg	2340
15	aaaacacctt tgaagctcgc caatggaacc aaatatagac ctctgcaaa actattaaag	2400
	gaaaggggtt tcttcggagc tattgctggt ttctagaag gaggatggga aggaatgatt	2460
	gcaggctggc acggatacac atctcacgga gcacatggag tggcagtggc ggcggacctt	2520
20	aagagtacgc aagaagctat aaacaagata acaaaaaatc tcaattcttt gagtgaagcta	2580
	gaagtaaaga atcttcaaag actaagtggg gccatggatg aactccacaa cgaaatactc	2640
25	gagctggatg agaaagtgga tgatctcaga gctgacacta taagctcgca aatagaactt	2700
	gcagtcttgc tttccaacga aggaataata aacagtgaag atgagcatct attggcactt	2760
	gagagaaaac taaagaaaat gctgggtccc tctgctgtag agataggaaa tggatgcttc	2820
30	gaaaccaaac acaagtgcaa ccagacctgc ttagacagga tagctgctgg cacctttaat	2880
	gcaggagaat tttctctccc cacttttgat tcaactgaaca ttactgctgc atctttaaat	2940
	gatgatggat tggataacta ccaaatactg tcaatttatt caacagtggc gagttcccta	3000
35	gcactggcaa tcatgatggc tggctctatct ttatggatgt gctccaatgg atcgttacaa	3060
	tgcagaatth gcatttaaag gcctatthtc tttagtttga atttactgtt attcgggtgtg	3120
40	catttctatg tttggtgagc ggttttctgt gctcagagtg tgtttattht atgtaattta	3180
	atttctttgt gagctcctgt ttagcaggtc gtcccttcag caaggacaca aaaagattht	3240
	aattttatta aaaaaaaaaa aaaaaaagac cgggaattcg atatcaagct tatcgacctg	3300
45	cagatcgthc aaacatttgg caataaagtt tottaagatt gaatcctgtt gccgggtotg	3360

	cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat	3420
50	gcatgacggt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat	3480
	acgcataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat	3540
	ctatgttact agattotaga gtctcaagct tcggcgcgcc	3580

55

<210> 94

<211> 569

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Influenza B virus B/Florida/4/2006

5

<400> 94

10

15

20

25

30

35

40

45

50

	Asp	Arg	Ile	Cys	Thr	Gly	Ile	Thr	Ser	Ser	Asn	Ser	Pro	His	Val	Val	
	1				5					10					15		
5	Lys	Thr	Ala	Thr	Gln	Gly	Glu	Val	Asn	Val	Thr	Gly	Val	Ile	Pro	Leu	
					20				25					30			
	Thr	Thr	Thr	Pro	Thr	Lys	Ser	Tyr	Phe	Ala	Asn	Leu	Lys	Gly	Thr	Arg	
			35					40					45				
10	Thr	Arg	Gly	Lys	Leu	Cys	Pro	Asp	Cys	Leu	Asn	Cys	Thr	Asp	Leu	Asp	
		50					55					60					
	Val	Ala	Leu	Gly	Arg	Pro	Met	Cys	Val	Gly	Thr	Thr	Pro	Ser	Ala	Lys	
	65					70					75					80	
15																	
	Ala	Ser	Ile	Leu	His	Glu	Val	Lys	Pro	Val	Thr	Ser	Gly	Cys	Phe	Pro	
					85					90					95		
	Ile	Met	His	Asp	Arg	Thr	Lys	Ile	Arg	Gln	Leu	Pro	Asn	Leu	Leu	Arg	
20					100				105					110			
	Gly	Tyr	Glu	Asn	Ile	Arg	Leu	Ser	Thr	Gln	Asn	Val	Ile	Asp	Ala	Glu	
			115					120					125				
25	Lys	Ala	Pro	Gly	Gly	Pro	Tyr	Arg	Leu	Gly	Thr	Ser	Gly	Ser	Cys	Pro	
		130					135					140					
	Asn	Ala	Thr	Ser	Lys	Ser	Gly	Phe	Phe	Ala	Thr	Met	Ala	Trp	Ala	Val	
	145					150					155					160	
30																	
	Pro	Lys	Asp	Asn	Asn	Lys	Asn	Ala	Thr	Asn	Pro	Leu	Thr	Val	Glu	Val	
					165					170				175			
	Pro	Tyr	Ile	Cys	Thr	Glu	Gly	Glu	Asp	Gln	Ile	Thr	Val	Trp	Gly	Phe	
				180					185					190			
35																	
	His	Ser	Asp	Asp	Lys	Thr	Gln	Met	Lys	Asn	Leu	Tyr	Gly	Asp	Ser	Asn	
			195					200					205				
	Pro	Gln	Lys	Phe	Thr	Ser	Ser	Ala	Asn	Gly	Val	Thr	Thr	His	Tyr	Val	
40			210				215					220					
	Ser	Gln	Ile	Gly	Ser	Phe	Pro	Asp	Gln	Thr	Glu	Asp	Gly	Gly	Leu	Pro	
	225					230					235					240	
45	Gln	Ser	Gly	Arg	Ile	Val	Val	Asp	Tyr	Met	Met	Gln	Lys	Pro	Gly	Lys	
					245					250				255			
	Thr	Gly	Thr	Ile	Val	Tyr	Gln	Arg	Gly	Val	Leu	Leu	Pro	Gln	Lys	Val	
				260					265					270			

50

Trp	Cys	Ala	Ser	Gly	Arg	Ser	Lys	Val	Ile	Lys	Gly	Ser	Leu	Pro	Leu
		275					280					285			
Ile	Gly	Glu	Ala	Asp	Cys	Leu	His	Glu	Lys	Tyr	Gly	Gly	Leu	Asn	Lys

55

	290		295		300	
	Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys					
	305		310		315	320
5						
	Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr					
		325		330		335
	Arg Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile					
10		340		345		350
	Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His					
		355		360		365
15						
	Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu					
		370		375		380
	Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser					
		385		390		395
20						
	Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met					
		405		410		415
	Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp					
25		420		425		430
	Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu					
		435		440		445
30						
	Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu					
		450		455		460
	Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly					
		465		470		475
35						
	Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp					
		485		490		495
	Arg Ile Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr					
		500		505		510
40						
	Phe Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu					
		515		520		525
	Asp Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu					
45		530		535		540
	Ala Val Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg					
		545		550		555
50						
	Asp Asn Val Ser Cys Ser Ile Cys Leu					
		565				

<211> 570

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Influenza B virus B/Malaysia/2506/2004

<400> 95

5

10

15

20

25

30

35

40

45

50

55

Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
 1 5 10 15

5

Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
 20 25 30

Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu
 35 40 45

10

Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
 50 55 60

Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Asn Ile Pro Ser Ala Arg
 65 70 75 80

15

Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
 85 90 95

Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Lys Leu Leu Arg
 100 105 110

20

Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
 115 120 125

Asn Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
 130 135 140

25

Asn Val Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
 145 150 155 160

30

Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser Leu Thr Ile Glu
 165 170 175

Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly
 180 185 190

35

Phe His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
 195 200 205

Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
 210 215 220

40

Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
 225 230 235 240

Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
 245 250 255

Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
 260 265 270

Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
 275 280 285

Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
 290 295 300

Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
305 310 315 320

	Cys	Pro	Ile	Trp	Val	Lys	Thr	Pro	Leu	Lys	Leu	Ala	Asn	Gly	Thr	Lys	
					325					330					335		
	Tyr	Arg	Pro	Pro	Ala	Lys	Leu	Leu	Lys	Glu	Arg	Gly	Phe	Phe	Gly	Ala	
5				340					345				350				
	Ile	Ala	Gly	Phe	Leu	Glu	Gly	Gly	Trp	Glu	Gly	Met	Ile	Ala	Gly	Trp	
			355					360					365				
10	His	Gly	Tyr	Thr	Ser	His	Gly	Ala	His	Gly	Val	Ala	Val	Ala	Ala	Asp	
		370					375					380					
	Leu	Lys	Ser	Thr	Gln	Glu	Ala	Ile	Asn	Lys	Ile	Thr	Lys	Asn	Leu	Asn	
	385					390					395					400	
15																	
	Ser	Leu	Ser	Glu	Leu	Glu	Val	Lys	Asn	Leu	Gln	Arg	Leu	Ser	Gly	Ala	
				405						410					415		
	Met	Asp	Glu	Leu	His	Asn	Glu	Ile	Leu	Glu	Leu	Asp	Glu	Lys	Val	Asp	
20				420						425				430			
	Asp	Leu	Arg	Ala	Asp	Thr	Ile	Ser	Ser	Gln	Ile	Glu	Leu	Ala	Val	Leu	
			435					440					445				
25	Leu	Ser	Asn	Glu	Gly	Ile	Ile	Asn	Ser	Glu	Asp	Glu	His	Leu	Leu	Ala	
	450						455					460					
	Leu	Glu	Arg	Lys	Leu	Lys	Lys	Met	Leu	Gly	Pro	Ser	Ala	Val	Glu	Ile	
	465				470						475					480	
30																	
	Gly	Asn	Gly	Cys	Phe	Glu	Thr	Lys	His	Lys	Cys	Asn	Gln	Thr	Cys	Leu	
					485					490					495		
	Asp	Arg	Ile	Ala	Ala	Gly	Thr	Phe	Asp	Ala	Gly	Glu	Phe	Ser	Leu	Pro	
				500					505					510			
35																	
	Thr	Phe	Asp	Ser	Leu	Asn	Ile	Thr	Ala	Ala	Ser	Leu	Asn	Asp	Asp	Gly	
			515					520					525				
	Leu	Asp	Asn	His	Thr	Ile	Leu	Leu	Tyr	Tyr	Ser	Thr	Ala	Ala	Ser	Ser	
40																	
		530					535					540					
	Leu	Ala	Val	Thr	Leu	Met	Ile	Ala	Ile	Phe	Val	Val	Tyr	Met	Val	Ser	
	545					550					555					560	
45																	
	Arg	Asp	Asn	Val	Ser	Cys	Ser	Ile	Cys	Leu							
					565					570							

<210> 96

<211> 548

50 <212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Influenza A virus A/Brisbane/59/2007(H1N1)

55

<400> 96

Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val

1

5

10

15

5

10

15

20

25

30

35

40

45

50

Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
 20 25 30
 5
 Leu Glu Asn Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
 35 40 45
 Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 50 55 60
 10
 Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
 65 70 75 80
 Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
 85 90 95
 15
 Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
 100 105 110
 Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
 115 120 125
 20
 Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
 130 135 140
 25
 Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
 145 150 155 160
 Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
 165 170 175
 30
 Gly Val His His Pro Pro Asn Ile Gly Asn Gln Lys Ala Leu Tyr His
 180 185 190
 Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys
 195 200 205
 35
 Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
 210 215 220
 Arg Ile JI.sn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile
 225 230 235 240
 40
 Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu
 245 250 255
 Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp
 260 265 270
 45
 Lys Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser
 275 280 285
 50
 Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro Lys
 290 295 300

Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile
305 310 315 320

55

Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
325 330 335

Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His
 340 345 350
 Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln
 5 355 360 365
 Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys
 370 375 380
 10 Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu
 385 390 395 400
 Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp
 405 410 415
 15
 Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
 420 425 430
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 20 435 440 445
 Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 450 455 460
 25 Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn
 465 470 475 480
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 485 490 495
 30 Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
 500 505 510
 Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
 515 520 525
 35
 Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540
 Arg Ile Cys Ile
 40 545

<210> 97

<211> 548

<212> PRT

45 <213> Artificial Sequence

<220>

<223> Synthesized Influenza A virus A/Solomon Islands/3/2006(H1N1)

50 <400> 97

Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val
1				5					10					15	

Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

5			20					25						30	
---	--	--	----	--	--	--	--	----	--	--	--	--	--	----	--

Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala
		35					40					45			

10

15

20

25

30

35

40

45

50

Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 50 55 60
 Pro Glu Cys Glu Leu Leu Ile Ser Arg Glu Ser Trp Ser Tyr Ile Val
 5 65 70 75 80
 Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
 85 90 95
 10 Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
 100 105 110
 Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Thr
 115 120 125
 15 Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
 130 135 140
 Lys Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
 145 150 155 160
 20 Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
 165 170 175
 Gly Val His His Pro Pro Asn Ile Gly Asp Gln Arg Ala Leu Tyr His
 180 185 190
 25 Lys Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys
 195 200 205
 Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
 30 210 215 220
 Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile
 225 230 235 240
 Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu
 35 245 250 255
 Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp
 260 265 270
 40 Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser
 275 280 285
 Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro Lys
 290 295 300
 45 Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile
 305 310 315 320
 Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
 325 330 335
 50 Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His
 340 345 350

Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr Gln
355 360 365
Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys

	370	375	380
	Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu		
	385	390	395 400
5	Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp		
	405	410	415
	Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg		
10	420	425	430
	Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val		
	435	440	445
15	Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe		
	450	455	460
	Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn		
	465	470	475 480
20	Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg		
	485	490	495
	Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile		
25	500	505	510
	Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser		
	515	520	525
	Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys		
30	530	535	540
	Arg Ile Cys Ile		
	545		
35	<210> 98		
	<211> 548		
	<212> PRT		
	<213> Artificial Sequence		
40	<220>		
	<223> Synthesized Influenza A virus A/New Caledonia/20/1999(H1N1)		

<400> 98

45

50

55

	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val
	1				5					10					15	
	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu
5					20				25					30		
	Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala
			35					40					45			
10	Pro	Leu	Gln	Leu	Gly	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly	Asn
	50					55						60				
	Pro	Glu	Cys	Glu	Leu	Leu	Ile	Ser	Lys	Glu	Ser	Trp	Ser	Tyr	Ile	Val
	65					70					75					80
15																
	Glu	Thr	Pro	Asn	Pro	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	Tyr	Phe	Ala

20

25

30

35

40

45

50

					85						90						95
	Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	
					100				105						110		
5																	
	Arg	Phe	Glu	Ile	Phe	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Val	
					115			120					125				
	Thr	Gly	Val	Ser	Ala	Ser	Cys	Ser	His	Asn	Gly	Lys	Ser	Ser	Phe	Tyr	
10					130			135				140					
	Arg	Asn	Leu	Leu	Trp	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	
	145					150					155					160	
	Ser	Lys	Ser	Tyr	Val	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	
15						165				170					175		
	Gly	Val	His	His	Pro	Pro	Asn	Ile	Gly	Asn	Gln	Arg	Ala	Leu	Tyr	His	
					180				185						190		
20	Thr	Glu	Asn	Ala	Tyr	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Arg	
			195					200					205				
	Phe	Thr	Pro	Glu	Ile	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Gln	Glu	Gly	
		210					215					220					
25	Arg	Ile	Asn	Tyr	Tyr	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	
	225					230					235					240	
	Phe	Glu	Ala	Asn	Gly	Asn	Leu	Ile	Ala	Pro	Trp	Tyr	Ala	Phe	Ala	Leu	
				245					250						255		
30																	
	Ser	Arg	Gly	Phe	Gly	Ser	Gly	Ile	Ile	Thr	Ser	Asn	Ala	Pro	Met	Asp	
				260					265					270			
	Glu	Cys	Asp	Ala	Lys	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	
			275					280					285				
35																	
	Leu	Pro	Phe	Gln	Asn	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	
		290					295					300					
	Tyr	Val	Arg	Ser	Ala	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	
40		305				310					315					320	
	Pro	Ser	Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	
				325					330					335			
45																	
	Glu	Gly	Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	
				340					345					350			
	Gln	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	
			355					360					365				
50																	
	Asn	Ala	Ile	Asn	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	
		370					375					380					
	Met	Asn	Thr	Gln	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu	
	385				390						395				400		

Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
405 410 415

Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
 420 425 430
 5 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 435 440 445
 Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 450 455 460
 10 Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys Asn
 465 470 475 480
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 485 490 495
 15 Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
 500 505 510
 Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
 20 515 520 525
 Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540
 25 Arg Ile Cys Ile
 545

<210> 99

<211> 547

30 <212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized Influenza A virus A/Singapore/1/1957(H2N2)

35

<400> 99

40

45

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Lys	Val
	1				5					10					15	
5	Asp	Thr	Ile	Leu	Glu	Arg	Asn	Val	Thr	Val	Thr	His	Ala	Lys	Asp	Ile
				20					25					30		
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Lys	Leu	Asn	Gly	Ile	Pro
			35					40					45			
10	Pro	Leu	Glu	Leu	Gly	Asp	Cys	Ser	Ile	Ala	Gly	Trp	Leu	Leu	Gly	Asn
		50					55					60				
	Pro	Glu	Cys	Asp	Arg	Leu	Leu	Ser	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Met
	65					70					75					80
15																
	Glu	Lys	Glu	Asn	Pro	Arg	Asp	Gly	Leu	Cys	Tyr	Pro	Gly	Ser	Phe	Asn
				85						90					95	
	Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Ser	Val	Lys	His	Phe	Glu
20				100					105					110		
	Lys	Val	Lys	Ile	Leu	Pro	Lys	Asp	Arg	Trp	Thr	Gln	His	Thr	Thr	Thr
			115					120					125			

25

30

35

40

45

50

	Gly	Gly	Ser	Arg	Ala	Cys	Ala	Val	Ser	Gly	Asn	Pro	Ser	Phe	Phe	Arg	
	130							135					140				
5	Asn	Met	Val	Trp	Leu	Thr	Lys	Lys	Glu	Ser	Asn	Tyr	Pro	Val	Ala	Lys	
	145							150					155				160
	Gly	Ser	Tyr	Asn	Asn	Thr	Ser	Gly	Glu	Gln	Met	Leu	Ile	Ile	Trp	Gly	
					165						170				175		
10	Val	His	His	Pro	Asn	Asp	Glu	Thr	Glu	Gln	Arg	Thr	Leu	Tyr	Gln	Asn	
					180						185				190		
	Val	Gly	Thr	Tyr	Val	Ser	Val	Gly	Thr	Ser	Thr	Leu	Asn	Lys	Arg	Ser	
			195					200					205				
15	Thr	Pro	Asp	Ile	Ala	Thr	Arg	Pro	Lys	Val	Asn	Gly	Leu	Gly	Ser	Arg	
					210				215				220				
	Met	Glu	Phe	Ser	Trp	Thr	Leu	Leu	Asp	Met	Trp	Asp	Thr	Ile	Asn	Phe	
20	225						230				235					240	
	Glu	Ser	Thr	Gly	Asn	Leu	Ile	Ala	Pro	Glu	Tyr	Gly	Phe	Lys	Ile	Ser	
					245					250					255		
	Lys	Arg	Gly	Ser	Ser	Gly	Ile	Met	Lys	Thr	Glu	Gly	Thr	Leu	Glu	Asn	
25					260				265					270			
	Cys	Glu	Thr	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Ala	Ile	Asn	Thr	Thr	Leu	
			275					280					285				
30	Pro	Phe	His	Asn	Val	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	
		290					295				300						
	Val	Lys	Ser	Glu	Lys	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Val	Pro	
	305					310					315					320	
35	Gln	Ile	Glu	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	
					325					330				335			
	Gly	Gly	Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Ser	
40				340					345					350			
	Asn	Asp	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Lys	Glu	Ser	Thr	Gln	Lys	
			355					360					365				
	Ala	Phe	Asp	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	Met	
45					370				375					380			
	Asn	Thr	Gln	Phe	Glu	Ala	Val	Gly	Lys	Glu	Phe	Ser	Asn	Leu	Glu	Arg	
	385					390					395					400	
50	Arg	Leu	Glu	Asn	Leu	Asn	Lys	Lys	Met	Glu	Asp	Gly	Phe	Leu	Asp	Val	

				405						410					415	
Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Met	Glu	Asn	Glu	Arg	Thr	
			420					425					430			
Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Asp	Lys	Val	Arg	
			435					440					445			

55

Met Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu
 450 455 460
 Phe Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly
 5 465 470 475 480
 Thr Tyr Asp Tyr Pro Lys Tyr Glu Glu Glu Ser Lys Leu Asn Arg Asn
 485 490 495
 10 Glu Ile Lys Gly Val Lys Leu Ser Ser Met Gly Val Tyr Gln Ile Leu
 500 505 510
 Ala Ile Tyr Ala Thr Val Ala Gly Ser Leu Ser Leu Ala Ile Met Met
 515 520 525
 15 Ala Gly Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg
 530 535 540
 Ile Cys Ile
 20 545

<210> 100

<211> 541

<212> PRT

25 **<213> Artificial Sequence**

<220>

<223> Synthesized [Influenza A virus A/Brisbane/10/2007(H3N2)]

30 **<400> 100**

35

40

45

	Thr	Ala	Thr	Leu	Cys	Leu	Gly	His	His	Ala	Val	Pro	Asn	Gly	Thr	Ile
	1				5					10					15	
	Val	Lys	Thr	Ile	Thr	Asn	Asp	Gln	Ile	Glu	Val	Thr	Asn	Ala	Thr	Glu
5				20				25						30		
	Leu	Val	Gln	Ser	Ser	Ser	Thr	Gly	Glu	Ile	Cys	Asp	Ser	Pro	His	Gln
			35					40					45			
10	Ile	Leu	Asp	Gly	Glu	Asn	Cys	Thr	Leu	Ile	Asp	Ala	Leu	Leu	Gly	Asp
		50					55					60				
	Pro	Gln	Cys	Asp	Gly	Phe	Gln	Asn	Lys	Lys	Trp	Asp	Leu	Phe	Val	Glu
	65					70					75					80
15																
	Arg	Ser	Lys	Ala	Tyr	Ser	Asn	Cys	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr
					85					90					95	
	Ala	Ser	Leu	Arg	Ser	Leu	Val	Ala	Ser	Ser	Gly	Thr	Leu	Glu	Phe	Asn
20																
				100						105					110	
	Asn	Glu	Ser	Phe	Asn	Trp	Thr	Gly	Val	Thr	Gln	Asn	Gly	Thr	Ser	Ser
			115					120					125			
	Ala	Cys	Ile	Arg	Arg	Ser	Asn	Asn	Ser	Phe	Phe	Ser	Arg	Leu	Asn	Trp
25																
				130				135					140			
	Leu	Thr	His	Leu	Lys	Phe	Lys	Tyr	Pro	Ala	Leu	Asn	Val	Thr	Met	Pro
	145					150					155					160

30

35

40

45

50

	Asn	Asn	Glu	Lys	Phe	Asp	Lys	Leu	Tyr	Ile	Trp	Gly	Val	His	His	Pro	
					165					170						175	
	Gly	Thr	Asp	Asn	Asp	Gln	Ile	Phe	Leu	Tyr	Ala	Gln	Ala	Ser	Gly	Arg	
5					180					185						190	
	Ile	Thr	Val	Ser	Thr	Lys	Arg	Ser	Gln	Gln	Thr	Val	Ile	Pro	Asn	Ile	
			195					200						205			
10	Gly	Ser	Arg	Pro	Arg	Val	Arg	Asn	Ile	Pro	Ser	Arg	Ile	Ser	Ile	Tyr	
		210					215						220				
	Trp	Thr	Ile	Val	Lys	Pro	Gly	Asp	Ile	Leu	Leu	Ile	Asn	Ser	Thr	Gly	
	225					230					235					240	
15	Asn	Leu	Ile	Ala	Pro	Arg	Gly	Tyr	Phe	Lys	Ile	Arg	Ser	Gly	Lys	Ser	
					245					250						255	
	Ser	Ile	Met	Arg	Ser	Asp	Ala	Pro	Ile	Gly	Lys	Cys	Asn	Ser	Glu	Cys	
			260						265						270		
20																	
	Ile	Thr	Pro	Asn	Gly	Ser	Ile	Pro	Asn	Asp	Lys	Pro	Phe	Gln	Asn	Val	
			275						280							285	
	Asn	Arg	Ile	Thr	Tyr	Gly	Ala	Cys	Pro	Arg	Tyr	Val	Lys	Gln	Asn	Thr	
		290					295						300				
25																	
	Leu	Lys	Leu	Ala	Thr	Gly	Met	Arg	Asn	Val	Pro	Glu	Lys	Gln	Thr	Arg	
	305					310					315					320	
	Gly	Ile	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Asn	Gly	Trp	Glu	Gly	
30					325					330						335	
	Met	Val	Asp	Gly	Trp	Tyr	Gly	Phe	Arg	His	Gln	Asn	Ser	Glu	Gly	Ile	
			340						345						350		
	Gly	Gln	Ala	Ala	Asp	Leu	Lys	Ser	Thr	Gln	Ala	Ala	Ile	Asp	Gln	Ile	
35			355						360							365	
	Asn	Gly	Lys	Leu	Asn	Arg	Leu	Ile	Gly	Lys	Thr	Asn	Glu	Lys	Phe	His	
		370					375						380				
40	Gln	Ile	Glu	Lys	Glu	Phe	Ser	Glu	Val	Glu	Gly	Arg	Ile	Gln	Asp	Leu	
	385				390						395					400	
	Glu	Lys	Tyr	Val	Glu	Asp	Thr	Lys	Ile	Asp	Leu	Trp	Ser	Tyr	Asn	Ala	
				405						410						415	
45	Glu	Leu	Leu	Val	Ala	Leu	Glu	Asn	Gln	His	Thr	Ile	Asp	Leu	Thr	Asp	
				420						425						430	
	Ser	Glu	Met	Asn	Lys	Leu	Phe	Glu	Lys	Thr	Lys	Lys	Gln	Leu	Arg	Glu	
			435						440						445		

50

Asn Ala Glu Asp Met Gly J:..sn Gly Cys Phe Lys Ile Tyr His Lys Cys
 450 455 460

Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp
 465 470 475 480

Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Val

				485					490					495					
				Glu	Leu	Lys	Ser	Gly	Tyr	Lys	Asp	Trp	Ile	Leu	Trp	Ile	Ser	Phe	Ala
							500					505				510			
5																			
				Ile	Ser	Cys	Phe	Leu	Leu	Cys	Val	Ala	Leu	Leu	Gly	Phe	Ile	Met	Trp
						515					520				525				
				Ala	Cys	Gln	Lys	Gly	Asn	Ile	Arg	Cys	Asn	Ile	Cys	Ile			
10					530					535					540				

<210> 101

<211> 541

<212> PRT

15 <213> Artificial Sequence

<220>

<223> Synthesized A/Wisconsin/67e5/2005(H3)

20 <400> 101

25

30

35

40

45

50

	Thr	Ala	Thr	Leu	Cys	Leu	Gly	His	His	Ala	Val	Pro	Asn	Gly	Thr	Ile	
	1				5					10					15		
	Val	Lys	Thr	Ile	Thr	Asn	Asp	Gln	Ile	Glu	Val	Thr	Asn	Ala	Thr	Glu	
5				20				25						30			
	Leu	Val	Gln	Ser	Ser	Ser	Thr	Gly	Gly	Ile	Cys	Asp	Ser	Pro	His	Gln	
			35					40					45				
10	Ile	Leu	Asp	Gly	Glu	Asn	Cys	Thr	Leu	Ile	Asp	Ala	Leu	Leu	Gly	Asp	
		50					55					60					
	Pro	Gln	Cys	Asp	Gly	Phe	Gln	Asn	Lys	Lys	Trp	Asp	Leu	Phe	Val	Glu	
	65					70					75					80	
15																	
	Arg	Ser	Lys	Ala	Tyr	Ser	Asn	Cys	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	
					85					90					95		
	Ala	Ser	Leu	Arg	Ser	Leu	Val	Ala	Ser	Ser	Gly	Thr	Leu	Glu	Phe	Asn	
20				100				105					110				
	Asp	Glu	Ser	Phe	Asn	Trp	Thr	Gly	Val	Thr	Gln	Asn	Gly	Thr	Ser	Ser	
			115					120					125				
25	Ala	Cys	Lys	Arg	Arg	Ser	Asn	Asn	Ser	Phe	Phe	Ser	Arg	Leu	Asn	Trp	
		130					135					140					
	Leu	Thr	His	Leu	Lys	Phe	Lys	Tyr	Pro	Ala	Leu	Asn	Val	Thr	Met	Pro	
	145					150					155					160	
30	Asn	Asn	Glu	Lys	Phe	Asp	Lys	Leu	Tyr	Ile	Trp	Gly	Val	His	His	Pro	
				165				170						175			
	Gly	Thr	Asp	Asn	Asp	Gln	Ile	Phe	Leu	His	Ala	Gln	Ala	Ser	Gly	Arg	
				180					185					190			
35																	
	Ile	Thr	Val	Ser	Thr	Lys	Arg	Ser	Gln	Gln	Thr	Val	Ile	Pro	Asn	Ile	
			195					200					205				
	Gly	Ser	Arg	Pro	Arg	Ile	Arg	Asn	Ile	Pro	Ser	Arg	Ile	Ser	Ile	Tyr	
40																	
45																	
50																	

210 215 220
 Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly
 225 230 235 240
 5
 Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys Ser
 245 250 255
 Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu Cys
 10 260 265 270
 Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn Val
 275 280 285
 Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn Thr
 15 290 295 300
 Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr Arg
 305 310 315 320
 20 Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly
 325 330 335
 Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Ile
 340 345 350
 25 Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asn Gln Ile
 355 360 365
 Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe His
 370 375 380
 30 Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp Leu
 385 390 395 400
 Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn Ala
 405 410 415
 35 Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp
 420 425 430
 Ser Glu Met Asn Lys Leu Phe Glu Arg Thr Lys Lys Gln Leu Arg Glu
 40 435 440 445
 Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys
 450 455 460
 45 Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp
 465 470 475 480
 Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Val
 485 490 495
 50 Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe Ala
 500 505 510

Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met Trp
515 520 525

55

Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
530 535 540

<210> 102

<211> 551

<212> PRT

<213> Artificial Sequence

5

<220>

<223> Synthesized Influenza A virus A/Anhui/1/2005(H5N1)

<400> 102

10

15

20

25

30

35

40

45

50

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val	
	1				5					10					15		
	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile	
5				20					25					30			
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys	
			35					40					45				
10	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn	
		50					55					60					
	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val	
	65					70					75					80	
15																	
	Glu	Lys	Ala	Asn	Pro	Ala	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Asn	Phe	Asn	
					85					90					95		
	Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu	
20				100					105					110			
	Lys	Ile	Gln	Ile	Ile	Pro	Lys	Ser	Ser	Trp	Ser	Asp	His	Glu	Ala	Ser	
			115					120					125				
25	Ser	Gly	Val	Ser	Ser	Ala	Cys	Pro	Tyr	Gln	Gly	Thr	Pro	Ser	Phe	Phe	
		130					135					140					
	Arg	Asn	Val	Val	Trp	Leu	Ile	Lys	Lys	Asn	Asn	Thr	Tyr	Pro	Thr	Ile	
	145					150					155					160	
30																	
	Lys	Arg	Ser	Tyr	Asn	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Ile	Leu	Trp	
					165					170					175		
	Gly	Ile	His	His	Ser	Asn	Asp	Ala	Ala	Glu	Gln	Thr	Lys	Leu	Tyr	Gln	
35				180					185					190			
	Asn	Pro	Thr	Thr	Tyr	Ile	Ser	Val	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg	
			195					200					205				
40	Leu	Val	Pro	Lys	Ile	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	Gly	
		210					215						220				
	Arg	Met	Asp	Phe	Phe	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn	
	225					230					235					240	
45	Phe	Glu	Ser	Asn	Gly	Asn	Phe	Ile	Ala	Pro	Glu	Tyr	Ala	Tyr	Lys	Ile	
				245						250				255			
	Val	Lys	Lys	Gly	Asp	Ser	Ala	Ile	Val	Lys	Ser	Glu	Val	Glu	Tyr	Gly	
				260					265					270			

50

55

	Asn	Cys	Asn	Thr	Lys	Cys	Gln	Thr	Pro	Ile	Gly	Ala	Ile	Asn	Ser	Ser	
			275					280					285				
	Met	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	
5		290					295					300					
	Tyr	Val	Lys	Ser	Asn	Lys	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser	
	305					310					315					320	
10	Pro	Leu	Arg	Glu	Arg	Arg	Arg	Lys	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	
					325					330					335		
	Gly	Phe	Ile	Glu	Gly	Gly	Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	
				340					345					350			
15	Tyr	His	His	Ser	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Lys	Glu	
			355					360					365				
	Ser	Thr	Gln	Lys	Ala	Ile	Asp	Gly	Val	Thr	Asn	Lys	Val	Asn	Ser	Ile	
20		370					375					380					
	Ile	Asp	Lys	Met	Asn	Thr	Gln	Phe	Glu	Ala	Val	Gly	Arg	Glu	Phe	Asn	
	385					390					395					400	
25	Asn	Leu	Glu	Arg	Arg	Ile	Glu	Asn	Leu	Asn	Lys	Lys	Met	Glu	Asp	Gly	
					405					410					415		
	Phe	Leu	Asp	Val	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Met	Glu	
				420					425					430			
30	Asn	Glu	Arg	Thr	Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	
			435				440						445				
	Asp	Lys	Val	Arg	Leu	Gln	Leu	Arg	Asp	Asn	Ala	Lys	Glu	Leu	Gly	Asn	
35		450					455					460					
	Gly	Cys	Phe	Glu	Phe	Tyr	His	Lys	Cys	Asp	Asn	Glu	Cys	Met	Glu	Ser	
	465					470					475					480	
	Val	Arg	Asn	Gly	Thr	Tyr	Asp	Tyr	Pro	Gln	Tyr	Ser	Glu	Glu	Ala	Arg	
40					485					490					495		
	Leu	Lys	Arg	Glu	Glu	Ile	Ser	Gly	Val	Lys	Leu	Glu	Ser	Ile	Gly	Thr	
				500					505					510			
45	Tyr	Gln	Ile	Leu	Ser	Ile	Tyr	Ser	Thr	Val	Ala	Ser	Ser	Leu	Ala	Leu	
		515						520					525				
	Ala	Ile	Met	Val	Ala	Gly	Leu	Ser	Leu	Trp	Met	Cys	Ser	Asn	Gly	Ser	
		530					535					540					
50	Leu	Gln	Cys	Arg	Ile	Cys	Ile										
	545					550											

55

<210> 103

<211> 552

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthesized [Influenza A virus ANietnam/1194/2004(H5N1)

<400> 103

5

10

15

20

25

30

35

40

45

50

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val
	1			5						10					15	
5	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile
				20					25					30		
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys
			35					40					45			
10	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn
		50				55						60				
	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val
	65				70					75					80	
15	Glu	Lys	Ala	Asn	Pro	Val	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Asp	Phe	Asn
				85					90					95		
	Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu
20				100					105					110		
	Lys	Ile	Gln	Ile	Ile	Pro	Lys	Ser	Ser	Trp	Ser	Ser	His	Glu	Ala	Ser
				115				120					125			
25	Leu	Gly	Val	Ser	Ser	Ala	Cys	Pro	Tyr	Gln	Gly	Lys	Ser	Ser	Phe	Phe
		130				135						140				
	Arg	Asn	Val	Val	Trp	Leu	Ile	Lys	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile
	145				150					155				160		
30	Lys	Arg	Ser	Tyr	Asn	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp
				165					170					175		
	Gly	Ile	His	His	Pro	Lys	Asp	Ala	Ala	Glu	Gln	Thr	Lys	Leu	Tyr	Gln
35				180					185					190		
	Asn	Pro	Thr	Thr	Tyr	Ile	Ser	Val	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg
				195				200					205			
40	Leu	Val	Pro	Arg	Ile	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	Gly
		210						215					220			

	Arg	Met	Glu	Phe	Phe	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn	
	225							230						235			240
45	Phe	Glu	Ser	Asn	Gly	Asn	Phe	Ile	Ala	Pro	Glu	Tyr	Ala	Tyr	Lys	Ile	
					245						250					255	
	Val	Lys	Lys	Gly	Asp	Ser	Thr	Ile	Met	Lys	Ser	Glu	Leu	Glu	Tyr	Gly	
					260					265					270		
50	Asn	Cys	Asn	Thr	Lys	Cys	Gln	Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser	
					275					280					285		
	Met	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	
55					290					295					300		

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
 305 310 315 320
 Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
 5 325 330 335
 Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
 340 345 350
 10 Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
 355 360 365
 Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
 370 375 380
 15 Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
 385 390 395 400
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
 20 405 410 415
 Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
 420 425 430
 25 Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
 435 440 445
 Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
 450 455 460
 30 Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
 465 470 475 480
 Ser Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala
 485 490 495
 35 Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
 500 505 510
 Ile Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
 40 515 520 525

Leu Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly

530

535

540

45

Ser Leu Gln Cys Arg Ile Cys Ile

545

550

<210> 104

<211> 553

50

<212> PRT

<213> Artificial sequence

<220>

<223> Synthesized A/Indonesia/5/2005(H5N1)

55

<400> 104

Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val
1				5					10					15	

5

10

15

20

25

30

35

40

45

50

	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile	
				20					25					30			
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys	
5				35				40					45				
	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn	
		50					55					60					
10	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val	
	65					70					75					80	
	Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Ser	Phe	Asn	
					85					90					95		
15	Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu	
					100				105					110			
	Lys	Ile	Gln	Ile	Ile	Pro	Lys	Ser	Ser	Trp	Ser	Asp	His	Glu	Ala	Ser	
			115					120					125				
20	Ser	Gly	Val	Ser	Ser	Ala	Cys	Pro	Tyr	Leu	Gly	Ser	Pro	Ser	Phe	Phe	
			130				135					140					
	Arg	Asn	Val	Val	Trp	Leu	Ile	Lys	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile	
	145					150					155					160	
25	Lys	Lys	Ser	Tyr	Asn	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp	
					165					170					175		
	Gly	Ile	His	His	Pro	Asn	Asp	Ala	Ala	Glu	Gln	Thr	Arg	Leu	Tyr	Gln	
30					180				185					190			
	Asn	Pro	T r	Thr	Tyr	Ile	Ser	Ile	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg	
			195					200					205				
	Leu	Val	Pro	Lys	Ile	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	Gly	
35					210			215				220					
	Arg	Met	Glu	Phe	Phe	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn	
	225					230					235					240	
40	Phe	Glu	Ser	Asn	Gly	Asn	Phe	Ile	Ala	Pro	Glu	Tyr	Ala	Tyr	Lys	Ile	
					245					250				255			
	Val	Lys	Lys	Gly	Asp	Ser	Ala	Ile	Met	Lys	Ser	Glu	Leu	Glu	Tyr	Gly	
				260					265				270				
45	Asn	Cys	Asn	Thr	Lys	Cys	Gln	Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser	
					275			280				285					
	Met	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	
		290					295					300					
50	Tyr	Val	Lys	Ser	Asn	Arg	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser	
	305					310					315					320	

Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
325 330 335

55

Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr

	340	345	350
	Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys		
	355	360	365
5	Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser		
	370	375	380
	Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe		
10	385	390	395 400
	Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp		
	405	410	415
15	Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met		
	420	425	430
	Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu		
	435	440	445
20	Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly		
	450	455	460
	Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu		
25	465	470	475 480
	Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala		
	485	490	495
30	Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly		
	500	505	510
	Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala		
	515	520	525
35	Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly		
	530	535	540
	Ser Leu Gln Cys Arg Ile Cys Ile Lys		
	545	550	
40			

<211> 548

<212> PRT

45

<213> Artificial Sequence

<220>

<223> Construct 900

50

<400> 105

55

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
1 5 10 15

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
5 20 25 30

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
35 40 45

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn

15

20

25

30

35

40

45

50

	50	55	60
	Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val 65 70 75 80		
5	Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Tyr Asp Val Pro 85 90 95		
	Asp Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu 100 105 110		
10	Phe Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr 115 120 125		
	Ser Ser Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu 130 135 140		
15	Asn Trp Leu Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr 145 150 155 160		
	Met Pro Asn Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly Val His 165 170 175		
	His Pro Gly Thr Asp Asn Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser 180 185 190		
25	Gly Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ile Pro 195 200 205		
	Asn Ile Gly Ser Arg Pro Arg Val Arg Asn Ile Pro Ser Arg Ile Ser 210 215 220		
30	Ile Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser 225 230 235 240		
	Thr Gly Asn Leu Ile Ala Pro Arg Cly Tyr Phe Lys Ile Arg Ser Gly 245 250 255		
35	Lys Ser Ser Ile Met Lys Ser Glu Leu Glu Tyr Gly Asn Cys Asn Thr 260 265 270		
	Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser Met Pro Phe His 275 280 285		
40	Asn Ile His Pro Leu Thr Ile Cly Glu Cys Pro Lys Tyr Val Lys Ser 290 295 300		
	Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu 305 310 315 320		
45	Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile 325 330 335		
	Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Cly Tyr His His 340 345 350		
50	Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln 355 360 365		
55	Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys 370 375 380		

	Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu	
	385	390 395 400
5	Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp	
	405	410 415
	Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg	
	420	425 430
10	Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val	
	435	440 445
	Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe	
	450	455 460
15	Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn	
	465	470 475 480
	Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg	
20	485	490 495
	Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile	
	500	505 510
25	Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met	
	515	520 525
	Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys	
	530	535 540
30	Arg Ile Cys Ile	
	545	
	<210> 106	
35	<211> 565	
	<212> PRT	
	<213> Artificial Sequence	
	<220>	
40	<223> Construct 745	

<400> 106

45

50

55

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val	
	1				5					10					15		
5	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile	
				20					25					30			
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys	
			35					40					45				
10	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn	
		50					55					60					
	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val	
	65					70					75				80		
15	Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Ile	Met	His	Asp	
					85					90					95		
20																	
25																	
30																	
35																	
40																	
45																	
50																	
55																	

	Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu Asn	
	100	105 110
5	Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly	
	115	120 125
	Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser	
	130	135 140
10	Lys Ser Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn	
	145	150 155 160
	Asn Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys	
	165	170 175
15	Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn	
	180	185 190
	Lys Thr Gln Met Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe	
	195	200 205
20	Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly	
	210	215 220
	Ser Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg	
25	225	230 235 240
	Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile	
	245	250 255
30	Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser	
	260	265 270
	Gly Arg Ser Lys Ile Met Lys Ser Glu Leu Glu Tyr Gly Asn Cys Asn	
	275	280 285
35	Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser Met Pro Phe	
	290	295 300
	His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr Val Lys	
	305	310 315 320
40	Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg	

		325		330		335										
	Glu	Ser	Arg	Arg	Lys	Lys	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe
45					340				345					350		
	Ile	Glu	Gly	Gly	Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His
					355				360					365		
50	His	Ser	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Lys	Glu	Ser	Thr
					370				375					380		
	Gln	Lys	Ala	Ile	Asp	Gly	Val	Thr	Asn	Lys	Val	Asn	Ser	Ile	Ile	Asp
	385					390					395				400	
55	Lys	Met	Asn	Thr	Gln	Phe	Glu	Ala	Val	Gly	Arg	Glu	Phe	Asn	Asn	Leu
					405					410				415		

Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu
 420 425 430
 Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu
 435 440 445
 Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys
 450 455 460
 Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys
 465 470 475 480
 Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg
 485 490 495
 Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys
 500 505 510
 Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln
 515 520 525
 Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile
 530 535 540
 Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln
 545 550 555 560
 Cys Arg Ile Cys Ile
 565

<210> 107

<211> 545

<212> PRT

<213> Artificial Sequence

<220>

<223> Construct 910

<400> 107

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val	
	1				5					10					15		
5	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile	
				20					25					30			
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Ser	Pro	His	Gln	Ile	
			35					40					45				
10	Leu	Asp	Gly	Glu	Asn	Cys	Thr	Leu	Ile	Asp	Ala	Leu	Leu	Gly	Asp	Pro	
	50						55					60					
	Gln	Cys	Asp	Gly	Phe	Gln	Asn	Lys	Lys	Trp	Asp	Leu	Phe	Val	Glu	Arg	
	65					70				75						80	
15																	
	Ser	Lys	Ala	Tyr	Ser	Asn	Cys	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	
					85					90					95		
	Ser	Leu	Arg	Ser	Leu	Val	Ala	Ser	Ser	Gly	Thr	Leu	Glu	Phe	Asn	Asn	
20					100				105					110			
25																	
30																	
35																	
40																	
45																	
50																	

Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser Ala
 115 120 125
 Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu
 5 130 135 140
 Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro Asn
 145 150 155 160
 10 Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly Val His His Pro Gly
 165 170 175
 Thr Asp Asn Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile
 180 185 190
 15 Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile Gly
 195 200 205
 Ser Arg Pro Arg Val Arg Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp
 210 215 220
 20 Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn
 225 230 235 240
 Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser
 245 250 255
 25 Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Thr Lys Cys Gln
 260 265 270
 Thr Pro Met Gly Ala Ile Asn Ser Ser Met Pro Phe His Asn Ile His
 30 275 280 285
 Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr Val Lys Ser Asn Arg Leu
 290 295 300
 Val Leu Ala Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Ser Arg Arg
 35 305 310 315 320
 Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly
 325 330 335
 40 Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Ser Asn Glu
 340 345 350
 Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala Ile
 355 360 365
 45 Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys Met Asn Thr
 370 375 380
 Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu Arg Arg Ile
 385 390 395 400
 50 Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp Val Trp Thr
 405 410 415

Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr Leu Asp
420 425 430

55

Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Leu Gln

435 440 445
 Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr
 450 455 460
 5
 His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn Gly Thr Tyr
 465 470 475 480
 Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg Glu Glu Ile
 10 485 490 495
 Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile Leu Ser Ile
 500 505 510
 15 Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly
 515 520 525
 Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
 530 535 540
 20
 Ile
 545

<210> 108

25 <211> 556

<212> PRT

<213> Artificial Sequence

<220>

30 <223> Construct 920

<400> 108

35

40

45

50

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val
	1				5					10					15	
5	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile
				20					25					30		
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Pro	Asp	Cys	Leu	Asn	Cys
			35					40					45			
10	Thr	Asp	Leu	Asp	Val	Ala	Leu	Gly	Arg	Pro	Met	Cys	Val	Gly	Thr	Thr
		50					55					60				
	Pro	Ser	Ala	Lys	Ala	Ser	Ile	Leu	His	Glu	Val	Lys	Pro	Val	Thr	Ser
	65					70					75					80
15																
	Gly	Cys	Phe	Pro	Ile	Met	His	Asp	Arg	Thr	Lys	Ile	Arg	Gln	Leu	Pro
					85					90					95	
	Asn	Leu	Leu	Arg	Gly	Tyr	Glu	Asn	Ile	Arg	Leu	Ser	Thr	Gln	Asn	Val
20				100					105					110		
	Ile	Asp	Ala	Glu	Lys	Ala	Pro	Gly	Gly	Pro	Tyr	Arg	Leu	Gly	Thr	Ser
			115					120					125			
25	Gly	Ser	Cys	Pro	Asn	Ala	Thr	Ser	Lys	Ser	Gly	Phe	Phe	Ala	Thr	Met
		130					135					140				
	Ala	Trp	Ala	Val	Pro	Lys	Asp	Asn	Asn	Lys	Asn	Ala	Thr	Asn	Pro	Leu

30

35

40

45

50

	145		150		155		160
	Thr Val Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr						
		165			170		175
5							
	Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met Lys Asn Leu Tyr						
		180		185			190
	Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr						
10		195		200			205
	Thr His Tyr Val Ser Gln Ile Gly Ser Phe Pro Asp Gln Thr Glu Asp						
		210		215		220	
	Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln						
15		225		230		235	240
	Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu						
		245			250		255
20							
	Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly						
		260		265			270
	Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Gln Thr Pro Met Gly Ala						
		275		280			285
25							
	Ile Asn Ser Ser Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly						
		290		295		300	
	Glu Cys Pro Lys Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly						
		305		310		315	320
30							
	Leu Arg Asn Ser Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu						
		325		330			335
	Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val						
		340		345			350
35							
	Asp Gly Trp Tyr Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr						
		355		360			365
	Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn						
40		370		375			380
	Lys Val Asn Ser Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val						
		385		390		395	400
45							
	Gly Arg Glu Phe Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys						
		405		410			415
	Lys Met Glu Asp Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu						
		420		425			430
50							
	Leu Val Leu Met Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn						
		435		440			445

Val Lys Asn Leu Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala
450 455 460

55

Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
465 470 475 480

	Glu	Cys	Met	Glu	Ser	Ile	Arg	Asn	Gly	Thr	Tyr	Asn	Tyr	Pro	Gln	Tyr	
							485					490				495	
	Ser	Glu	Glu	Ala	Arg	Leu	Lys	Arg	Glu	Glu	Ile	Ser	Gly	Val	Lys	Leu	
5							500					505				510	
	Glu	Ser	Ile	Gly	Thr	Tyr	Gln	Ile	Leu	Ser	Ile	Tyr	Ser	Thr	Val	Ala	
				515					520					525			
10	Ser	Ser	Leu	Ala	Leu	Ala	Ile	Met	Met	Ala	Gly	Leu	Ser	Leu	Trp	Met	
			530					535						540			
	Cys	Ser	Asn	Gly	Ser	Leu	Gln	Cys	Arg	Ile	Cys	Ile					
	545						550				555						

15

<210> 109

<211> 548

<212> PRT

<213> Artificial Sequence

20

<220>

<223> Construct 930

<400> 109

25

30

35

40

45

50

	Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val
	1				5					10					15	
	Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu
5				20					25					30		
	Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys
			35					40					45			
10	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn
		50					55					60				
	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val
	65					70					75					80
15																
	Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Ser	Phe	Asn
				85						90					95	
	Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu
20				100					105					110		
	Lys	Ile	Gln	Ile	Ile	Pro	Lys	Ser	Ser	Trp	Ser	Asp	His	Glu	Ala	Ser
			115					120					125			
25	Ser	Gly	Val	Ser	Ser	Ala	Cys	Pro	Tyr	Leu	Gly	Ser	Pro	Ser	Phe	Phe
		130					135					140				
	Arg	Asn	Val	Val	Trp	Leu	Ile	Lys	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile
	145					150					155					160
30	Lys	Lys	Ser	Tyr	Asn	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp
					165					170					175	
	Gly	Ile	His	His	Pro	Asn	Asp	Ala	Ala	Glu	Gln	Thr	Arg	Leu	Tyr	Gln
				180					185						190	

35

40

45

50

	Asn	Pro	Thr	Thr	Tyr	Ile	Ser	Ile	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg	
	Leu	Val	Pro	Lys	Ile	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	G:'.y	
5																	
	Arg	Met	Glu	Phe	Phe	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn	
	225																
10	Phe	Glu	Ser	Asn	Gly	Asn	Phe	Ile	Ala	Fro	Glu	Tyr	Ala	Tyr	Lys	Ile	
	Val	Lys	Lys	Gly	Asp	Ser	Ala	Ile	Met	Lys	Ser	Glu	Leu	Glu	Tyr	Gly	
15																	
	Asn	Cys	Asp	Ala	Lys	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	
	Leu	Pro	Phe	Gln	Asn	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	
	290																
20																	
	Tyr	Val	Arg	Ser	Ala	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	
	305																
	Pro	Ser	Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	
25																	
	Glu	Gly	Sly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	
30																	
	Gln	Asn	Slu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln	
	Asn	Ala	Ile	Asn	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys	
	370																
35																	
	Met	Asn	Thr	Gln	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu	
	385																
	Arg	Arg	Met	Glu	Asn	Leu	Asn	Lys	Lys	Val	Asp	Asp	Gly	Phe	Leu	Asp	
40																	
	Ile	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu	Arg	
	Thr	Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys	Val	
45																	
	Lys	Ser	Gln	Leu	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys	Phe	
	450																
50	Glu	Phe	Tyr	His	Lys	Cys	Asn	Asn	Glu	Cys	Met	Glu	Ser	Val	Lys	Asn	
	465																

Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
485 490 495

55

Glu Lys Ile Asp Sly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
500 505 510

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser

515

520

525

Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys

5

530

535

540

Arg Ile Cys Ile

545

10

<210> 110

<400> 110

000

15

<210> 111

<211> 552

<212> PRT

<213> Artificial Sequence

20

<220>

<223> Construct 690 and 734

<400> 111

25

30

35

40

45

	Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val	
	1				5					10					15		
	Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile	
5					20				25					30			
	Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys	
			35					40					45				
10	Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn	
		50					55					60					
	Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val	
	65					70					75					80	
15																	
	Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	His	Phe	Ala	
					85					90					95		
	Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	
20					100				105					110			
	Arg	Phe	Glu	Ile	Phe	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Val	
			115					120					125				
	Thr	Gly	Val	Ser	Ala	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe	Tyr	
25					130			135				140					
	Arg	Asn	Leu	Leu	Trp	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	
	145					150					155					160	
30	Ser	Lys	Ser	Tyr	Ala	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	
					165					170				175			
	Gly	Val	His	His	Pro	Pro	Asn	Ile	Gly	Asp	Gln	Lys	Ala	Leu	Tyr	His	
				180					185					190			
35																	
	Thr	Glu	Asn	Ala	Tyr	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	
40																	
45																	
50																	

55

195

200

205

Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
 210 215 220

5

Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile
 225 230 235 240

Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu

10

245

250

255

Ser Arg Gly Phe Gly Ser Gly Ile Met Lys Ser Glu Leu Glu Tyr Gly
 260 265 270

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser

15

275

280

285

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
 290 295 300

20

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
 305 310 315 320

Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
 325 330 335

25

Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr

340

345

350

Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
 355 360 365

30

Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
 370 375 380

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
 385 390 395 400

35

Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Val Glu Asp
 405 410 415

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
 420 425 430

40

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
 435 440 445

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
 450 455 460

45

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
 465 470 475 480

50

Ser Ile Arg Ile Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala

485

490

495

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
500 505 510

55

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
515 520 525

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly

530

535

540

Ser Leu Gln Cys Arg Ile Cys Ile

5

545

550

<210> 112

<211> 548

<212> PRT

10

<213> Artificial Sequence

<220>

<223> Construct 696

15

<400> 112

20

25

30

35

40

45

50

Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
1 5 10 15

5

Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
20 25 30

Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
35 40 45

10

Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
50 55 60

Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
65 70 75 80

15

Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Ser Phe Asn
85 90 95

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
100 105 110

20

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
115 120 125

25

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
130 135 140

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
145 150 155 160

30

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
165 170 175

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
180 185 190

35

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
195 200 205

40

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
210 215 220

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
225 230 235 240

45

50

55

	Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile		
	245	250	255
5	Val Lys Lys Gly Asp Ser Ala Ile Ile Thr Ser Asn Ala Pro Met Asp		
	260	265	270
	Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser Ser		
	275	280	285
10	Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro Lys		
	290	295	300
	Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn Ile		
	305	310	315 320
15	Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile		
	325	330	335
	Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His His		
	340	345	350
20	Gln Asn Glu Gln Gly Ser Gly r,yr Ala Ala Asp Gln Lys Ser Thr Gln		
	355	360	365
	Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys		
25	370	375	380
	Met Asn T r Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu		
	385	390	395 400
30	Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp		
	405	410	415
	Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg		
	420	425	430
35	Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val		
	435	440	445
	Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe		
	450	455	460

40

Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys Asn
 465 470 475 480

Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 485 490 495

Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
 500 505 510

Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
 515 520 525

Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540

Arg Ile Cys Ile
 545

<210> 113

<211> 552

<212> PRT

<213> Artificial Sequence

5

<220>

<223> Construct 691

<400> 113

10

15

20

25

30

35

40

45

50

Asp Gln Ile Cys Ile Gly Tyr His A:_a Asn Asn Ser Thr Glu Gln Val
 1 5 10 15
 5 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 20 25 30
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
 35 40 45
 10 Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 50 55 60
 Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
 65 70 75 80
 15 Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
 85 90 95
 Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
 100 105 110
 20 Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
 115 120 125
 Thr Gly Val Ser Ala Ser Cys Ser H:_s Asn Gly Glu Ser Ser Phe Tyr
 130 135 140
 Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
 145 150 155 160
 30 Ser Lys Ser Tyr Ala Asn Asn Lys G:_u Lys Glu Val Leu Val Leu Trp
 165 170 175
 Gly Val His His Pro Pro Asn Ile G -Y Asp Gln Lys Ala Leu Tyr His
 180 185 190
 35 Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys
 195 200 205
 Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
 210 215 220
 40 Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile

225 230 235 240

45 Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu
245 250 255

Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp
260 265 270

50

55

Lys Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile
 Asn Ser Ser 275 280 285
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu
 Cys Pro Lys 290 295 300
 Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu
 Arg Asn Ser 305 310 315 320
 10
 Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala
 Ile
 325 330 335
 Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp
 Gly Trp Tyr 340 345 350
 15
 Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala
 Ala Asp Lys 355 360
 365
 Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn
 Ser
 20
 370 375 380
 Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly
 Arg Glu Phe 385 390 395 400
 25
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu
 Asp
 405 410 415
 Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu
 Val Leu Met 420 425 430
 30
 Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn
 Leu
 435 440 445
 Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys
 Glu Leu Gly 450 455 460
 35
 Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu
 Cys Met Glu 465 470 475 480
 Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu
 Ala
 40
 485 490 495
 Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu
 Ser Ile Gly 500 505 510

45 Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser
 Ser Leu Ala 515 520 525

 Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys
 Ser Asn Gly 530 535 540

 50

 Ser Leu Gln Cys Arg
 Ile Cys Ile 545 550

REIVINDICAÇÕES

1. Ácido nucleico que pode ser expresso numa planta hospedeira compreendendo uma ou mais regiões reguladoras operando numa planta, e operacionalmente ligado a uma sequência codificando um polipéptido de hemaglutinina (HA) de gripe quimérico compreendendo um agregado de domínio de haste (SDC), uma cabeça de agrupamento de domínio (HDC) e um grupo de domínios transmembranares (TDC), uma ou mais regiões reguladoras compreendendo um promotor e um 5'UTR, 3'UTR ou 5'UTR e 3'UTR, e em que a) o SDC compreende um subdomínio F'1, F'2 e F; b) o HDC compreende um subdomínio de ligação ao recetor (RB), E1 e E2; c) o CDT compreende um subdomínio transmembrana (TmD) e terminal C (CT); e
 - i) em que o subdomínio RB é de um primeiro polipéptido de HA de gripe e os subdomínios de SDC e E1 e E2 são de um segundo polipéptido de HA de gripe e,
 - ii) em que o primeiro polipéptido de HA de gripe é de gripe H1 ou H5 e o segundo polipéptido de HA de gripe é de gripe H1 ou H5 e o segundo polipéptido de HA de gripe é derivado de uma estirpe de gripe diferente do primeiro polipéptido de HA de gripe.
2. Ácido nucleico da reivindicação 1, em que:
 - i) a sequência que codifica o polipéptido de HA de gripe quimérico compreende ainda uma sequência peptídica de sinal selecionada do grupo de uma sequência peptídica de sinal nativa de HA e uma sequência peptídica de sinal de PDI de alfafa e/ou
 - ii) em que o 5'UTR, 3'UTR ou 5'UTR e 3'UTR são obtidos a partir de uma UTR de plastocianina ou UTR e/ou
 - iii) em que o promotor é obtido a partir de uma região reguladora de plastocianina, uma 1,5-bisfosfato de ribulose região reguladora carboxilase / oxigenase (RuBisCO), uma região reguladora da proteína de ligação clorofila a / b (CAB), uma região reguladora CaMV 35S, uma região reguladora da actina, uma região reguladora da ubiquitina, uma região reguladora da

triosofosfato isomerase 1, um fator de iniciação translacional 4A região reguladora e uma região reguladora ST-LS1.

3. Método de produzir partículas semelhantes a vírus da gripe quiméricas (VLPs) numa planta compreendendo:
 - a) introduzir o ácido nucleico da reivindicação 1 ou 2 na planta, ou sua porção, e
 - b) incubar a planta, ou porção da mesma, sob condições que permitem a expressão do ácido nucleico, produzindo assim as VLPs, e
 - c) opcionalmente colheita da planta e purificação das VLPs.
4. Método de acordo com a reivindicação 3, em que no passo de introdução (passo a), i) o ácido nucleico é introduzido na planta de um modo transitório ou ii) o ácido nucleico é introduzido na planta de modo a ser integrado de forma estável em um genoma.
5. Polipéptido codificado pelo ácido nucleico da reivindicação 1.
6. Partícula semelhante a vírus (VLP) compreendendo o polipéptido da reivindicação 5.
7. VLP da reivindicação 6 compreendendo ainda N-glicanos específicos de plantas, ou N-glicanos modificados.
8. Composição compreendendo uma dose eficaz da VLP da reivindicação 6 ou 7 e um veículo farmacologicamente aceitável.
9. Partícula semelhante a vírus, de acordo com a reivindicação 6 ou 7, ou composição de acordo com a reivindicação 8, para uso na indução de imunidade a uma infecção por vírus gripe em um indivíduo.

10. Partícula semelhante a vírus ou a composição para uso de acordo com a reivindicação 9, que é adequada para administração oral, intradérmica, intranasal, intramuscular, intraperitoneal, intravenosa ou subcutânea a um sujeito.
11. Planta, ou sua porção, compreendendo um polipéptido codificado pelo ácido nucleico da reivindicação 1.

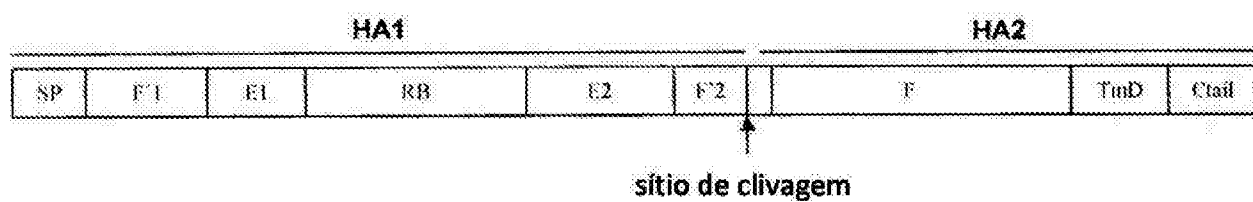


Figura 1a

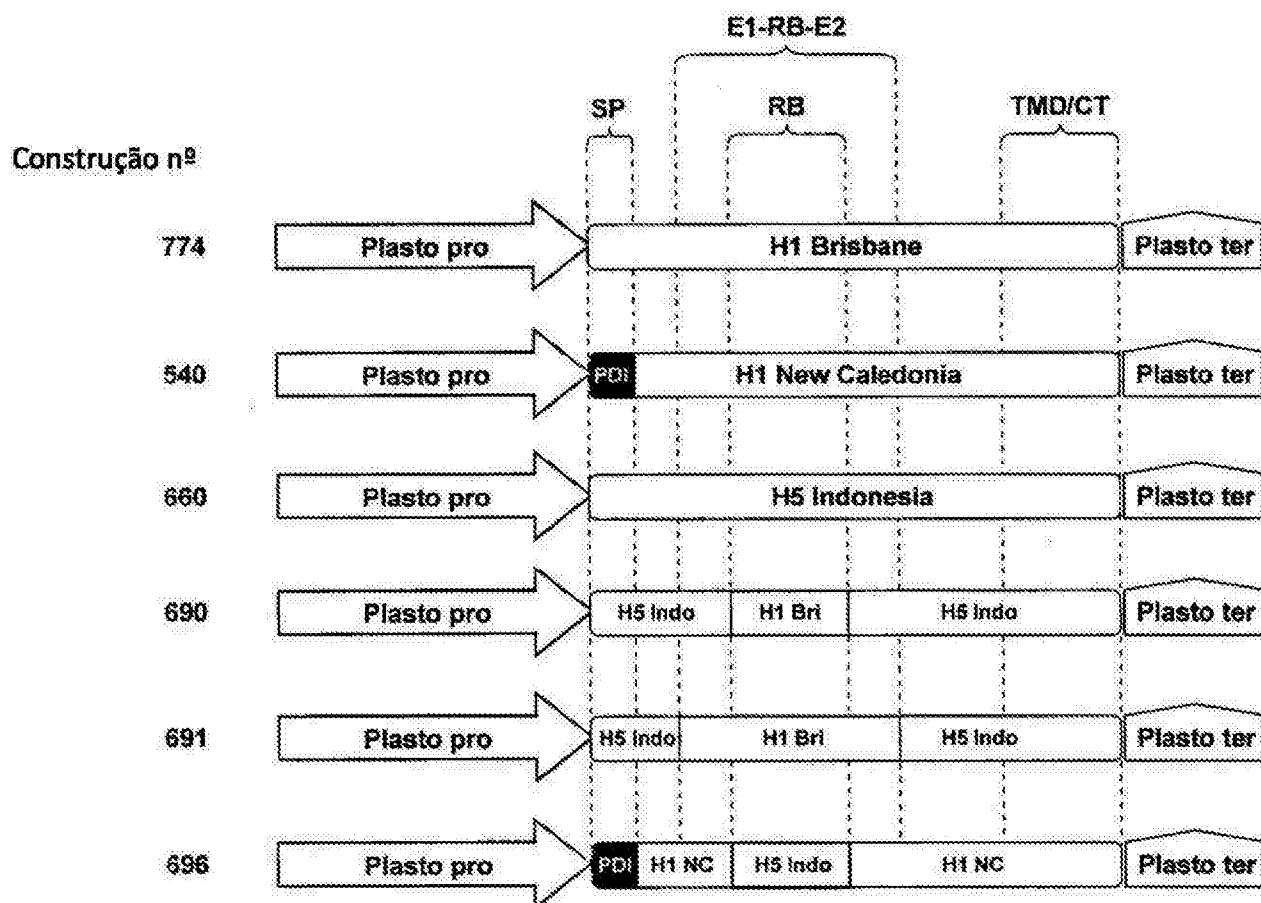


Figura 1b

770 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
840 H1A_NewCal
780 H2A_Singapore
776 H3A_Brisbane
777 H3A_WCN
781 H5A_Anhui
782 H5A_Vietnam
880 H5_Indo

10 20 30
- - - I C T G I T S - S N S P V V T A T Q G V N V T G V I P
- - - I C T G I T S - S N S P V V T A T Q G V N V T G V I P
- - - I C T G I Y A - N N S T T V T V L N V T V T S V N
- - - I C T G I Y A - N N S T T V T V L N V T V T S V N
- - - I C T G I Y A - N N S T T V T V L N V T V T S V N
- - - I C T G I Y A - N N S T T V T V L N V T V T A
T A T L C L G A V P N G T I V T I T N G I V T N A T
T A T L C L G A V P N G T I V T I T N G I V T N A T
- - - I C T G I Y A - N N S T T V T I T N N V T V T A G
- - - I C T G I Y A - N N S T T V T I T N N V T V T A G
- - - I C T G I Y A - N N S T T V T I T N N V T V T A G
T X I C T G I Y A V N N S T X X V T I X N V T V T A X X

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
840 H1A_NewCal
780 H2A_Singapore
776 H3A_Brisbane
777 H3A_WCN
781 H5A_Anhui
782 H5A_Vietnam
880 H5_Indo

40 50 60
L T T T R T S Y F A M L G T T G L C P - C L N - - -
L T T T R T S Y F A M L G T T G L C P - C L N - - -
L L - - - S - - - - - - - - - - - G L C L L - G I A P - L
L L - - - S - - - - - - - - - - - G L C L L - G I A P - L
L L - - - S - - - - - - - - - - - G L C L L - G I A P - L
L L - - - S - - - - - - - - - - - G L C L L - G I A P - L
L V - G - S - S - - - - - - - - - - - G G I C S P - - - G I L
L V - G - S - S - - - - - - - - - - - G G I C S P - - - G I L
L L - - - T - - - - - - - - - - - G L C L - G V P L I
L L - - - T - - - - - - - - - - - G L C L - G V P L I
L L - - - T - - - - - - - - - - - G L C L - G V P L I
I I A X R X - - - F A M L G T X T G I C X I X C g X X P I I

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
840 H1A_NewCal
780 H2A_Singapore
776 H3A_Brisbane
777 H3A_WCN
781 H5A_Anhui
782 H5A_Vietnam
880 H5_Indo

70 80 90
- - - C T E V A L G P M C V G T T H S A A S - L L V
- - - C T E V A L G P M C V G T T H S A A S - L L V
Q L G N C S V A G W L L G N P C L L I S S W S Y I V P
Q L G N C S V A G W L L G N P C L L I S S W S Y I V P
Q L G N C S V A G W L L G N P C L L I S S W S Y I V P
Q L G N C S V A G W L L G N P C L L I S S W S Y I V P
G N C T L I A L L G P C C G G N - W L V V S
G N C T L I A L L G P C C G G N - W L V V S
- L C S V A G W L L G N P M C C I N V P W S Y I V A
- L C S V A G W L L G N P M C C I N V P W S Y I V A
- L C S V A G W L L G N P M C C I N V P W S Y I V A
X I X R C a v a g w l l g n p x c X X I X X X X w s y i v X

Length: 609

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
840 H1A_NewCal
780 H2A_Singapore
776 H3A_Brisbane
777 H3A_WCN
781 H5A_Anhui
782 H5A_Vietnam
880 H5_Indo

100 110 120 130
P - V T S G C F P I M T T L L L L G - - Y T
P - V T S G C F P I M T T L L L L G - - Y T
P - G T C Y P G A Y L L - S S V S S S
P - G T C Y P G A Y L L - S S V S S S
P - G T C Y P G A Y L L - S S V S S S
P - G T C Y P G A Y L L - S S V S S S
P - A Y S N C Y P Y V P Y A S L S L V A S S G - L L
P - A Y S N C Y P Y V P Y A S L S L V A S S G - L L
P - A N L C Y P G N Y Y L L - S I N P I
P - V N L C Y P G N Y Y L L - S I N P I
P - T N L C Y P G S N Y Y L L - S I N P I
T P X X X C Y P g X X X Y I X X I X a S X X X P X I

Figura 1C-1

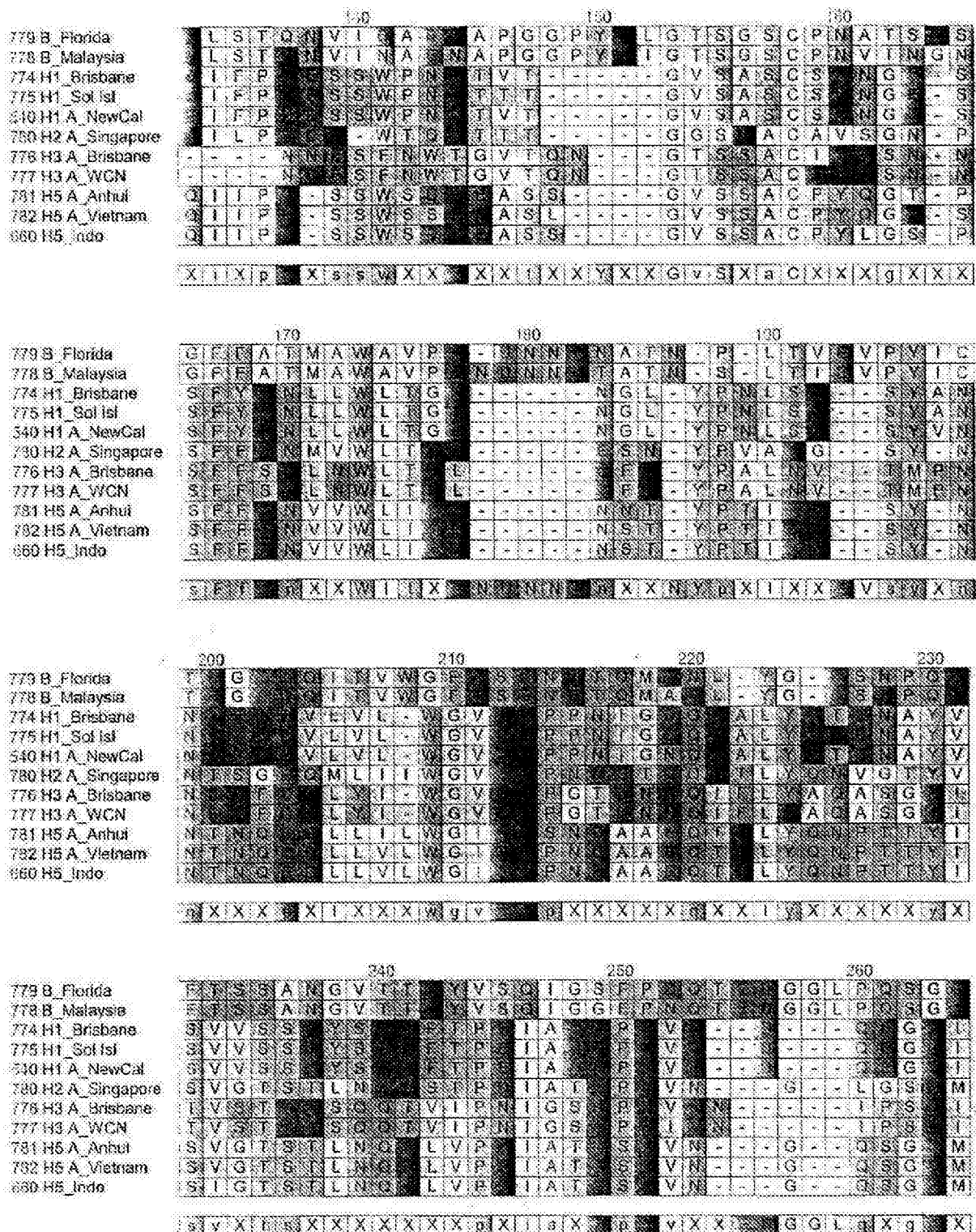
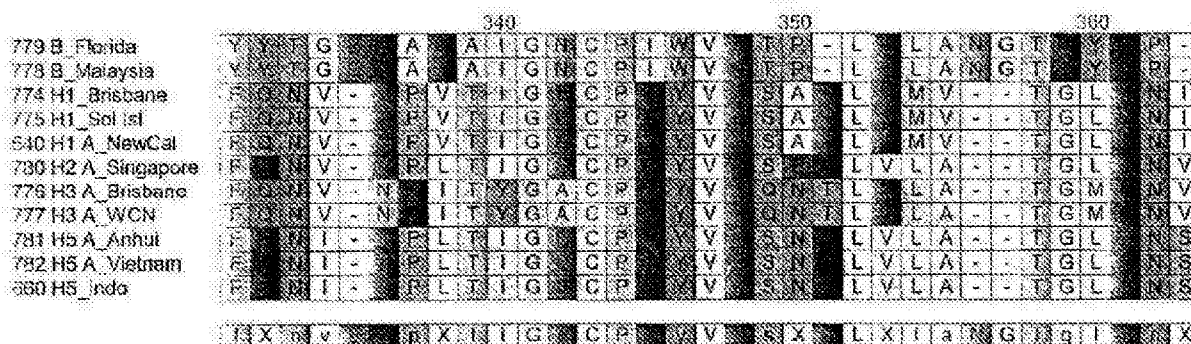


Figura 1C-2



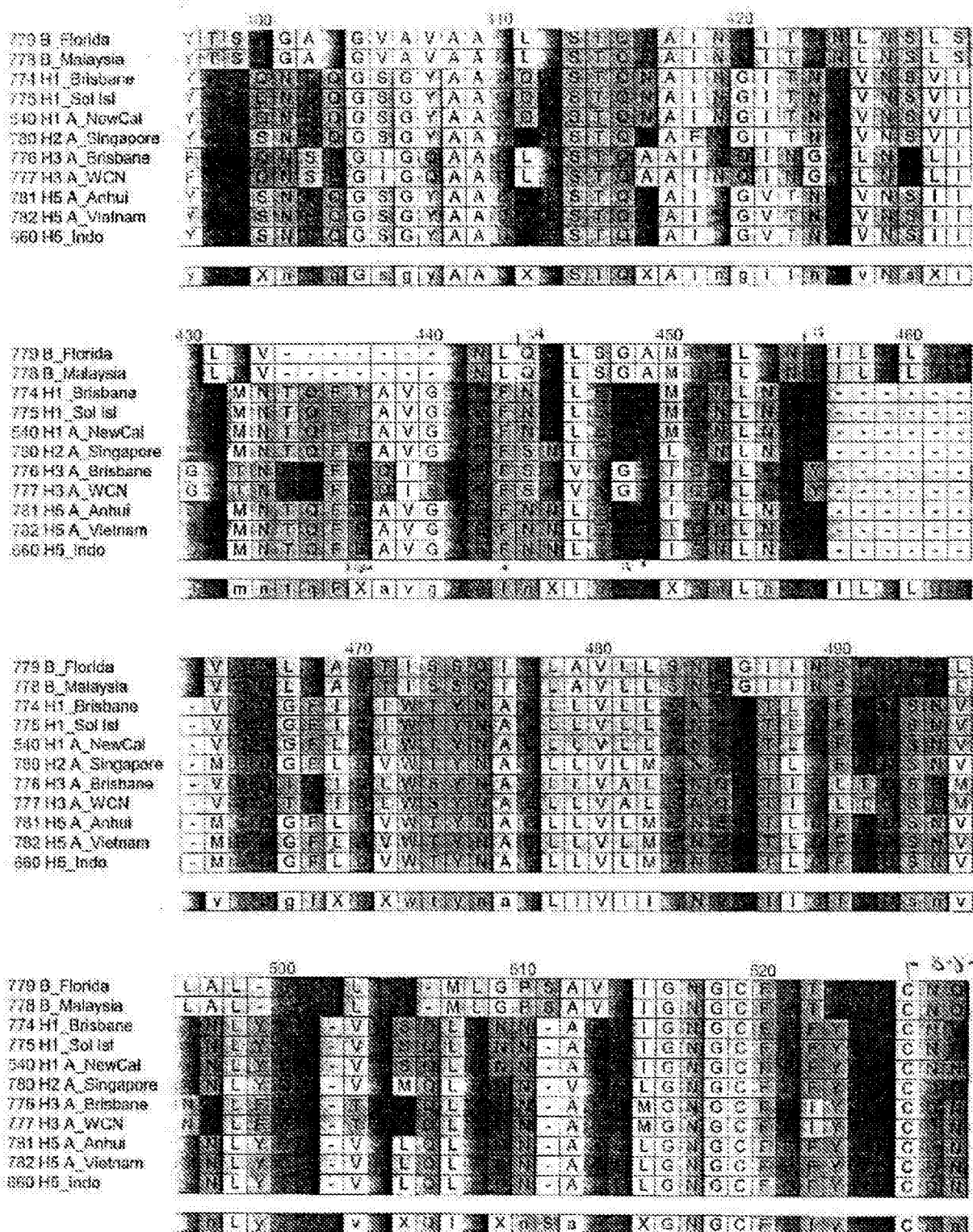


Figura 1C-4

	20	40	60	80	100	120	140	160	180	200	220	240	260	280	300	320	340	360	380	400	420	440	460	480	500	520	540	560	580	600
770 B_Florida	I	C	L	I	A	A	G	T	P	M	A	G	F	S	L	P	T	F	S	-	L	N	I	T	A	A	S	L	N	
778 B_Malaysia	I	C	L	I	A	A	G	T	P	M	A	G	F	S	L	P	T	F	S	-	L	N	I	T	A	A	S	L	N	
774 H1_Brisbane	C	M	S	V	N	G	T	Y	Y	P	Y	-	-	-	S	S	L	N	I	I	T	A	A	S	L	N				
775 H1_Sol Isl	C	M	S	V	N	G	T	Y	Y	P	Y	-	-	-	S	S	L	N	I	I	T	A	A	S	L	N				
540 H1 A_NewCal	C	M	S	V	N	G	T	Y	Y	P	Y	-	-	-	S	S	L	N	I	I	T	A	A	S	L	N				
780 H2 A_Singapore	C	M	N	S	V	N	G	T	Y	Y	P	Y	-	-	-	S	L	N	I	I	T	A	A	S	L	N				
776 H3 A_Brisbane	A	C	I	G	S	I	N	G	T	Y	V	Y	-	-	-	S	A	L	N	I	I	T	A	A	S	L	N			
777 H3 A_WCN	A	C	I	G	S	I	N	G	T	Y	V	Y	-	-	-	S	A	L	N	I	I	T	A	A	S	L	N			
781 H5 A_Anhui	C	M	S	V	N	G	T	Y	Y	P	Q	Y	-	-	-	S	A	L	N	I	I	T	A	A	S	L	N			
782 H5 A_Vietnam	C	M	S	V	N	G	T	Y	Y	P	Q	Y	-	-	-	S	A	L	N	I	I	T	A	A	S	L	N			
560 H5 Indo	C	M	S	V	N	G	T	Y	Y	P	Q	Y	-	-	-	S	A	L	N	I	I	T	A	A	S	L	N			

I C M S V X N G T Y Y P X Y S L P S S X I N X I X G V

	470															500															530														
770 B_Florida	-	-	-	G	L	A	T	I	L	L	Y	Y	S	T	A	A	-	S	S	L	A	V	T	L	M	I	A	I	F	I															
770 B_Malaysia	-	-	-	G	L	A	T	I	L	L	Y	Y	S	T	A	A	-	S	S	L	A	V	T	L	M	I	A	I	F	I															
774 H1_Brisbane	L	S	M	G	V	Y	G	I	L	A	I	-	Y	S	T	V	A	-	S	S	L	V	L	L	V	S	L	G	A	I	S														
775 H1_Sol Isl	L	S	M	G	V	Y	G	I	L	A	I	-	Y	S	T	V	A	-	S	S	L	V	L	L	V	S	L	G	A	I	S														
540 H1 A_NewCal	L	S	M	G	V	Y	G	I	L	A	I	-	Y	S	T	V	A	-	S	S	L	V	L	L	V	S	L	G	A	I	S														
780 H2 A_Singapore	L	S	S	M	G	V	Y	G	I	L	A	I	-	Y	A	T	V	A	G	-	S	L	S	L	A	I	M	M	A	G	I	S													
776 H3 A_Brisbane	L	S	-	G	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-														
777 H3 A_WCN	L	S	-	G	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-														
781 H5 A_Anhui	L	S	I	G	T	Y	G	I	L	S	I	-	Y	S	T	V	A	-	S	S	L	A	L	A	I	M	V	A	G	L	S														
782 H5 A_Vietnam	L	S	I	G	T	Y	G	I	L	S	I	-	Y	S	T	V	A	-	S	S	L	A	L	A	I	M	V	A	G	L	S														
560 H5 Indo	L	S	I	G	T	Y	G	I	L	S	I	-	Y	S	T	V	A	-	S	S	L	A	L	A	I	M	M	A	G	L	S														

L S X G X Y G I I X I Y Y S S V A I S S I X I X X M I A G X S

Length: 609

Queens cys

	600														
779 B_Florida	V	Y	M	V	S		N	V	S	C	S	I	C	L	
778 B_Malaysia	V	Y	M	V	S		N	V	S	C	S	I	C	L	
774 H1_Brisbane	F	W	M	C	S	N	G	S	L	Q	C		I	C	I
775 H1_Sol Isl	F	W	M	C	S	N	G	S	L	Q	C		I	C	I
540 H1 A_NewCal	F	W	M	C	S	N	G	S	L	Q	C		I	C	I
780 H2 A_Singapore	F	W	M	C	S	N	G	S	L	Q	C		I	C	I
776 H3 A_Brisbane	M	W	A	C	Q		G	N	I		C	N	I	C	I
777 H3 A_WCN	M	W	A	C	Q		G	N	I		C	N	I	C	I
781 H5 A_Anhui	L	W	M	C	S	N	G	S	L	Q	C		I	C	I
782 H5 A_Vietnam	L	W	M	C	S	N	G	S	L	Q	C		I	C	I
560 H5_Indo	L	W	M	C	S	N	G	S	L	Q	C		I	C	I

X W M C S N G S I Q C I C I

Figura 1C-5

Troca do subdomínio do recetor de ligação na haste HA. Sequência de proteínas maduras

Construção nº haste N-terminal : F'1+E1		cabeça RB	haste C-terminal: E2+F'2 stop
690 and 734	H5/Indo DQICIGYHANNST TEQVDTIMEKNV TVTHAQDILEKT HNGKLCDLGV KPLILRDCSVAG WLLGNPMCDEF INVPEWSYIVEK ANPTNDLCYP	H1/Brisbane GHFADYEELREQLSSVSSFERFEIF PKESSWPNTHTVTGVSASC SHNGES SFYRNLLWLTGKNGLYPNLSKSYA NNKEKEVLVLWG VHHPPNIGDQKA LYHTENAYVSVVSSHYSRKFTPEIA KRPKVRDQEGRINYWTLLPEGDTI IFEANGNLIAPRYAFALSRGFGSG	H5/Indo IMKSELEYGNCNTKCQTPMGAINSSM PFHNIHPLTIGEC PKYVKS NRLVLATGL RNSPQRESRRKRGLFGAIA GFIEGG WQGMVDGWYGYHHSNEQGS GYAAD KESTQKAIDGVTNKVNSIIDKMNTQFE AVGREFNLERRIENLNKKMEDGFLDV WTYNAELLVLMENERTLDFHDSNVKN LYDKVRLQLRDNAKELGNGCFEFYHK CDNECMESIRNGTYNYPQYSEEARLK REEISGVKLESIGTYQILSIYSTVASSLA LAIMMAGLSLWMCSNGSLQCRICI
696	H1/NC DTICIGYHANNST TDYDVTLEKNV TVTHSVNLLED HNGKLC LLKGIA PLQLGNCSVAG WILGNPECELLI SKESWSYIVETP NPENGTCYP	H5/Indo GSFNDYEELKHLISRINHFEKIQIP KSSWSDHEASSGVSSACPYLGSPS FFRNVAWLKKNSTYPTIKKSYNNT NQEDLLVLWG IHHPNDAAEQTRLY QNPTTYISIGTSTLNQRLVPKIATRS KVNGQSSRMEFFWTILKPNDAINFE SNGNFIAPEYAYKIVKKGDSA	H1/NC IITSNAPMDECDACKQTPGAINSSLP FQNVHPVTIGEC PKYVRS AKLRMV TGL RNIPSIQSRGLFGAIA GFIEGGWTGMV DGWYGYHHQNEQGS GYAADQKSTQ NAINGITNKVNSVIEKMNTQFTAVGKEF NKLERRMENLNKKVDDGFLDWTYNA ELLVLENERTLDFHDSNVKNLYEKVK SQLKNNAKEIGNGCFEFYHKCDNECM ESVKNGTYDYPKYSEESKLNREKIDGV KLESMGVYQILAIYSTVASSLVLLVSLG AISFWMCSNGSLQCRICI

Troca do recetor de ligação e subdomínios de esterase na haste HA. Sequência de proteínas maduras

Construção nº haste N-terminal : F'1		cabeça E1-RB-E2	haste C-terminal: F'2 stop
691	H5/Indo DQICIGYHANNST TEQVDTIMEKNV TVTHAQDILEKT HNGKLC	H1/Brisbane LLKGIA PLQLGNCSVAGWIL GNPECELLISKESWSYIVEK PNPENGTCYPGHFADYEEL REQLSSVSSFERFEIFPKES SWPNHTVTGVSASC SHNG ESSFYRNLLWLTGKNGLYP NLSKSYANNKEKEVLVLWG VHHPPNIGDQKALYHTENA YVSVSSHYSRKFTPEIAKR PKVRDQEGRINYWTLLPE GDTIIFEANGNLIAPRYAFAL SRGFGSGIINSNAPMCK	H5/Indo CNTKCQTPMGAINSSMPFHNIHPLTIGEC PKYVKS NRLVLATGLRNSPQRESRRKRGLFGAIA GFIEGGW QGMVDGWYGYHHSNEQGS GYAADKESTQKAIDG VTNKVNSIIDKMNTQFEAVGREFNLERRIENLNKK MEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNL YDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIR NGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIY STVASSLALAIMMAGLSLWMCSNGSLQCRICI

Figura 2

Translação de aminoácido da sequência de codificação na cassette de expressão da construção 690 e 734 (SEQ ID NO: 80)

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTTHAQDILEKTHNGKLCDL DGVKPLILR
DCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGHFADYEELREQLSSVSSFERFEIFPK
ESSWPNHVTGTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKEKEVLVLWGVHHPPI
GDQKALYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYWTLLPGDTIIFEANGNLIAPRY
AFALSRGFGSGIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKS NRLVLATGLRN
SPQRESRRKKRGLFGAIA GFIEGGWQGMVDGWYGYHHSNEQSGSYAADKESTQKAIDGV TNKVN
SIIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYD
KVRLQLRDN AKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQIL
SIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Figura 3

Translação de aminoácido da sequência de codificação na cassette de expressão da construção 691 (SEQ ID NO: 81)

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTTHAQDILEKTHNGKLCLLKGIA PLQL
GNCSVAGWILGNPECELLISKESWSYIVEKPNPENGTCYPGHFADYEELREQLSSVSSFERFEIF
PKESSWPNHVTGTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKEKEVLVLWGVH
PPNIGDQKALYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYWTLLPGDTIIFEANGN
LIAPRYAFALSRGFGSGIINSNAPMDKCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKS NRLV
LATGLRNSPQRESRRKKRGLFGAIA GFIEGGWQGMVDGWYGYHHSNEQSGSYAADKESTQKA I
DGV TNKVN SIIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDF
HDSNVKNLYDKVRLQLRDN AKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEIS
GVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Figura 4

Translação de aminoácido da sequência de codificação na cassette de expressão da construção 696 (SEQ ID NO: 82)

MAKNVAIFGLLFSLVLVPSQIFADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLEDSHNGKLCLLKGIA
PLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGSFNDYEELKHLLSRINHF EKIQIIPK
SSWSDHEASSGVSSACPYLGSPSFFRN VWLIKKNSTYPTIKKSYNNNTNQEDLLVLWGIHHPNDAAEQ T
RLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKK
GDSAIITSNAPMDECDACQTPQGAINSSLPFQNVHPVTIGECPKYVRS AKLRMVTGLRNIPS IQSRGLF
GAIA GFIEGGWTGMVDGWYGYHHQNEQSGSYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEF
NKLERRMENLNKKVDDGFLDIWTYNAELLVLENER TLDFHDSNVKNLYEKVKS QLKNNAKEIGNGC FE
FYHKCNNECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFW
MCSNGSLQCRICI

Figura 5

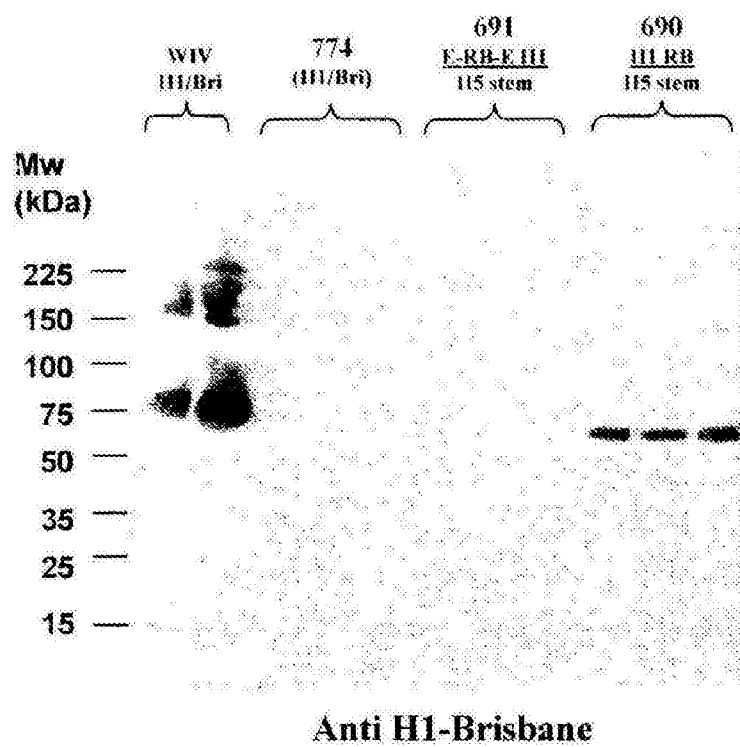


Figura 6

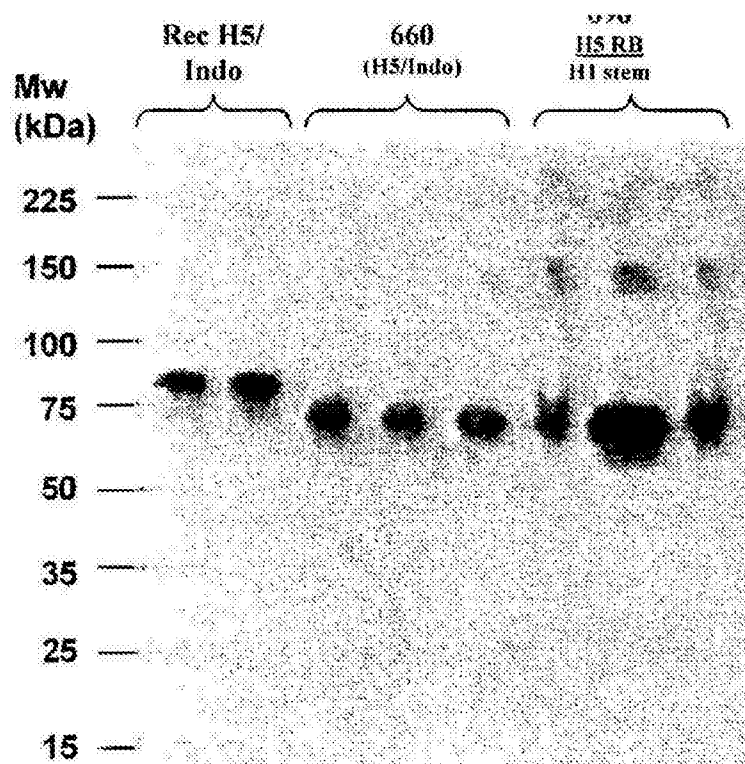


Figura 7

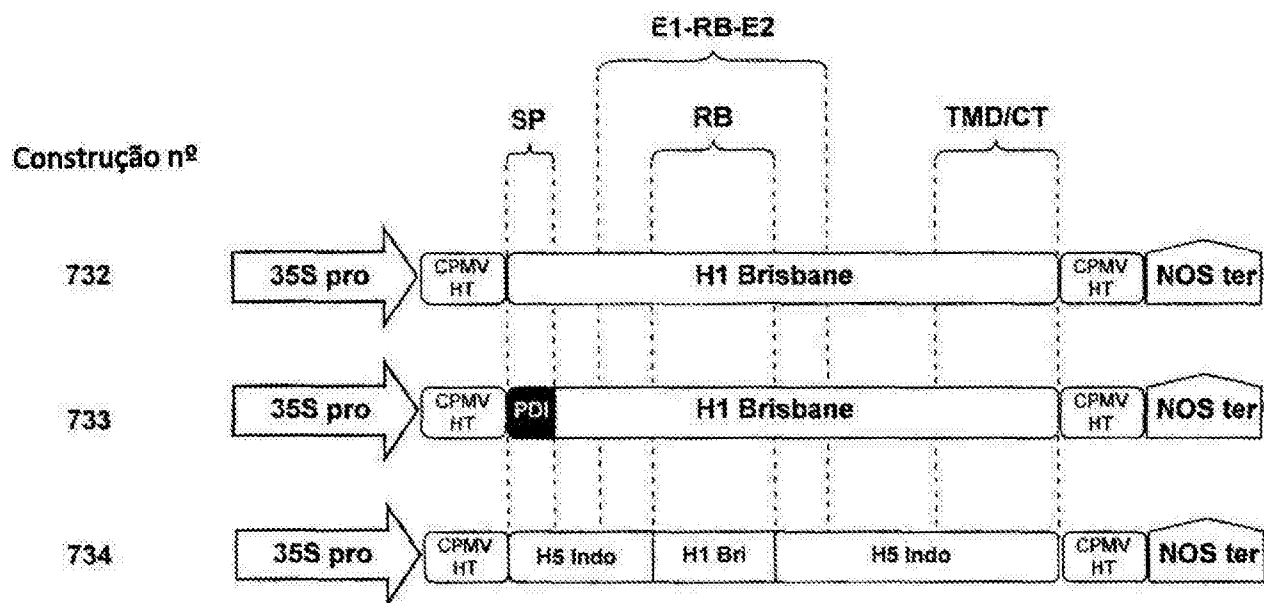
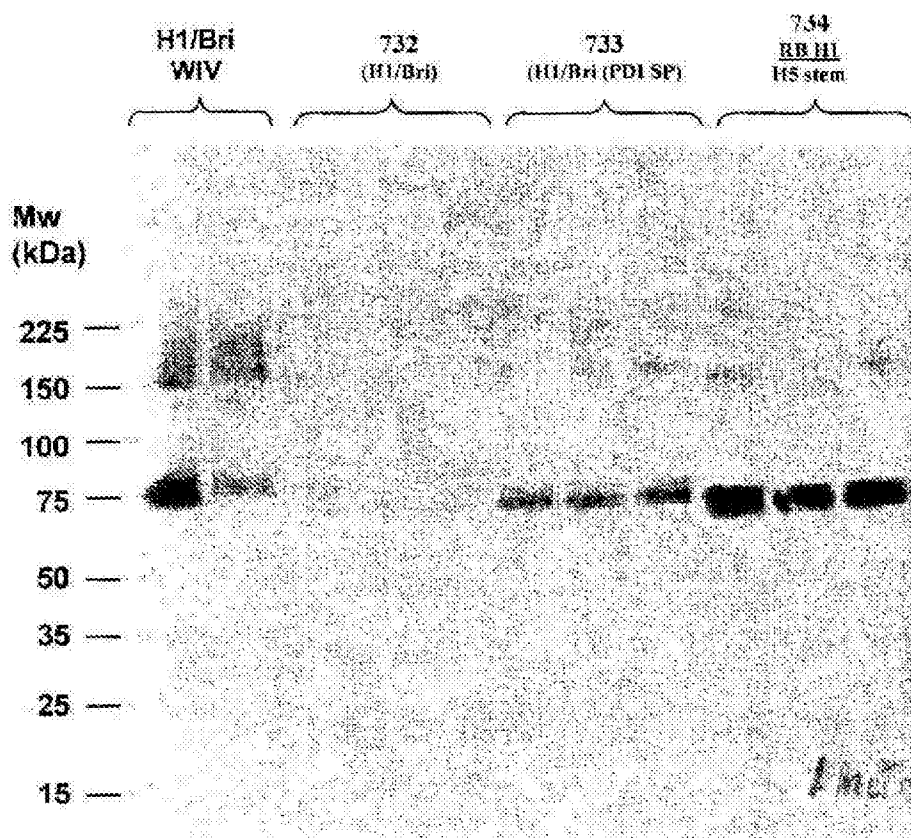


Figura 8

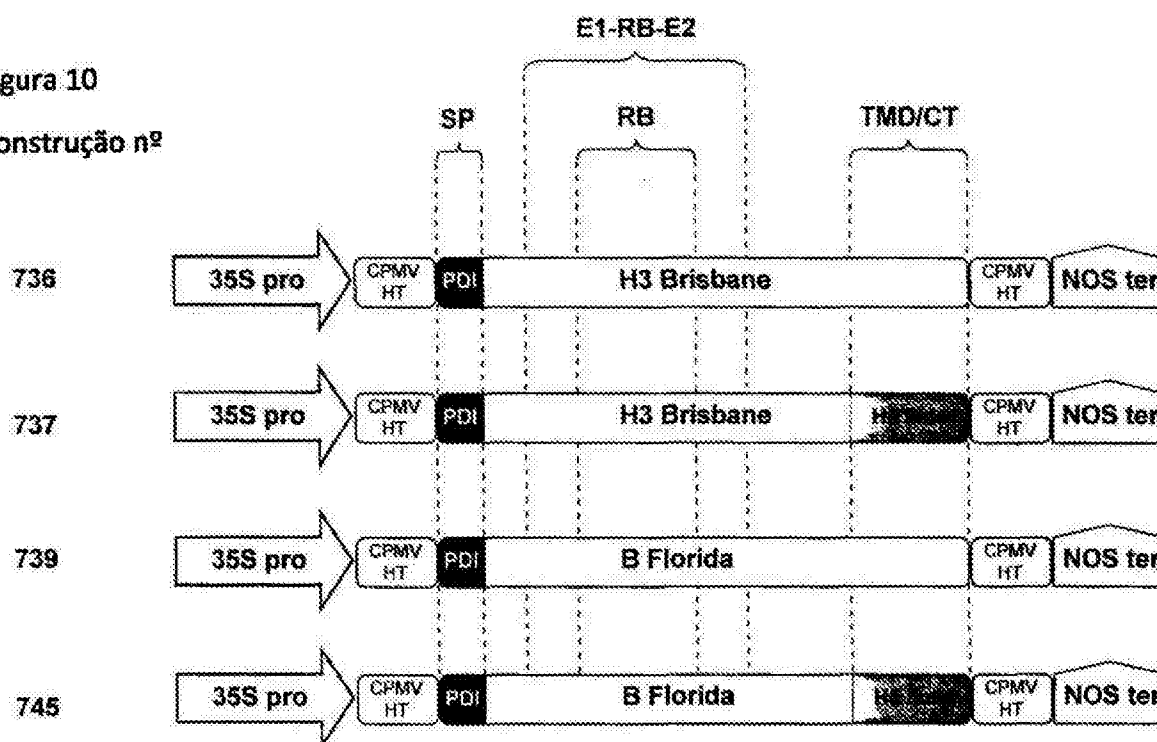


Anti H1-Brisbane

Figura 9

Figura 10

Construção nº



Construção nº

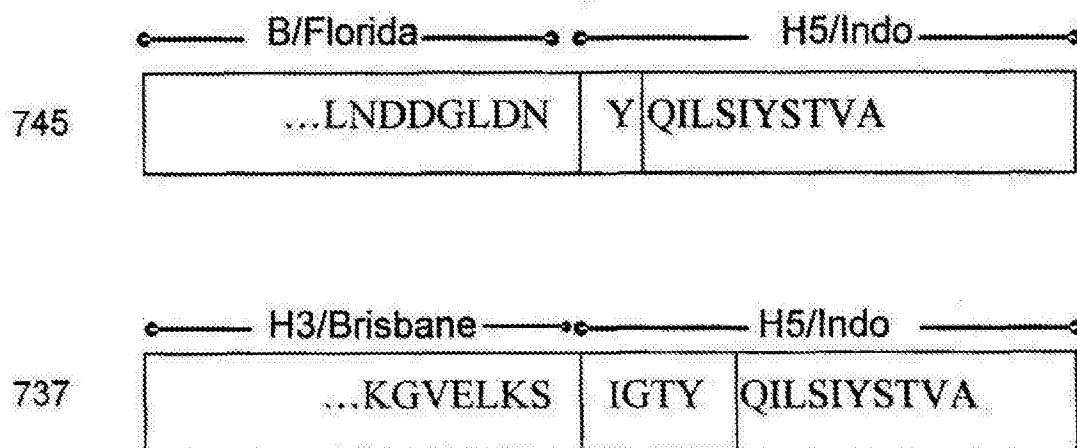


Figura 11

**Translação de aminoácido da sequência de codificação na cassette de expressão da construção 737
(SEQ ID NO: 83)**

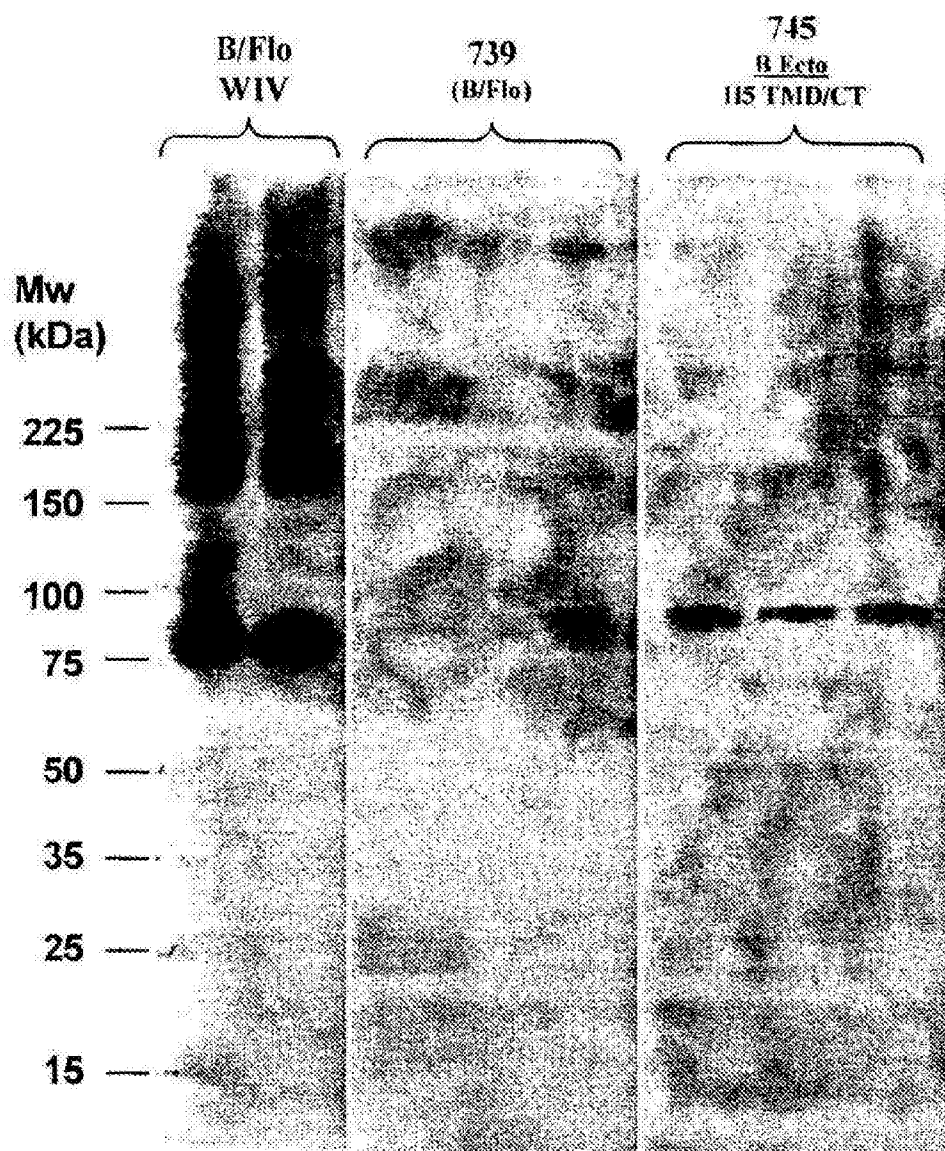
MAKNVAIFGLLFSLLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSS
STGEICDSPHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVA
SSGTLEFNNEFNNWTGVTQNGTSSACIRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFDKLYIW
GVHHPGTDNDQIFLYAQASGRITVSTKRSQQTVIPNIGSRPRVRNIPSRISYWTIVKPGDILLINSTGNL
IAPRGYFKIRSGKSSIMRSDAIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGM
RNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGIGQAADLKSTQAAIDQINGKLNRLIGK
TNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSMNKLFKTKKQLR
ENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSIGTYQILSIYSTVASS
LALAIMMAGLSLWMCSNGSLQCRICI

Figura 12

**Translação de aminoácido da sequência de codificação na cassette de expressão da construção 745
(SEQ ID NO: 84)**

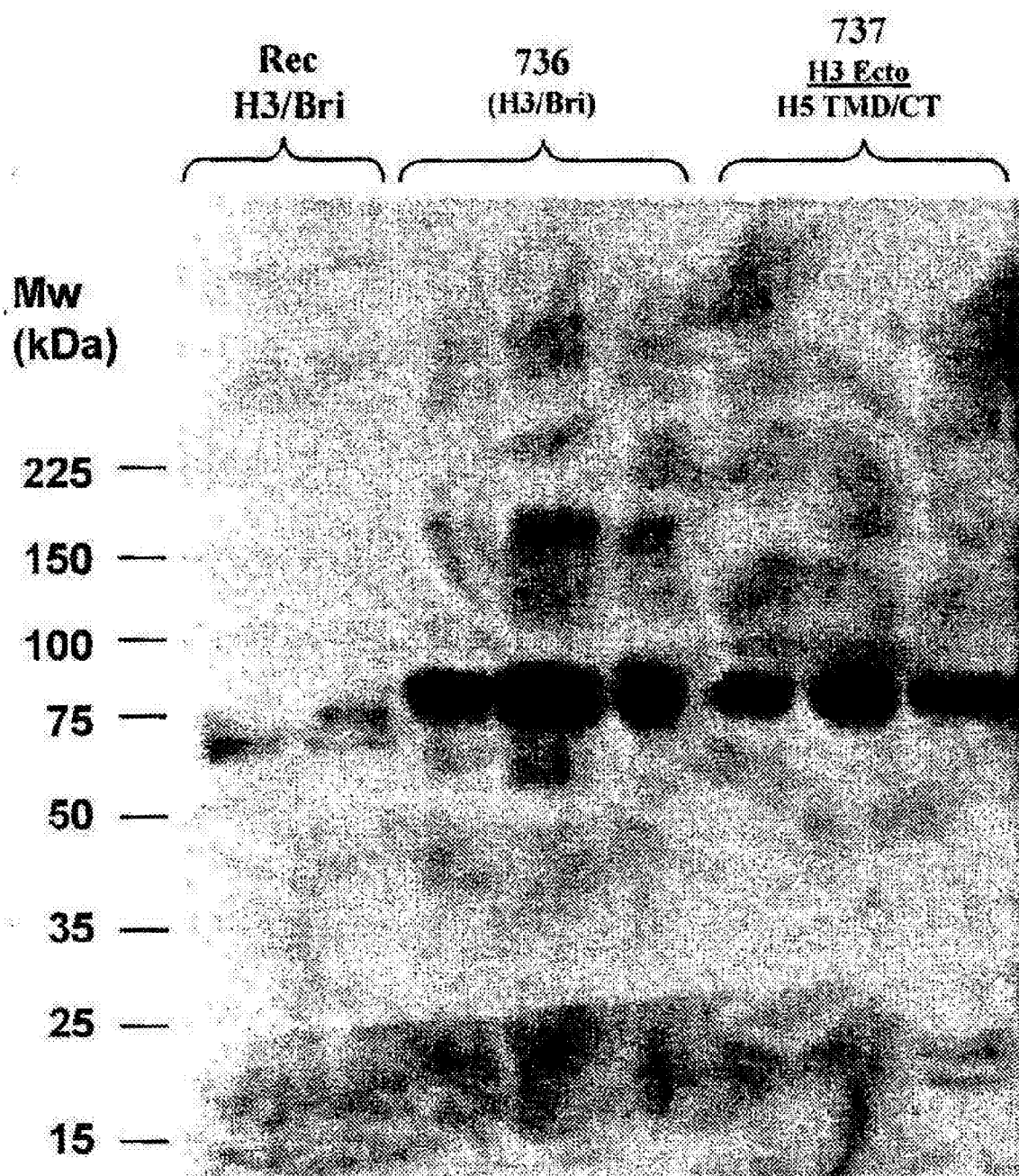
MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKATQGEVNVTVGVIPLTTTPTKSYFANLKGTR
TRGKLCPCDCLNCTDLVALGRPMC VGTTPSAKASILHEVKPVTSGCFPIMHDR TKIRQLPNLLRGYENI
RLSTQNVIDAEKAPGGPYRLGTSGSCP NATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEGED
QITVWGFHSDNKTQMKNLYGDSNPQKFTSSANGVTTHYVSQIGSFDPDQTEDGGLPQSGRIVVDYMMQ
KPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSPLIGEADCLHEKYGGLNKS KPYTGEHAKAIGNCP
IWKTPKLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQE
AINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEGIINSEDEHLL
ALERKLLKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSL NITAASLNDDGLDNY
QILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Figura 13



Anti B-Florida

Figura 14



Anti H3-Brisbane

Figura 15

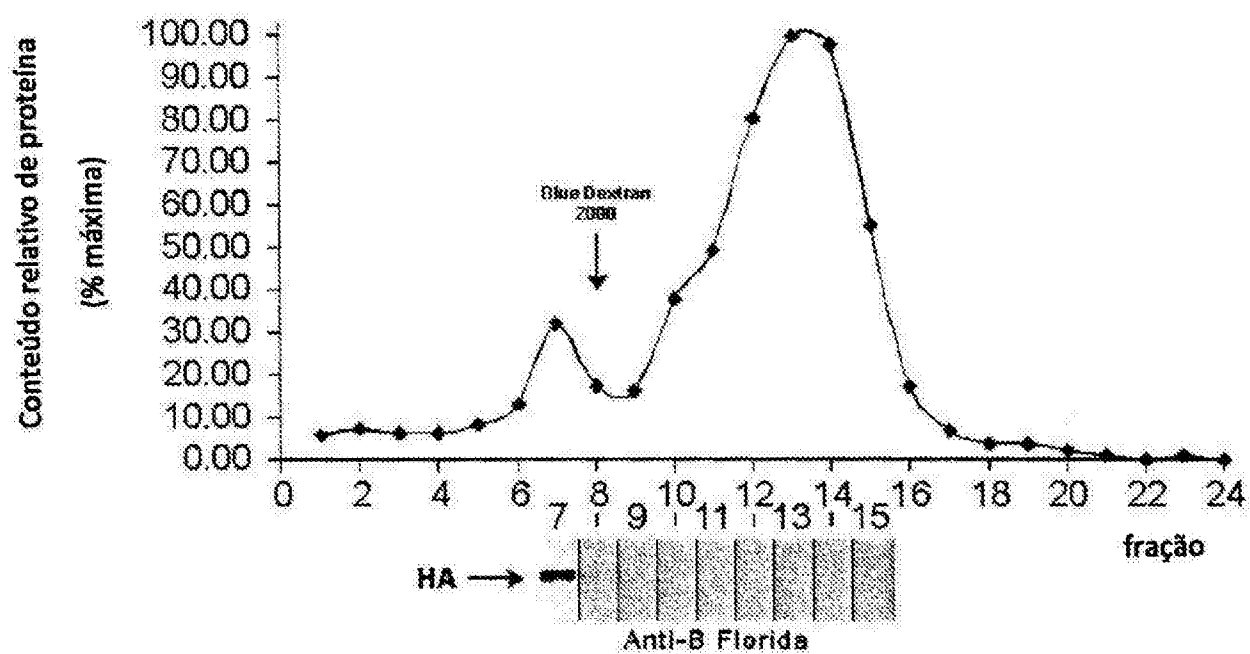


Figura 16

SEQ. ID. NO. 52.

B-Plasto-Native SP-H5 A/Indonesia/5/05 (HindIII-Native Sp-H5 A/Indo/5/05-SacI)

AAGCTTATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATT
TGCATTGGTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGT
TACTGTTACACATGCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAG
ATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCA
ATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGAAGGCCAATCCAAC
CAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAACTGAAACACCTATTGAGCA
GAATAAACCATTTTGAGAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCT
CATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTA
TGGCTTATCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAA
GAGGATCTTTTGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGC
TATATCAAAACCCAACCACTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTAC
CAAAAATAGCTACTAGATCCAAAGTAAACGGGCAAGTGGAAGGATGGAGTTCTTCTGGACA
ATTTTAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAAATTTCAATTGCTCCAGAATAT
GCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAGTGAATTGGAATATGGTAA
CTGCAACACCAAGTGTCAAACCTCAATGGGGGCGATAAACTCTAGTATGCCATTCCACAACA
TACACCCTCTCACCATCGGGGAATGCCCAAATATGTGAAATCAAACAGATTAGTCCTTGCA
ACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAAGAAAAAGAGAGGACTATTTGGAG
CTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCA
CCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAGGCAATA
GATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGCCGT
TGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACG
GGTTTCTAGATGTCTGGACTTATAATGCCGAACCTCTAGGACAAAGGTCCGACTACAGCTTAGGGAT
TAGACTTTCATGACTCAAATGTTAAGAACCCTCTACGACAAGGTCCGACTACAGCTTAGGGAT
AATGCAAAGGAGCTGGGTAAACGGTTGTTTCGAGTTCTATCACAATGTGATAATGAATGTATG
GAAAGTATAAGAAACGGAACGTACAACTATCCGCAGTATTCAGAAGAAGCAAGATTAAAAAG
AGAGGAAATAAGTGGGGTAAATTGGAATCAATAGGAACTTACCAAATACTGTCAATTTATTC
AACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCT
CCAATGGATCGTTACAATGCAGAATTTGCATTTAAGAGCTC

Figura 17

SEQ ID NO: 53

Cassete de expressão 660 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). H5 a partir da sequência de codificação A/Indonésia/5/2005 sublinhada

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATA
TGTTGTCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAAT
ATTACCTACTACTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGT
AGTGATATTTTGACAACAATTTTGTGCAACATTTGAGAAAAATTTGTTGTTCTCTCTTTTCATTGGTCAA
AAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTA
CAAAAGTTGTACCAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTT
AATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTA
AAAAGAAAGAATAAATTATTTTAAAATTAAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAAT
TGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATC
TTTTCTATATATTGCCCCATAGAGTCAGTTAACTCATTTTATATTTTATAGATCAAATAAGAGAAATAA
CGGTATATTAATCCCTCCAAAAAAGGATATTTACTAAAAATCTAAGCCACGTAGGAGGATA
ACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTT
TAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACA
AAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCT
TCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAAATGGAGAAAAAGTG
CTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAA
CAGAGCAGGTTGACACAATCATGGAAAAGAAGCTTACTGTTACACATGCCCAAGACATACTGGAAAAGA
CACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTAAGAGATTGTAGTGTAGCTG
GATGGCTCCTCGGGAACCCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAG
AAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAAGTGAACACCTA
TTGAGCAGAATAAACCATTTTGAGAAAAATCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCT
CATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTAGAAATGGGTATGGCTTA
TCAAAAAGAACAGTACATACCCAACAATAAAGAAAGCTACAATAATACCAACCAAGAGGATCTTTGGT
ACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCCCAACCACCT
ATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGTAA
CGGGCAAAGTGGAAGGATGGAGTTCTCTGGACAATTTAAAACCTAATGATGCAATCAACTTCGAGAG
TAATGGAAATTTCAATGCTCCAGAATATGCATACAAAATTTGTCAAGAAAGGGGACTCAGCAATTATGAAA
AGTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACCTCCAATGGGGGCGATAAACTCTAGTATG
CCATTCCACAACATACACCTCTCACCATCGGGGAATGCCCAAATATGTGAAATCAAACAGATTAGTC
CTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAGAAGAAAAAGAGAGGACTATTTGGAGC
TATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCACCATAGCA
ATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCAATCAAAAGGCAATAGATGGAGTCACCAAT
AAGGTCAAACCAATCATTGACAAAATGAACACTCAGTTTGAGGCGGTTGGAAGGGAATTTAATAACTTA
GAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTTCTAGATGTCTGGACTTATAATGCC
GAACCTCTGGTTCTCATGGAAGATGAGAGAACTCTAGACTTTTATGACTCAAATGTTAAGAACCTCTACG
ACAAGGTCCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACA
AATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACCTATCCGCAGTATTCAGAAGAAG
CAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAATTTGGAATCAATAGGAACCTACCAATACTGTCAAT
TTATTCACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTC
CAATGGATCGTTACAATGCAGAAATTTGCATTTAAGAGCTCTAAGTTAAAATGCTTCTTCGTCTCCTATTTA
TAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTA
TTTGATGAGATGAAGTGGTGAATGTAAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCC
ATAACTAAGTACATGAAGACCTGCCGCGTACAATTTGTCTTATATTTGAACAACCTAAAATTTGAACATCT
TTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAG
TATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACCTACTAATTTTATATCATCCCTTTGAT
AAATGATAGTACCAATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Figura 18

SEQ ID NO: 54

Sequência de codificação H1/NC do tipo selvagem (H1N1) com falta de TmD e Ctail (nº de acesso GenBank AY289929)

ATGAAAGCAAACTACTGGTCCTGTTATGTACATTTACAGCTACATATGCAGACACAATATGTAT
AGGCTACCATGCCAACAACTCAACCGACACTGTTGACACAGTACTTGAGAAGAATGTGACAGT
GACACACTCTGTCAACCTACTTGAGGACAGTCACAATGGAAAACCTATGTCTACTAAAAGGAATA
GCCCCACTACAATTGGGTAAATTGCAGCGTTGCCGGATGGATCTTAGGAAACCCAGAATGCGAA
TACTGATTTCCAAGGAATCATGGTCCTACATTGTAGAAACACCAAATCCTGAGAATGGAACAT
GTTACCCAGGGTATTTCCGCCGACTATGAGGAACTGAGGGAGCAATTGAGTTCAGTATCTTCAT
TTGAGAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAACCCAGCGTAACCGGAGTAT
CAGCATCATGCTCCATAATGGGAAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGA
AGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGTAAACAACAAGAGAAAGAAGTCCTTGT
ACTATGGGGTGTTTCATCACCCGCCCTAACATAGGGAACCAAAGGGGCCCTCTATCATACAGAAAA
TGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAGATTCACCCAGAAATAGCCAAAAGA
CCCAAAGTAAGAGATCAGGAAGGAAGAATCAACTACTACTGGACTCTGCTGGAACCTGGGGAT
ACAATAATATTTGAGGCAAATGGAAATCTAATAGCGCCATGGTATGCTTTTGCAGTGAAGTAGAG
GCTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATGCGAAGTGTCAA
CACCTCAGGGAGCTATAAACAGCAGTCTTCCCTTCCAGAATGTACACCCAGTCACAATAGGAG
AGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCAT
CCATTCAATCCAGAGGTTTGTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAA
TGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATC
AAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTGAATTCTGTAATTGAGAAAAT
GAACACTCAATTCACAGCTGTGGGCAAAGAATTCAACAAATTGGAAAGAAGGATGGAAAACCTTA
AATAAAAAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAATTGTTGGTTCTACT
GGAAAATGAAAGGACTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAA
AGCCAATTAAGAATAATGCCAAAGAAATAGGAAACGGGTGTTTTGAATTCATCACAAGTGTA
ACAATGAATGCATGGAGAGTGTGAAAAATGGAACCTTATGACTATCCAAAATATTCCGAAGAATC
AAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTCTATCAGATTCTG
GCGATCTACTCAACTGTCGCCAGTTCCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTC
TGGATGTGTTCCAATGGGTCTTTGCAGTGTAGAATATGCATCTGAGACCAGAATTTCA

Figura 19

SEQ ID NO: 55

Bg111-H1 A/NC/20/99-SacI/StuI sintetizado

AGATCTTCGCTGACACAATATGTATAGGCTACCATGCCAACAACTCAACCGACACTGTTGACAC
AGTACTTGAGAAGAATGTGACAGTGACACACTCTGTCAACCTACTTGAGGACAGTCACAATGGA
AAACTATGTCTACTAAAAGGAATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGATGGA
TCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGAATCATGGTCCTACATTGTAGAAACA
CCAAATCCTGAGAATGGAACATGTTACCCAGGGTATTTCCGCCGACTATGAGGAACTGAGGGAGC
AATTGAGTTCAGTATCTTCATTTGAGAGATTCGAAAATATCCCCAAAGAAAGCTCATGGCCCAAC
CACACCGTAACCGGAGTATCAGCATCATGCTCCCATAAATGGGAAAAGCAGTTTTTACAGAAATTT
GCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGTAAACAACAAA
GAGAAAGAAGTCCTTGTACTATGGGGTGTTCATCACCCGCCTAACATAGGGAACCAAAGGGCAC
TCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAGATTCACCCCA
GAAATAGCCAAAAGACCCAAAGTAAGAGATCAGGAAGGAAGAATCAACTACTACTGGACTCTGC
TGGAACCTGGGGATACAATAATTTTGGGCAAATGGAAATCTAATAGCGCCATGGTATGCTTTT
GCACTGAGTAGAGGCTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATG
CGAAGTGTCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCCTTTCCAGAATGTACACCCAGT
CACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGG
AACATCCCATCCATTCAATCCAGAGGTTTGTGGAGCCATTGCCGGTTTCATTGAAGGGGGGT
GGACTGGAATGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGC
TGCAGATCAAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTCAATTCTGTAATTG
AGAAAATGAACACTCAATTCACAGCTGTGGGCAAAGAGTTCAACAAATTGGAAAGAAGGATGGA
AACTTAAATAAAAAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAATTGTTGGT
TCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAG
TAAAAAGCCAATTAAAGAATAATGCCAAAGAAATAGGAAACGGGTGTTTGGAGTTCTATCACAAG
TGTAACAATGAATGCATGGAGAGTGTGAAAAATGGTACCTATGACTATCCAAAATATTCCGAAGA
ATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTATACTAAGAG
CTCAGGCCT

Figura 20

SEQ ID NO: 56

KpnI-H1 A/NC/20/99 TmD+Ctail-SacI/StuI sintetizado

GGTACCTATGACTATCCAAAATATTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTG
AAATTGGAATCAATGGGAGTATACCAGATTCTGGCGATCTACTCAACTGTCGCCAGTTCCCTGGT
TCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGGGTCTTGCAGTGTAGAA
TATGCATCTAAGAGCTCAGGCCT

Figura 21

SEQ ID NO: 57

**Péptido de sinal do gene da proteína de alfa dissulfido isomerase
(PDI) (nucleótidos 32-103: nº de acesso Z11499)**

CCAAATCCTTAACATTCTTTCAACACCAACAATGGCGAAAAACGTTGCGATTTTCGGTTTATTGTTT
TCTCTTCTTCTGTTGGTTCCCTTCTCAGATCTTCGCTGAGGAATCATCAACTGACGCTAAGGAATTT
GTTCTTACATTGGATAACACTAATTTCCATGACACTGTTAAGAAGCACGATTTTCATCGTCGTTGAAT
TCTACGCACCTTGGTGTGGACACTGTAAGAAGCTAGCCCCAGAGTATGAGAAGGCTGCTTCTATC
TTGAGCACTCACGAGCCACCAGTTGTTTTGGCTAAAGTTGATGCCAATGAGGAGCACAACAAAGA
CCTCGCATCGGAAAAATGATGTTAAGGGATTCCCAACCATTAAAGATTTTATAGGAATGGTGGAAAGAA
CATTCAAGAATACAAAGGTCCCCGTGAAGCTGAAGGTATTGTTGAGTATTTGAAAAACAAAGTGG
CCCTGCATCCACAGAAATTAAATCTGCTGATGATGCGACCGCTTTTGTGGTGACAACAAAGTTGT
TATTGTGCGAGTTTTCCCTAAATTTTCTGGTGAGGAGTACGATAACTTCATTGCATTAGCAGAGAA
GTTGCGTTCTGACTATGACTTTGCTCACACTTTGAATGCCAAACACCTTCCAAAGGGAGACTCATC
AGTGTCTGGGCCTGTGGTTAGGTTATTTAAGCCATTTGACGAGCTCTTTGTTGACTCAAAGGATTT
CAATGTAGAAGCTCTAGAGAAATTCATTGAAGAATCCAGTACCCCAATTGTGACTGTCTTCAACAA
TGAGCCTAGCAATCACCCCTTTTGTGTCAAATCTTTAACTCTCCCAACGCAAAGGCTATGTTGTT
CATCAACTTTACTACCGAAGGTGCTGAATCTTTCAAACAAAATACCATGAAGTGGCTGAGCAATA
CAAACAACAGGGAGTTAGCTTTCTTGTGGAGATGTTGAGTCTAGTCAAGGTGCCTTCCAGTATTT
TGGACTGAAGGAAGAACAAGTACCTCTAATTATTATTCAGCATAATGATGGCAAGAAGTTTTTCAA
ACCCAATTTGGAACCTTGATCAACTCCCAACTTGGTTGAAGGCATACAAGGATGGCAAGGTTGAAC
CATTTGTCAAGTCTGAACCTATTCTGAAACTAACAACGAGCCTGTTAAAGTGGTGGTTGGGCAA
CTCTTGAGGACGTTGTTTTCAAGTCTGGGAAGAATGTTTTGATAGAGTTTTATGCTCCTTGGTGTG
GTCAGTGAAGCAGTTGGCTCCAATCTTGGATGAAGTTGCTGTCTCATTCCAAAGCGATGCTGAT
GTTGTTATTGCAAACTGGATGCAACTGCCAACGATATCCCAACCGACACCTTTGATGTCCAAGG
CTATCCAACCTTGACTTCAGGTCAGCAAGTGGAAAACTATCACAATACGACGGTGGTAGGACAA
AGGAAGACATCATAGAATTCATTGAAAAGAACAAGGATAAACTGGTGCTGCTCATCAAGAAGTAG
AACAACCAAAAGCTGCTGCTCAGCCAGAAGCAGAACAACCAAAAGATGAGCTTTGAAAAGTTCCG
CTTGAGGATATCGGCACACAGTCATCTGCGGGCTTTACAACCTTTTTGTATCTCAGAATCAGAA
GTTAGGAAATCTTAGTGCCAATCTATCTATTTTTGCGTTTCATTTATCTTTTTGGTTTACTCTAATG
TATTACTGAATAATGTGAGTTTTGGCGGAGTTTAGTACTGGAACTTTTGTTTCTGTAAAAA
AA

Figura 22

SEQ. ID. NO. 58.

PromPlasto-PDI SP-BgIII

CTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATATAAAGTTTAAGTTAGCAAGTGTGTACATT
TTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATA
TGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTGCAACATTTGAGAAAAATTTG
TTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGGAGAATAAAAAACATA
ATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAATAGTTGTACAAATATCATTGAGGA
ATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGA
TGTACCATTAGAGAATTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTAAAATTAAGTT
GAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGT
AATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCATAGAGTCAG
TTAACTCATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAA
AAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
ACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGG
CACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATC
TTTATCACCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTCAAACACATACA
AAGAGAAGAGACTAATTAATTAATCATCTTGAGAGAAAAATGGCGAAAAACGTTGCGATTTTCG
GCTTATTGTTTTCTCTCTTGTGTTGGTTCCTTCTCAGATCT

Figura 23A

SEQ ID NO: 85

Sequência de plasmídeo SacI-Plasto 3'UTR

GAGCTCTAAGTTAAATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTC
TGTAAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAAATG
TAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAAGTACATGAAGACC
TGCCGCGTACAATTGTCTTATATTTGAACAACATAAATTGAACATCTTTTGCCACAACCTTTATAAGT
GGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAATT
CATTACAATCAACTTAACGTTATTAATACTACTAATTTTATATCATCCCCTTTGATAAATGATAGTACA

Figura 23B

SEQ ID NO: 59

Cassete de expressão 540 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). H5 a partir da sequência de codificação A/Nova Caledónia/20/1999 sublinhada

AAGCTTGCTAGCGGCCCTCAATGGCCCTGCAGGTGCACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTTGT
CAAATAACTCAAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTA
CTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAGAGTAGTGATATTTTGACAAC
AATTTTGTGCAACATTTGAGAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGG
AAGAGGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAA
ATATCATTGAGGAATTTGACAAAAGCTACACAAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGAGGCATTAGA
GAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAGAAAGAAAGAAATAATTATTTTAAATTTAAAGTTGAGTC
ATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGT
GTCAAATTTAATTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTTTTATATTTTCATAG
ATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTATATTTACTAAAAAATCTAAGCCA
CGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAA
CCCCTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACA
CAAAAACCAATCCACATCTTTATCACCCTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAA
ACACATACAAAGAGAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGCGAAAAACGTTGCGATTTTCG
GCTTATTGTTTTCTCTCTTGTTGGTTTCCTTCTCAGATCTTCCGTGACACAATATGTATAGGCTACCATGCCAA
CAACTCAACCGACACTGTTGACACAGTACTTGAGAAGAAATGTGACAGTGACACACTCTGTCAACCTACTTGAGG
ACAGTCACAATGGAAAACTATGTCTACTAAAAGGAATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGAT
GGATCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGAATCATGGTCTACATTGTAGAAACACCAAATC
CTGAGAATGGAAACATGTTACCCAGGGTATTTCCGCCGACTATGAGGAACCTGAGGGAGCAATTGAGTTTCAGTATCT
TCATTTGAGAGATTGGAATATTCCCCAAGAAAGCTCATGGCCCAACCCACACCGTAACCGGAGTATCAGCATCA
TGCTCCCATAAATGGGAAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAAC
CTGAGCAAGTCTCTATGTAACAACAAAGAGAAAGAAAGTCTTGTACTATGGGGTGTTCATCACCCGCCTAACATA
GGGAACCAAGGGGCACTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAAGATTC
ACCCAGAAATAGCCAAAGACCCAAAGTAAGAGATCAGGAAGGAAGAATCACTACTACTGGACTCTGCTGGA
ACCTGGGGATACAATAATTTGAGGCAAATGGAATCTAATAGCGCCATGGTATGCTTTTGCCTGAGTAGAGG
CTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATGCGAAGTGTCAAACACCTCAGGGAG
CTATAAACAGCAGTCTTCTTCCAGAATGTACACCCAGTCACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTG
CAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCAATCCAGAGGTTTGTGTTGGAGCCATTGCCG
GTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCT
GGCTATGCTGCAGATCAAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTCAATTCTGTAATTGAG
AAAATGAACACTCAATTACAGCTGTGGGCAAGAGTTCAACAAATTGGAAAGAAGGATGGAAAACCTAAATAAA
AAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAAATTGTTGGTTCTACTGGAAAATGAAAGGACTT
TGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAGAATAATGCCAAAGAAAT
AGGAAACGGGTGTTTTGAGTTCTATCACAAGTGAACAATGAATGCATGGAGAGTGTGAAAAATGGTACCTATGA
CTATCCAAAATATTCGAAGAATCAAAGTTAAACAGGGAGAAAAATTGATGGAGTGAAATTGGAATCAATGGGAGT
ATACGAGATTCTGGCGATCTACTCAACTGTCCGAGTCCCTGGTTCTTTGGTCTCCCTGGGGGCAATCAGCTT
CTGGATGTGTTCCAATGGGTCTTTGCAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAAAATGCTTCTTCGTCT
CCTATTTATAATATGGTTTGTATTGTTAATTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATAC
TATTTGTATGAGATGAACTGGTGAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCTCCATAA
CTAACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGAACAACATAAAATTGAACATCTTTGCCACAA
CTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAAT
TCATTAACAATCAACTTAACGTTATTAATACTACTAATTTTATATCATCCCTTTGATAAATGATAGTACACCAATTAG
GAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Figura 24

SEQ ID NO: 60

Fragment sintetizado Dralll-Plasto(-84+1)-H1 A/Brisbane/59/07-Sac1

CAC TTT GTG AGT CTAC ACT TTT GAT TCC CTT CAA ACAC ATAC AAA GAGA AGAG ACTA ATTA ATTA AT
TAAT CAT CTT GAG AGA AAA ATG AAA GTAAA ACTACT GGT CCT GTT ATG CAC ATTTAC AGCTAC ATAT
GCAG ACACA ATATGT ATAGG CTACCAT GCTAACAACTCG ACCGACACTGTTGACACAGTACTTGA
AAAGA ATGTGACAGTGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAA ACTATGTC
TATTA AAAAGGAATAGCCCCACTACAATTGGGTAA TTGCAGCGTTGCCGGGTGGATCTTAGGAAAC
CCAGA ATGCGAATTACTGATTTCCAAGGAGTCATGGTCCTACATTGTAGAAAAACCAAATCCTGA
GAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAACTGAGGGGAGCAATTGAGTTCA
GTATCTTCATTTGAGAGGTTT CGAAATATTCCCCAAAGAAAGCTCATGGCCCAACCACACCCGTAAC
CGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTTACAGAAATTTGCTATGGCTGA
CGGGGAAGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAAACAACAAAGAAAAAGAAGT
CCTTGTACTATGGGGTGTTCATCACCCGCCAAACATAGGTGACCAAAGGCCCTCTATCATACA
GAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAAATTCACCCAGAAATAGCCAAA
AGACCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGACTCTGCTTGAACCCGGGG
ATACAATAATATTTGAGGCCAAATGGAATCTAATAGCGCCAAGATATGCTTTCGCACTGAGTAGA
GGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAAGTGCCAAAC
ACCTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAACGTACACCCAGTCACAATAGGAGAG
TGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCAT
TCAATCCAGAGGTTTGTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTA
GATGGTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAAG
CACACAAAATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCA
ATTCACAGCAGTGGGCAAAGAGTTCAACAAATTGGAAGAAGGATGGAAA ACTTGAATAAAAAAG
TTGATGATGGGTTTATAGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAA
GGACTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAAGTAAAAAGCCAGTTAAAG
AATAATGCTAAAGAAATAGGAAATGGGTGTTTTGAGTTCTATCACAAGTGTAACGATGAATGCAT
GGAGAGTGTAAGAATGGA ACTTATGACTATCCAAAATATTCCGAAGAATCAAAGTTAAACAGGG
AGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTCTATCAGATTCTGGCGATCTACTCAACA
GTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATG
GGTCTTTACAGTGTAGAATATGCATCTAAGAGCTC

Figura 25

SEQ ID NO: 61

Cassete de expressão 774 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). H1 a partir da sequência de codificação A/Brisbane/59/2007 sublinhada

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTGCTAGAGGTACCCCGGGCTGGTATATTTATATGTT
GTCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAATATTCACC
TACTACTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTT
GACAACAATTTTGTGCAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGA
GAAAAAGGAAGAGGGAGAAATAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAA
TAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGG
ATGACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTAA
AATTAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTAT
TAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTA
ACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTAT
ATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCA
ATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCAC
ACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCATTCTATAAAAAATCACACT
TTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGA
GAGAAAATGAAAGTAAACTACTGGTCCTGTTATGCACATTTACAGCTACATATGCAGACACAATATGTATAG
GCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGAAAAGAATGTGACAGTGACACACTCTG
TCAACCTGCTTGAGAACAGTCACAATGGAAAACCTATGTCTATTAAGGAATAGCCCCACTACAATTGGGTAA
TTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCCAAGGAGTCATGGTCCTA
CATTGTAGAAAAACCAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAACTGAG
GGAGCAATTGAGTTCAGTATCTTCATTGAGAGGTTGGAATATTTCCCAAGAAAGCTCATGGCCCAACCA
CACCGTAACCGGAGTGTGAGCATCATGCTCCCATATGGGGAAGCAGTTTTTACAGAAATTTGCTATGGCT
GACGGGGAAGAATGGTTGTACCCAAACCTGAGCAAGTCCTATGCAACAAACAAAGAAAAAGAGTCCTTGT
ACTATGGGGTGTTCATCACC CGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGT
CTCTGTAGTGTCTTACATTATAGCAGAAAAATTCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAA
GAAGGAAGAATCAATTACTACTGGACTCTGCTTGAACCCGGGATACAATAATATTTGAGGCAATGGAAAT
CTAATAGCGCCAAGATATGCTTTCCGACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCA
ATGGATAAATGTGATGCGAAGTGCCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAACGTA
CACCCAGTCACAATAGGAGAGTGTCAAAGTATGTGAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGG
AACATCCCATCCATTCAATCCAGAGGTTTGTGTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGA
ATGGTAGATGGTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAAGC
ACACAAAATGCCATTAAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCAATTCACAG
CAGTGGGCAAGAGTTCAACAAATGGAAAGAAGGATGGAAAACCTGAATAAAAAAGTTGATGATGGGTTTAT
AGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCC
AATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGGGTGT
TTGAGTTCTATCACAAGTGTAAACGATGAATGCATGGAGAGTGTAAAGAATGGAATCAATGGGAGTCTATCAGATT
TTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTTGATGGAGTGAAATGGAATCAATGGGAGTCTATCAGATT
CTGGCGATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATG
TGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAAAATGCTTCTTCGTCTCCTATT
TATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATT
TGTATGAGATGAACTGGTGTAAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCTCCATAACT
AACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGAACAACATAAAATGAACATCTTTTGCCACAA
CTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAATATCAGTTATCGA
AATTCATTAAACAATCAACTTAACGTTATTAATACTAATTTTATATCATCCCTTTGATAAATGATAGTACACCA
ATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Figura 26

SEQ ID NO: 62

Cassete de expressão 828 a partir de PacI (a montante do promotor) para Ascl (imediatamente a jusante do terminador NOS). sequência CPMV HT 3'UTRH1 sublinhada com ATG mutante a negrito. Sítio de restrição ApaI (imediatamente a jusante de ATG para a sequência de codificação de proteína, neste caso a cadeia leve C5-1 kapa, está sublinhada em *itálico*)

TTAATTAAGAATTGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTT
TATTGAGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGC
CATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCG
TGGAAAAAGAAGACGTTCCAACACAGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGT
AAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTAT
TTGGAGAGGTATTAAATCTTAATAGGTTTTGATAAAGCGAACGTGGGGAAACCCGAACCAAACCT
TCTTCTAAACTCTCTCTCATCTCTCTTAAGCAAACCTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTT
CAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATC
TCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTC
TTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCATACATTACTTGTACGATTCTGCTGA
CTTTGCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTT
CTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCG
AACTTGAGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGTTTTACA
CCTCAGATACTTGGACTTATGCTTTTTTGGATTTGAGCCTCCAGAGGTGATATTGTGCTAACTCAGT
CTCCAGCCACCCTGTCTGTGACTCCAGGAGATAGTGTGAGTCTTTCTGCAGGGCCAGCCAAAGTA
TTAGCAACAACCTACACTGGTTTCAACAAAAATCGCATGAGTCTCCAAGGCTTCTCATCAAGTATGC
TTCCAGTCCATATCTGGGATCCCCTCCAGGTTGAGTGGCAGTGGATCTGGGACAGATTTCACTCT
CAGTATCAACAGTGTGAAGACTGAAGATTTTGGAAATGTTTTCTGTCAACAGAGTAACAGCTGGCCT
CTCACGTTCCGGTGATGGGACAAAGCTGGAGCTGAAACGGGCTGATGCTGCACCAACTGTATCCAT
CTTCCCACCATCCAGTGAGCAGTTAACATCTGGAGGTGCCTCAGTCGTGTGCTTCTTGAACAACTT
CTACCCCAAAGACATCAATGTCAAGTGGAAGATTGATGGCAGTGAACGACAAAATGGCGTCCTGAA
CAGTTGGACTGATCAGGACAGCAAAGACAGCACCTACAGCATGAGCAGCACCCCTCACGTTGACCA
AGGACGAGTATGAACGACATAACAGCTATACCTGTGAGGCCACTCACAAGACATCAACTTCACCCA
TTGTCAAGAGCTTCAACAGGAATGAGTGTTAGAGGCCCTATTTTCTTTAGTTTGAATTTACTGTTATTC
GGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAAT
TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTCAGCAAGGACACAAAAAGATTTAATTTTATTA
AAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATT
TGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGA
TTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAA
ATTATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 27

SEQ ID NO: 63

Construção número 690 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). Sequência de codificação HA quimérico está sublinhada

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTGCACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTT
GTCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAATATTCACCT
ACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTG
ACAACAATTTTGTGCAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGA
AAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAATAG
TTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATG
ACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTAAAT
AAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGT
AATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTC
ATTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTATATTT
ACTAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCA
CAACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATT
CTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTG
AGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAA
AATGGAGAAAATAGTGCTTCTTCTTGAATAGTCAGTCTTGTAAAAGTGATCAGATTTGCATTGGTTACCATG
CAAACAATTCACAGAGCAGGTTGACACAATCATGGAAAAGAAGCTTACTGTTACACATGCCCAAGACATACT
GGAAAAGACACACAACGGGAAGCTCTGCCATCTAGATGGAGTGAAGCCTCTAATTTAAGAGATTGTAGTGT
AGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGA
GAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGCATTTCGCTGACTATGAGGAAGTGAAGGAGCAATT
GAGTTCAGTATCTTCATTTGAGAGGTTCCGAAATATTTCCCAAGAAAGCTCATGGCCCAACCACACCGTAACC
GGAGTGTGAGCATCATGCTCCCATATGGGGAAGCAGTTTACAGAAATTTCTATGGCTGACGGGGAAG
AATGGTTTGTACCCAAACCTGAGCAAGTCTATGCAACAACAAAGAAAAGAAGTCTTGTACTATGGGGTG
TTTCATCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGT
TTACATTTATAGCAGAAAATTCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATC
AATTACTACTGGACTCTGCTTGAACCCGGGGATACAATAATTTTGAAGCAAATGGAAATCTAATAGCGCCAA
GATATGCTTTGCACTGAGTAGAGGCTTTGGATCAGGAATTATGAAAAGTGAATTGGAATATGGTAACTGCAA
CACCAAGTGTCAAACCTCAATGGGGGGGATAAACTCTAGTATGCCATTCCACAACATACACCTCTCACCATC
GGGGAATGCCCCAAATATGTGAAATCAAACAGATTAGTCCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGA
GAGAGCAGAAAGAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAAT
GGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCA
CTCAAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGC
CGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTA
GATGTCTGCACCTATAATGCCGAACCTCTGTTCTCATGGAAAATGAGAGAACTCTAGACTTTTCATGACTCAA
ATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAGGAGCTGGGTAACGGTTGTT
TCCAGTTCTATCACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACATATCCGCAGTA
TTCAGAAGAAGCAAGATTTAAAGAGAGAGGAAATAAGTGGGGTAAATTTGGAATCAATAGGAACTTACCAATA
CTGTCAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGT
GCTCCAATGCATCGTTACAATGCAGAAATTCATTTAAGAGCTCTAAGTTAAATGCTTCTCGTCTCTATTT
ATAATATGGTTTGTATTGTTAATTTTGTCTGTAGAAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTT
GTATGAGATGAAGTGGTGAATGTAATTCATTTACATAAGTGGAGTCAGAAATCAGAAATGTTTCTCCATACTA
ACTAGACATGAAGACCTGCCGCTACAATTGTCTTATTTGAACAACATAAAATGAACATCTTTTGCCACAAC
TTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAATATCAGTTATCGAAA
TTCAATTAACAATCAACTTAACGTTATTAATACTAATTTTATATCATCCCTTTGATAAATGATAGTACACCAATT
AGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Figura 28

SEQ ID NO: 64

Construção número 691 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). Sequência de codificação HA quimérico está sublinhada

AAGCTTGCTAGCGGCCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTTGTC
AAATAACTCAAAAACCATAAAAGTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTACT
GTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATT
TTGTTGCAACATTTGAGAAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAG
GGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCAT
TGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGT
ACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAAATAAATTTTAAAAATTAAAGTTGAGTCATTTGATT
AAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTT
AATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTTTATATTTTCATAGATCAAATAAG
AGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTATATTTACTAAAAATCTAAGCCACGTAGGAGGAT
AACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCC
CACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCAC
ATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAA
GAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAGAAAATAGTGCTTCTTCTTGAATAGTCAGTCTTGTTA
AAAGTGATCAGATTTGCAATTTGGTTACCATGCAACAATTCACAGAGCAGGTTGACACAATCATGGAAAAGAACGT
TACTGTTACACATGCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCCTATTTAAAGGAATAGCCCC
ACTACAATTGGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAAACCCAGAATGCGAATTACTGATTTCCAAGGA
GTCATGGTCTACATTGTAGAAAACCAAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAG
GAACTGAGGGAGCAATTGAGTTCAGTATCTTCATTTGAGAGGTTGGAATATTTCCCAAGAAAGGCTCATGGCCCA
ACCACACCGTAACCGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCT
GACGGGGAAAGAAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAACAACAAGAAAAAGAGTCTTGTACT
ATGGGGTGTTCATCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTA
GTGTCTTCACATTATAGCAGAAAAATCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAA
TCAATTACTACTGGACTCTGCTTGAACCCGGGGATACAATAATTTGAGGCAAATGGAAATCTAATAGCGCCAAAG
ATATGCTTTCCCACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGCAACACC
AAGTGTCAAACCTCAATGGGGGCGATAAACTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAAT
GCCCAAAATATGTGAAATCAAACAGATTAGTCCCTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAA
GAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAAATGGTAGATGGTTGGT
ATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAGGCAATAGAT
GGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGGCCGTTGGAAGGGAATTTAATA
ACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGA
ACTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTTCATGACTCAAATGTTAAGAACCCTACGACAAGGTG
CGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAATGTGATAATGAAT
GTATGGAAAGTATAAGAAACGGAACGTACAACATCCGCAGTATTCAGAAGAAGCAAGATTAAAAAGAGAGGAAAT
AAGTGGGGTAAATTTGGAATCAATAGGAACTTACCAAAATCTGTCAATTTATTCAACAGTGGCGAGTTCCCTAGCA
CTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAAATTTGCATTTAAG
AGCTCTAAGTTAAATGCTTCTTCTGCTCTCTATTTATAATATGGTTTGTATTGTAAATTTGTTCTGTAGAAGAGC
TTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACCTGGTGTAAATGTAAATTCATTTACATAAGTGGAG
TCAGAAATCAGAAATGTTTCTCCATAACTAAGTACATGAAGACCTGCCGCTACAATTTGCTTATATTTGAACAAC
TAAATTTGAACATCTTTTGCACAACCTTTAAGTGGTTAATATAGTCAAATATATGGTCAAGTTCAATAGATTAAT
AATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACACTACTAATTTATATCATCCCTTTGA
TAAATGATAGTACCAATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Figura 29

SEQ ID NO: 65

Construção número 691 a partir de HindIII (no sítio de clonagem múltiplo, a montante do promotor de plastocianina) para EcoRI (imediatamente a jusante do terminador de plastocianina). Sequência de codificação HA quimérico está sublinhada

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATTTTATATGTTG
TCAAATAACTCAAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTAC
TACTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGAC
AACATTTTGTGCAACATTTGAGAAAATTTTGTGTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAA
AAGGAAGAGGGAGAAATAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTT
GTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGAC
GCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAATAAATTATTTTAAATTA
AAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAAT
TAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTT
TTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTATATTTACTAA
AAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCACAAACA
TCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCA
CACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTAC
ACTTTGATTCCCTTCAAACACATACAAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGCGA
AAAACGTTGCGATTTTCGGCTTAATGTTTTCTCTTCTTGTGTTGGTTCCCTTCTCAGATCTTCGCTGACACAATAT
GTATAGGCTACCATGCCAACCACTCAACCGACACTGTTGACACAGTACTTGAGAAGAATGTGACAGTGACACA
CTCTGTCAACCTACTTGAGGACAGTCACAATGGAAAACCTATGTCTACTAAAGGAATAGCCCCACTACAATTGG
GTAAATGCGAGCGTTGCCGGATGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCCAAGGAATCATGGTC
CTACATTTAGAAAACCAAATCCTGAGAAATGGAACATGTTACCCAGGGAGTTTCAACGACTATGAAGAATG
AAACACCTATTGAGCAGAAATAAACCATTTTGAGAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGA
AGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCTTTTAGAAATGTGGTATGGCTTA
TCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTTGGTACTG
TGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCCAACCATATATTTCCA
TTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGTAAACGGGGCAAAGTGG
AAGGATGGAGTTCTTCTGGACAATTTAAACCTAATGATGCAATCAACTCGAGAGTAATGGAATTTTATTG
CTCCAGAATATGCATACAAAATTTGCAAGAAAGGGGACTCAGCAATCATCACCTCAAATGCACCAATGGATGA
ATGTGATGCGAAGTGTCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAATGTACACCCAGTC
ACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCAT
CCATTCAATCCAGAGGTTTGTGTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATG
GGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAGTACACAAAATGC
CATTAAACGGGATTACAAACAAGGTCAATTCGTGAATTGAGAAAATGAACACTCAATTACAGCTGTGGGCAAAAG
AGTTCAACAAATTTGAAAGAAGGATGGAAAACCTAAATAAAAAAGTTGATGATGGGTTCTAGACATTTGGACA
TATAATGCAGAATTTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCCAATGTGAAGAATCT
GTATGAGAAAAGTAAAAAGCCAATTAAGAATAATGCCAAAGAAATAGGAAACGGGTGTTTGGAGTTCTATCACA
AGTGTAAACAATGAATGCATGGAGAGTGTGAAAAATGGTACCTATGACTATCCAAAATATTCCGAAGAATCAAAG
TTAAACAGGGGAGAAAATTTGATGGAGTGAAATTTGGAATCAATGGGAGTATACCAGATTCTGGCGATCTACTCAA
CTGTGCGCAGTTCCCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGGGTCTTT
GCAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAATGCTTCTCGTCTCCTATTTATAATATGGTTTGTAT
TGTTAATTTTGTCTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGATGAGATGAACGGT
GTAATGTAATTCATTTACATAAGTGGAGTCAGAAATCAGAATGTTTCTCCATAACTAAGACATGAAGACCT
GCCGCGTACAATTTGCTTATATTTGAACAACATAAATTTGAACATCTTTGCCACAACCTTTATAAGTGGTTAATAT
AGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAATATCAGTTATCGAAATTCATTAACAATCAACTTA
ACGTTATTAATACTAATTTTATATCATCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGA
GGCCTGGCTGGCCGAATTC

Figura 30

SEQ ID NO: 66

Construção 732 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de H1 A/Brisbane/59/2007 está sublinhada

TTAATTAAGAATTTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
GATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAAGATGACCCGCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGACAAATCCCACTATCCTTCGCAAGA
CCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAGGTTTTGATAAAAGCGAACGT
GGGGAAACCCGAACCAACCTTCTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTTGTCTTTCTTGC
GTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAG
ATCTCTTTGTGGACACGTAGTGCGGGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGGGTGTGGTCTTGGGA
AAAGAAAGCTTGCTGGAGGCTGCTGTTTCAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCA
ATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTTGTGATTGGTCTATAAGAA
ATCTAGTATTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCT
GCCCAAATTTGTGGGGCCCATGAAAGTAAACTACTGGTCTGTATGACACATTACAGCTACATATGCAGACACA
ATATGTATAGGCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGAAAAGAAATGTGACAGTGACAC
ACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAACTATGTCTATTAAAGGAATAGCCCCACTACAATTGGG
TAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCGAAGGAGTCATGGTCTCTAC
ATTGTAGAAAAACCAAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCCGCTGACTATGAGGAACCTGAGGGAG
CAATTGAGTTTCAGTATCTTCAATTTGAGAGGTTTCAAAATATTTCCCAAGAAAGCTCATGGCCCAACCACCCGTAA
CCGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAGA
ATGGTTTTGTACCCAAACCTGAGCAAGTCTATGCAAAACAACAAAGAAAAAGAGTCTTGTACTATGGGGTGTTC
TCACCCGGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAATGCTTATGTCTCTGTAGTGTCTTACATT
ATAGCAGAAAAATCACCCCAAGAAATAGCCAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTG
GACTCTGCTTGAACCCGGGGATACAATAATTTTGGGGCAATGGAAATCTAATAGCGCCCAAGATATGCTTCCCA
CTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAAGTGCCAAACAC
CTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAAGGTACACCCAGTCACAATAGGAGAGTGTCCAAAGTATGT
CAGGAGTGCAAAATTAAGGATCGTTACAGGACTAAGGAACATCCCATCCATTCAATCCAGAGGTTTGTTTGGAGCC
ATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATGTTGGTATGGTTATCATCATCAGAATGAGCAA
GGATCTGGCTATGCTGCAGATCAAAAAAGCACACAAATGCCATTAATGGGATTACAACAAGGTCAATTCTGTAA
TTGAGAAATGAACACTCAATTCACAGCAGTGGGCAAGAGTTCAACAAATTTGAAAGAAGGATGGAAAACCTTGAA
TAAAAAAGTTGATGATGGGTTTATAGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGA
CTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAA
ATAGGAAATGGGTGTTTTGAGTTCTATCACAAGTGTAAACGATGAATGCATGGAGAGTGTAAAGAATGGAACCTATG
ACTATCCAAAATATTCGAAGAATCAAAGTTAAACAGGGAGAAAAATTGATGGAGTAAATTTGGAATCAATGGGAGT
CTATCAGATTCTGGCGATCTACTCAACAGTCGCCAGTTCTCTGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTC
TGGATGTGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAAAGGCCATTTTCTTTAGTTGAATTTACTGTTA
TTCCGGTGTGCTTTCTATGTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTG
TGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTTATTAATAAAAAAAAAAAAAA
AGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAA
TCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG
CATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACCGGATAGAAAACAAAT
ATAGCCGCGAAACTAGGATAAATTATCGCGCGCGGTGTCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGC
GCGC

Figura 31

SEQ ID NO: 67

Sequência de codificação, a partir de ATG para stop, construção intermedia número 787

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTTGGTTCCTTCTCAGATCTTCG
CTGACACAATATGTATAGGCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGAAAAGA
ATGTGACAGTGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAAACCTATGTCTATTTAAAGG
AATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCCAATT
ACTGATTTCCAAGGAGTCATGGTCCTACATTGTAGAAAAACCAAATCCTGAGAATGGAACATGTTACCCA
GGGCATTTGCTGACTATGAGGAACTGAGGGAGCAATTGAGTTCAGTATCTTCATTTGAGAGGTTTCGAA
ATATTCCCCAAAGAAAGCTCATGGCCCAACCACACCGTAACCGGAGTGTGAGCATCATGCTCCCATAT
GGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAACCTGAGC
AAGTCCTATGCAACAACAAAGAAAAAGAAGTCCTTGTACTATGGGGTGTTCATCACCCGCCAAACATA
GGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAA
AATTCACCCCAGAAATAGCCAAAAGACCCAAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGA
CTCTGCTTGAACCCGGGGATACAATAATTTGAGGCAAATGGAAATCTAATAGCGCCAAGATATGCTTT
CGCACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAA
GTGCCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCCTTTCCAGAACGTACACCCAGTCACAATAGG
AGAGTGTCCAAAGTATGTGAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCAT
TCAATCCAGAGGTTTGTGTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATG
GTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAAGCACACAAA
ATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCAATTCACAGCAGT
GGGCAAAGAGTTCAACAAATTGGAAAGAAGGATGGAAAACCTTGAATAAAAAAGTTGATGATGGGTTTATA
GACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACT
CCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGG
GTGTTTTGAGTTCTATCACAAGTGTAACGATGAATGCATGGAGAGTGTAAGAATGGAACCTTATGACTAT
CCAAAATATTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGA
GTCTATCAGATTCTGGCGATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCA
ATCAGCTTCTGGATGTGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAA

Figura 32

SEQ ID NO: 68

Construção 733 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de SpPDI-H1 A/Brisbane/59/2007 está sublinhada

TTAATTAAGAATTTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAAG
ATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTG
CCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTT
CAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCAGTATCCTTCGCAAGACCGT
TCCTCTATATAAGGAAGTTTCATTTTCAATTTGGAGAGGTATTAATACTTAATAGGTTTTGATAAAAGCGAACGTGGGGA
AACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAGCAAACCTTCTCTCTTGTCTTTCTTCCGTGAGC
GATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCTT
GTGGACACGTAGTGGCGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTCTGGTGTGGTCTTGGGAAAAGAAAGC
TTGCTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTAC
TTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTTCTGATTGTTCTATAAGAAATCTAGTATT
TCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGT
CGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCCCTCTCAGATCTTCGC
TGACACAATATGTATAGGCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGAAAAGAATGTGACAG
TGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAAACCTATGTCTATTAAAAGGAATAGCCCCACTACAAT
TGGGTAATTGCAGCGTTGCGCGGTGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCCAAGGAGTCATGGTC
CTACATTGTAGAAAAACCAAATCCTGAGAATGGAACATGTTACCCAGGGCATTTGCTGACTATGAGGAACTGAGGG
AGCAATTGAGTTGAGTATCTTCATTTGAGAGGTTGAAATATTCCCCAAAGAAAGCTCATGGCCCCAACACACCGTA
ACCGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTACAGAAATTTGCTATGGCTGACGGGGAAGA
ATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAACAACAAAGAAAAAGAGTCTTTGACTATGGGGTGTTCAT
CACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAATGCTTATGTCTGTAGTGTCTTCACATTAT
AGCAGAAAAATTCACCCAGAAATAGCCAAAGAGCCCAAGTAAGAGATCAAGAAGGAAGATCAATTACTACTGGAC
TCTGTTGAACCCGGGGATACAATAATATTGAGCAAAATGGAAATCTAATAGCGCCAAGATATGCTTTGCACTGA
GTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAAGTGCCAAACACCTCAG
GGAGCTATAAACAGCAGTCTTCTTTCCAGAAGTACACCCAGTACAATAGGAGAGTGTCCAAAGTATGTCAGGAG
TGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCATTCAATCCAGAGGTTTGTTTGGAGCCATTGCCG
GTTTCATTGAAGSGGGGGTGGACTGGATGGTAGATGGTTGGTATGGTTATCATCATCAGAAATGAGCAAGGATCTGG
CTATGCTGCAGATCAAAAAAGCACACAAAATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAAT
GAACACTCAATTCACAGCAGTGGGCAAGAGTTCACAAAATTTGAAAGAAGGATGGAAAACCTTGAATAAAAAAGTTG
ATGATGGGTTTATAGACATTGGACATATAATGCAGAACGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTC
ATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGGGT
GTTTTGAGTTCTATCACAAGTGAACGATGAATGCATGGAGAGTGTAAAGAATGGAACCTTATGACTATCCAAAATATT
CCGAAGAATCAAAGTTAAACAGGGAGAAAATTTGATGGAGTGAAATTTGGAATCAATGGGAGTCTATCAGATTCTGGCG
ATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGG
GTCTTTACAGTGTAGAATATGCATCTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCCGTGTGCATTTCTAT
GTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGG
TCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAA
GCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATG
ATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGG
TTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATT
ATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 33

SEQ ID NO: 69

Construção 734 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de HA quimérico está sublinhada

TTAATTAAGAATTGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAAG
ATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCCATAAAGGAAAAGGCCATCGTTGAAGATGCCTCTG
CCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAACGTTCCAACCACGTCTT
CAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCC
TTCTCTATATAAGGAAGTTCATTTTCTTTGGAGAGGTATTAATAATCTTAATAGGTTTGGATAAAAGCGAACGTGGGG
AAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTTCTTGGCGTGAG
CGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCT
TTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCTTGGGAAAAGAA
AGCTTGCTGGAGGCTGCTGTTCAAGCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCT
CTACTTCTGCTTGACGAGGTATTGTTGCCCTGTACTTCTTCTTCTTCTTCTGCTGATTGGTCTATAAGAAATCTAG
TATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAAGTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAA
ATTTGTGGGCCCATTGGAGAAAATAGTGCTTCTTCTTGAATAGTCAGTCTTGTAAAAGTGATCAGATTGCAITG
GTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGA
CATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGT
GTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGA
AGGCCAATCCAACCAATGACCTCTGTTACCCAGGGCATTTCGCTGACTATGAGGAAGTGAAGGAGCAATTGAGTTC
AGTATCTTTCATTGAGAGGTTTCAAATATTCGCCAAAGAAAGCTCATGGCCCAACACACCGTAACCGGAGTGTCA
GCATCATGCTCCCATATGGGGAAAGCAGTTTTACAGAAAATTTGCTATGGCTGACGGGGAAGAATGGTTTGTACC
CAAACCTGAGCAAGTCTATGCAAAACAACAAAGAAAAAGAGTCTTGTACTATGGGGTGTTCATCACCGGCCAA
CTACCTGACCAAAAGGCCCTCTATCATAAGAAAATGCTTATGTCTCTGTAGTGTCTTACATTATAGCAGAAAAT
TCACCCCGAAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGACTCTGCTTGA
ACCCGGGGATACAATAATATTTGAGGCAAAATGGAAATCTAATAGCGCCAAGATATGCTTTCCGCACTGAGTAGAGGC
TTTGGATCAGGAATTATGAAAAGTGAATTGGAATATGGTAAGTGAACACCAAGTGTCAAACCTCAATGGGGCGGA
TAAACTCTAGTATGCCATCCACAACATACACCCTCTCACCATCGGGGAATGCCCAATATGTGAAATCAAACAGA
TTAGTCTTGAACAGGGCTCAGAAAATAGCCCTCAAAGAGAGAGCAGAGAAGAAAAAGAGAGGACTATTTGGAGCTA
TAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGG
GGAGTGGGTACGCTGCAGACAAAGAATCCAATCAAAGGCAATAGATGGAGTGACCAATAAGGTCAACTCAATCAT
TGACAAAATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTAATAACTTAGAAAGGAGAATAGAGAATTTAAACA
AGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAACCTCTGGTTCTCATGGAAAATGAGAGAAC
TCTAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTACGGATAATGCAAAGGAG
CTGGGTAACGGTGTGTTTGGAGTTCTATCACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAA
CTATCCGCGAGTATTCAGAAGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAATTTGGAATCAATAGGAACT
TACCAAACTACTGTCAATTTATTCACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATCGCTGGTCTATCTTTATG
GATGTGCTCCAATGGATCGTTACAATGCAGAAATTTGCATTTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATT
CGGTGTGCAATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTG
AGCTCCTGTTTAGCAGGTCTGCCCTTCAGCAAGGACACAAAAGATTTAATTTTATTAATAAAAAAAAAAAAAAAAAAGA
CCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCC
TGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCAT
GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACCGGATAGAAAACAAAATATA
GCGCGCAAACTAGGATAAATTTATCGCGCGCGGTGTCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGG
CC

Figura 34

SEQ ID NO: 70

Fragment sintetizado Dralll-Plasto(-84+1)-H1 A/Brisbane/59/07-Sac1

CACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAAT
CATCTTGAGAGAAAATGAAGACTATCATTGCTTTGAGCTACATTCTATGTCTGGTTTTCACTCAAAAAC
TCCCGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATA
GTGAAAACAATCACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTCCTCAACA
GGTGAAATATGCGACAGTCCTCATCAGATCCTTGATGGAGAAAACGACACTAATAGATGCTCTATTG
GGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTA
CAGCAACTGTTACCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCAGTGTGCTCATCCGGCA
CACTGGAGTTTAAACAATGAAAGTTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGCTCTGCTTGC
ATAAGGAGATCTAATAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAATTCAAATACCCAG
CATTGAACGTGACTATGCCAAACAATGAAAAATTTGACAAATTGTACATTTGGGGGGTTCACCACCCG
GGTACGGACAATGACCAAATCTTCCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAG
AAGCCAACAAACTGTAATCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCCAGCAGAATAA
GCATCTATTGGACAATAGTAAACCCGGGAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTC
CTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAA
TGCAATTCTGAATGCATCACTCCAAACGGAAGCATTCCCAATGACAAACCATTCCAAAATGTAAACAGG
ATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAGGGATGCGAAA
TGTACCAGAGAAACAACTAGAGGCATATTTGGCGCAATCGCGGGTTTCATAGAAAATGGTTGGGAGG
GAATGGTGGATGGTTGGTATGGTTTCAGGCATCAAATTTCTGAGGGAATAGGACAAGCAGCAGATCTC
AAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATAGGTTGATCGGGAAAACCAACGA
GAAATCCATCAGATTGAAAAAGAGTTCTCAGAAGTCGAAGGGAGAATCCAGGACCTTGAGAAATATG
TTGAGGACACCAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACAT
ACAATTGATCTAACTGACTCAGAAATGAACAACTGTTTGAAAAAACAAGAAGCAACTGAGGGAAAAT
GCTGAGGATATGGGCAATGGTTGTTTCAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATC
AGAAATGGAAC TTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGGG
CGTTGAGCTGAAGTCAGGATACAAAGATTGGATACTATGGATTTCTTTGCCATATCATGTTTTTTGCT
TTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAGGCAACATTAGGTGCAACATTTGCA
TTTGAGAGCTC

Figura 35

SEQ ID NO: 71

Construção 736 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de SpPDI-H1 A/Brisbane/59/2007 está sublinhada

TTAATTAAGAATTGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACCTTATTGAGAA
GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAG
ACCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAAATCCTAATAGGTTTTGATAAAGCGAAGC
TGGGGAAACCCGAACCAACCTTCTTCTAAACTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAAGTACAACGTTTTCTTCACTGAAGCGAAATCAA
GATCTCTTTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGTGTGGTCTTGGG
AAAAGAAAGCTTGTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTG
CAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTTGTGCTGATTGGTCTATAAG
AAATCTAGTATTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCCGAACCTGGAGAAAGATTGTTAAGCTTCTGTATATT
CTGCCCAATTTGTGGGCCCCATGGCGAAAAACGTTGCGATTTTGGGCTTATTGTTTCTCTCTTGTGTTGGTTC
CTTCTCAGATCTTCGCTCAAAAACTTCCCGGAAATGACAAACAGCACGGCAACGCTGTGCCTTGGGCAACCATGCGAG
TAGCAAACCGAAGCATAGTGAAAACAATCACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAG
TTCTCAACAGGTGAAATATGCGACAGTCTCATCAGATCCTTGATGGAGAAAACCTGCACACTAATAGATGCTCTA
TTGGGAGACCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCAGCAAAAGCCTACAGG
AACTGTTACCTTATGATGTGCGCGATTATGCCTCCCTTAGGTCACTAGTTGGCTCATCCGGCACACTGGAGTTTA
ACAATGAAAGTTTCAATTGGACTGGAGTCACTCAAAACCGAACAAGCTCTGCTTGATAAGGAGATCTAATAACAG
TTCTTTAGTAGATTGAATTGGTTGACCCACTTAAATTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATG
AAAAATTTGACAAATTTGTACATTTGGGGGTTTCAACCCCGGGTACGGACAATGACCAAATCTTCTGTATGCTCA
AGCATCAGGAAGAATCACAGTCTCTACCAAAGAAGCCAAACAACCTGTAATCCCGAATATCGGATCTAGACCCAG
AGTAAGGAATATCCCGAGCAGAATAAGCATCTATTGGACAATAGTAAACCGGGAGACATACTTTTGATTAAACAGC
ACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCAC
CCATTGGCAAATGCAATTTCTGAATGCATCACTCCAAACCGGAAGCATTCCCAATGACAAACCATTCCAAATGTAAA
CAGGATCACATACGGGGCTGTCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAGGGATGCGAAATGT
ACCAGAGAAACAAACTAGAGGCATATTTGGCGCAATCGCGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGG
ATGGTTGGTATGGTTTCAGGCATCAAAATTTCTGAGGGAATAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAG
CAATCGATCAAATCAATGGGAAGCTGAATAGGTTGATCGGGAAAACCAACGAGAAATCCATCAGATTGAAAAAGA
GTTCTCAGAAGTCCAAGGGGAGAATCCAGGACCTTGAGAAATATGTTGAGGACACCAAAATAGATCTCTGGTCATA
CAACGCGGAGCTTCTTGTTCCTGGGAGAACCAACATACAATTGATCTAACTGACTCAGAAATGAACAACTGTTT
GAAAAAACAAAGAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAAATGTG
ACAATGCCTGCATAGGATCAATCAGAAATGGAACCTTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCG
GTTCCAGATCAAGGGCGTTGAGCTGAAGTCAGGATACAAAGATTGGATACTATGGATTTCTTTGCCATATCATGT
TTTTGCTTTGTGTGCTTTGTTGGGGTTTCATCATGTGGGCTGCCAAAAAGGCAACATTAGGTGCAACATTTGCA
TTTGAAGGCCTATTTCTTAGTTTGAATTTACTGTTATTGCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTG
CTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACA
CAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGAT
CGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCT
GTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGT
CCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTG
TCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 36

SEQ ID NO: 72

Construção 737 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação HA quimérico está sublinhada

TTAATTAAGAATTTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGAGAAGA
TAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGC
CGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCCTTCA
AAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTC
CTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATCTTAATAGGTTTTGATAAAGCGAACGTGGGGAAA
CCCGAACCAACCTTCTTCTAACTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTCTTCTTCTTGGCTGAGCGAT
CTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTG
CTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCATATCTCTACTTCT
GCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTTGTGATTGGTCTATAAGAAATCTAGTATTTCTTT
GAAACAGAGTTTTCCCGTGGTTTTCGAAGTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGG
CCCATTGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTCTTGTGTTGGTTCTTCTCAGATCTTCGGCTCAA
AACTTCCCGGAAATGACAACAGCAAGGCAACGCTGTGCTTGGGCAACCATGCAGTACCAACGGAACGATAGTGA
AAACAATCACGAATGACCAAAATTGAAGTTACTAATGCTACTGAGCTGGTTTCAGAGTTTCTCAACAGGTGAAATATGCG
ACAGTCTCATCAGATCTTGTATGGAGAAAACCTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGC
TTCCAAAATAAGAAATGGGACCTTTTGTGTAACGCGAGCAAAGCCTACAGCAACTGTTACGCTTATGATGTGCCGGAT
TATGCTTCCCTTAGGTCAGTGTGCTCATCCGGCACACTGGAGTTAACAATGAAAGTTTCAATTGGACTGGAGT
CACTCAAAACGGAACAAGCTCTGCTTCATAAGGAGATCTAATAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCA
CTTAAATTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAAAAATTTGACAAATTTGATATTGGGGGGT
TCACCAACCGGGTACGGACAATGACCAAAATCTTCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAA
GAAGCCAAACAACTGTAATCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCGAGCAGAATAAGCATCTAT
TGGACAATAGTAAACCGGGAGACATACTTTGATTAAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAA
ATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAA
CGGAAGCATTCCCAATGACAAACCATTCCAAATGTAAACAGGATCACATACGGGGCCTGTCCAGATATGTTAAGC
AAAACACTCTGAAATTGGCAACAGGGATGCGAAATGTACCAGAGAAAACAACTAGAGGCATATTTGGCGCAATCGCG
GGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTATGGTTTCAGGCATCAAAATTTCTGAGGGAATAG
GACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAAATCAATGGGAAGCTGAATAGGTTGATCGGGAA
AACCAACGAGAAATTCATCAGATTGAAAAAGAGTTCTCAGAAGTCGAAGGGAGAATCCAGGACCTTGAGAAATATG
TTGAGGACACCAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTGCTTGGAGAACCAACATACAATTGAT
CTAACTGACTCAGAAATGAACAAACTGTTTGAACAAACAAAGAAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAA
TGGTTGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACCTTATGACCACGATGT
ATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGGGCGTTGAGCTGAAGTCAATAGGAACTTACCAAAATAC
TGTCATTTTATCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCA
ATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTT
CTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAG
CAGGTGCTCCCTTCAGCAAGGACACAAAAAGATTTAATTTTATTAACAAAAAAGACCGGGAATTCGAT
ATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTG
CGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGA
TGGGTTTTATGATTAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAACTAGGAT
AAATTATCGCGCGCGGTGTCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 37

SEQ ID NO: 73

Fragment sintetizado Dralll-Plasto(-84+1)-HA B/Florida/4/06-Sacl

CACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAA
TCATCTTGAGAGAAAATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAAT
CTGCACTGGAATAACATCTTCAAACCTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTCAAT
GTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAG
GACCAGAGGGAAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACC
AATGTGTGTGGGGACCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCC
GGGTGCTTTCTATAATGCACGACAGAACAACAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGA
AAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGGCACCAGGAGGACCTACAGACTT
GGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTG
TCCCAAAGGACAACAACAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAA
GGGGAAGACCAAATCACTGTTTGGGGTTCCATTGAGATAACAAAACCCAAATGAAGAACCTCTATG
GAGACTCAAATCCTCAAAAGTTACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATT
GGCAGCTTCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTTGATTACA
TGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTACCAAAAGAGGTGTTTTGTTGCCTCAAAAGGTG
TGGTGC GCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCAGATTGCC
TTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAAGCCATA
GGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTG
CAAACTATTAAAGGAAAAGGGTTTCTTCGGAGCTATTGCTGGTTTCTAGAAAGGAGGATGGGAAGG
AATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACCT
TAAGAGTACGCAAGAAGCTATAAACAAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAA
GAATCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAA
GTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAATAGAAGTTGCAGTCTTGCTTTCCAACGAAG
GAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACCTAAAGAAAATGCTGGGTCCC
TCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAACACAAGTGCAACCAGACCTGCTTAGACA
GGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCACTTTTGATTCACTGAACATTACT
GCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCT
AGTTTGGCTGTAACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGC
TCCATCTGTCTATAAGAGCTC

Figura 38

SEQ ID NO: 74

Construção 739 a partir de Pacl (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de SpPDI-HA B/Florida/4/2006 está sublinhada

TTAATTAAGAATTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAG
ACCTTTCCTCTATATAAGGAAGTTCATTTTCATTTGGAGAGGTATTAAAAATCTTAATAGGTTTTGATAAAAGCGAACG
TGGGGAAACCGGAACCAAACCTTCTTCTAAACTGTCTCTCATCTCTCTTAAAGCAAACCTTCTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAA
GATCTCTTTGTGACACGTAGTGGCGCGCCATTAAATAACGTGTACTTGTCTTATTCTTGTGGTGTGGTCTTGGG
AAAAGAAAGCTTCTGGAGGCTGCTGTTTCAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTG
CAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGTGCTGATTGGTCTATAAG
AAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTGGAACCTTGAGAAAGATTGTTAAGCTTCTGTATATT
CTGCCCAAATTTGTGGGGCCCATGGCGAAAAACGTTGCGATTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTC
CTTCTCAGATCTTCGCTGATCGAATCTGCACTGGAAATAACATCTTCAAACCTCACCTCATGTGGTCAAACAGCCAC
TCAAGGGGAGGTCAATGTGACTGGTGTGATACCCTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAA
GGAACAAGGACCCAGAGGGAACCTATGCCAGACTGTCTCAACTGCACAGATCTGGATGGCTTTGGGCAGACC
AATGTGTGTGGGGACCCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTTACATCCGGGTGCTT
TCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAAATCAGGCTATCA
ACCCAAAACGTCATCGATGCGGAAAAGGCCACGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAA
CGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAA
CCCCTAACAGTAGAAGTACCATACTTTGTACAGAAGGGGAAGACCAAAATCACTGTTTGGGGGTTCCATTTCAGAT
AACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCA
CACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGAAAGACGGAGGACTACCACAAAGCGGCAGGATTG
TTGTTGATTACATGATGCAAAAAACCTGGGAAAAACAGGAACAATTGTCTACCAAGAGGTGTTTTGTTGCCTCAAAA
GGTGTGGTGGCGAGTGGCAGGAGCAAAAGTAATAAAAGGGTCTTGCCTTTAATTGGTGAAGCAGATTGCCTTCA
TGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCC
AATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAATATAGACCTCCTGCAAAACCTATTAAAGGAAAGG
GGTTTCCTCGAGCTATTGCTGGTTTTCTAGAAGGAGGATGGGAAGGAATGATTGCAGGCTGGCAGCGATACAC
ATCTCACGGAGCACATGGAGTGGCAGTGGCGGGCGGACCTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAA
AAAATCTCAATTTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGACTAAGTGGTGCCATGGATGAACCTCCACAA
CGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAAATAGAAGCTTGCAGT
CTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACCTGAGAGAAAACTAAAGAAAATG
CTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTGGAAACCAAAACACAAGTGAACCCAGACCTGCTTAGAC
AGGATAGCTGCTGGCACCTTTAATGCAGGAGAAATTTCTCTCCCACCTTTTGATTCACTGAACATTACTGCTGCAT
CTTTAAATGATGATGGATTGGATAACCATATACTGCTCTATTACTCAACTGCTGCTCTAGTTTGGCTGTAACA
TTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAAAGGCCTAT
TTTTCTTTAGTTTGAATTTACTGTTATTGGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGT
TTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGCTCCCTTCAGCAAGGACACAAAAAGATTTTA
ATTTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTT
GGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTTTCGATGATTATCATATAATTTCTGTTGAATTACGTT
AAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATA
CATTTAATACCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTA
CTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 39

SEQ ID NO: 75

Construção 745 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação HA quimérico está sublinhada

TTAATTAAGAATTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
GATAGTGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGA
CCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGCGAACGT
GGGGAAACCCGAACCAACCTTCTCTAAACTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGC
GTGAGCGATCTTCAACGTTGTGATGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAG
ATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTTATTCTTGTCTGGTGTGGTCTTGGGA
AAAGAAAGCTTGTCTGGAGGCTGCTGTTGAGCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCA
ATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTCTATAAGAA
ATCTAGTATTTTCTTGAACAGAGTTTTCCCGTGGTTTTGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCT
GCCCAAATTTGTGCGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTCTTGTGTTGGTTCCTT
CTCAGATCTTGGCTGATCGAATCTGCACTGGAATAACATCTTCAAACCTCAGCTCATGTGGTCAAACAGCCACTCA
AGGGGAGGTCAATGTGACTGGTGTGATACCACTAACAACAACCAACAAAAATCTTATTTTCAAATCTCAAAGGA
ACAAGGACAGAGGGAACTATGCCAGACTGTCTCAACTGCAGAGATCTGGATGTGGCTTGGGCAGACCAATG
TGTGTGGGGACCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTTCT
ATAATGCACGACAGAACAAAAATCAGGCAACTACCAATCTTCTCAGAGGATATGAAAAATATCAGGCTATCAACCC
AAAACGTCATCGATGCGGAAAGGCCACGAGGAGGACCCTACAGACTTGAACCTCAGGATCTTGCCTAACGCTA
CCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCAC
TAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAATCACTGTTTGGGGGTTCATTTCAGATAACAA
AACCCAAATGAAGAACCCTCTATGGAGACTCAAATCCTCAAAGTTTACCTCATCTGCTAATGGAGTAACCACACAC
TATGTTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGT
GATTACATGATGCAAAAAACCTGGGAAAAACAGGAACAATTTGTCTACCAAGAGGTGTTTTGTTCCTCAAAGGTGT
GGTGGCGGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCCTGCCCTTAATTGGTGAAGCAGATTGCCCTCATGAAA
AATACGGTTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATG
GGTGAACACACCTTTGAAGCTCGCCAATGGAACCAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGGTTTC
TTCCGAGCTATTGCTGTTTTCTAGAGGAGGATGGCAAGGAATGATTGCAGGCTGGCAGCGATACACATCTCAC
GGAGCACATGGAGTGGCAGTGGCGGCGGACCTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTC
AATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGACTAAGTGGTGGCATGGATGAACCTCCACAACGAAATAC
TCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAATAGAAGTTGCAGTCTTGCTTTC
CAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAAATGCTGGGTCCC
TCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAACACAAGTGCAACCAGACCTGCTTAGACAGGATAGCT
GCTGGCACCTTTAATGCAGGAGAAATTTCTCTCCCCACTTTTGATTCACTGAACATTACTGCTGCATCTTTAAATGA
TGATGGATTGGATAACTACCAATACTGTCAATTTATTCAACAGTGGCGAGTTCCTTAGCACTGGCAATCATGATG
GCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCTATTTCTTTAG
TTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTATG
TAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCTGCCCTTCAGCAAGGACACAAAAAGATTTTAATTTATTAA
AAAAAAGGAAAGGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAG
TTTCTTAAGATTGAATCCTGTTGCCGGTCTTGGGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAAT
AATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGC
GATAGAAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGA
GTCTCAAGCTTCGGCGCGGCC

Figura 40

SEQ ID NO: 76

Sequência de codificação Msj1

ATGTTTGGGCGCGGACCAACAAGGAAGAGTGATAACACCAAATATTACGATATTCTTGGTGTTCAAA
AAGTGCTAGTGAAGATGAAATCAAGAAAGCCTATAGAAAGGCAGCGATGAAGAACCATCCAGATAAG
GGTGGGGATCCTGAGAAGTTCAAGGAGTTGGGCCAAGCATATGAAGTGTTGAGCGATCCTGAAAAG
AAAGAACTGTATGATCAATATGGTGAAGATGCCCTTAAAGAAGGAATGGGGGGAGGCGCAGGAAGC
TCATTTTCATAATCCGTTTGATATTTTCGAATCATTTTTTGGTGCAGGCTTTGGTGGTGGTGGTCCTTCA
CGCGCAAGAAGACAGAAGCAAGGAGAAGATGTGGTGCATTCTATAAAGGTTTCCTTGGAGGATGTGT
ATAACGGCACTACAAAGAAGCTATCACTTTCTAGGAATGCACTGTGCTCAAAATGTAAAGGGAAAGGT
TCAAAAAGTGGAAGCTGCTGGAAGGTGTTTTGGATGCCAGGGCACAGGTATGAAGATTACCAGAAGG
CAAATTGGAAGTGGGCATGATTCAACAAATGCAACACGTCTGTCTGACTGCAAAGGAACAGGCGAGG
TCATTAGTGAGAGAGATAGATGCCCTCAATGCAAGGGAAACAAGATTACTCAAGAAAAGAAGGTGCT
GGAGGTGCATGTGGAAAAGGGGATGCAGCAGGGTCACAAGATTGTATTGCAAGGACAAGCTGATGA
AGCTCCTGATACAATCACAGGAGACATAGTTTTTGTCTTGCAAGTAAAGGGACATCCGAAGTTTCGGA
GGGAGCGTGATGACCTCCACATTGAACACAATTTGAGCTTAAGTGAAGGCTCTCTGTGGCTTCCAGTT
TAATGTCACACATCTTGATGGAAGGCAACTATTGGTCAAATCGAACCCCGGCGAAGTCATCAAGCCA
GGTCAACATAAAGCTATAAATGATGAGGGAATGCCACAACATGGTAGGCCGTTTCATGAAGGGACGCC
TATACATCAAGTTTAGTGTTGATTTCCCGGATTCCGGGTTTTCTTTCCCAAGCCAAAGCCTGGAATTA
GAAAAGATATTACCTCAAAAGACAAGCAAGAAGTGTCCCAAAAGGAGGTAGATGATTGTGAGGAGA
CCACCCTGCATGATGTCAATATTGCAGAGGAGATGAGTCGAAAAGAAGCAACAATACCGTGAGGCATA
TGATGACGATGATGATGAAGATGATGAGCACTCGCAGCCTCGGGTGCAATGCGCTCAACAGTAG

Figura 41

SEQ ID NO: 77

Construção número R850 a partir de HindIII (no sítio de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). Sequência de codificação HSP40 está sublinhada

AAGCTTGCATGCCTGCAGGTCCGACTCTAGAGGATCCCCGGGCTGGTCTGTACATTCATCTTGCCGCCCTTGCAATT
CACTTGGCCACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGCAAGTGCACCTCG
AGGGTCAGAACTGTATATCATATCTATGTGAGAGAAAGGGGAACATTTGAGATGGAGTCCATTTACTTGAGGTAT
ACTTATTATTTTCATCAATAAATTTGTATACTTCTTATTTAGATCAATAAATTTGTCTTAAGCTATAATCCAAAATAA
ATTACGATCAAATATGCAAATGTTAGCCAGTACTTGTGTTAACTTGATGGCATCTCTTGGTTTCTTGGCAATCAC
ATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGATAAAC
AGCTTAATTAAGCGCTTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAGAGCATCTCTTGGTTTCTCTGGCA
ATCATATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACCTGTGTTGGG
TAAACAGCTTAATTAAGTGCCTTATAGAATAAGCGCTTATCATATAAGTGCCTTTGTACAGTTATTTCTATGAAAGTA
GAAGAAATAGTCATATTGTTTAAATATAAGCTATCCTGGAGAGCTTGTGGAAATAACCAGAAAGAACTTATGGAC
ACGTCATGAGCTGTTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTATGCCAGTAGATAAATCAATAAAGTC
AATCTAAACAGACCCATAATCCATTATGGTACCTATCATTATGCTTATCCATCTTTAATAAGATGTCATGAGATA
ACATAAATGACACACATATTTTGACACAAATGGGCAGATCTAGCAATTTAACTCTGGAGTGCCTTCAAGACTGCTG
TTCTTACGAAGTTCACGTCCTGAATCATGTTCTGTATGGAAGCCTGAAAGACCTCAAATTTCTAAAGGTGGCGA
TAAATTTGAAGTTTACAAAATATACCCTGCGGGCTTGACACAGAGGGCAAGCTCTTTATACCTTCCAGTTCAACGG
GGATGTTGATTTTCAAGTCACTTGGAGAGCAATCCTTGTGCCAAGTTTGAAGTAATTTTGTGTAGCATATGTTG
AGCTACCTACAATTTACATGATCACCTAGCATTAGCTCTTTCACTTAACTGAGAGAATGAAGTTTTAGGAATGAGTA
TGACCATGGAGTGGGCATGGCTTTGTAATGCCCTACCCTACTTTGCCCACTCATGGGGATTACATTGAGAAAA
TATACATGACTTCAACCATACTTAAACCCCTTTTGTAGATAACTGAATGTTTATTTAATGTTGGGTTGTAGTG
TTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTGGGTCTGTACTGTTATTTATTTATTTTT
TTTTAGCAGAAACACCTTATCTTTTGTGTTTGAATGTAGAATGAAATAAAAGAAAGAAATAACATCATCG
GCCCGCTTGTCTAATTTCCGGGCAGTTAGGATCCTCTCCGGTCACCGGAAAGTTTCAGTAGAAGAAACAAAACAC
CGTGACTAAAATGATACTATTATTTATTTATTTGTGTTTTCTTTTTCTACCGGAAGTTTGTAGAACGGATCCCAAC
TCGTTCCGGGGCCCTACAACGAAACAAAAGAGATATTTCTCTCTCTTCAGAAATGTAAGTTTTCTTTACAG
ATACCATTCACCATTTGATTGAGATGTGGTGAAGTATGAGATAAAGCATACTAATTTGACTCTTGGAAACCCATAAA
GTTTATGTTTCTTTTATAAATTTCTAATTTCTTCTGTGTAACCTATTTCATGTATTATTTTCTTTAAATTTACATGTCAT
GATTTATGCGGCGACTTGAATTTAAATCTAGGAGGGGCAGACATTGAACAATCCCAATATTTTAATAACTTATG
CAAGATTTTTTTTATTAATGAGATGATGTGTTTGTGACTGAGATTGAGTCATACATTTCACTAAGAAATGGTCCAA
GTACCAAACTATCATGACCCAGTTGCAACATGACGTTCCGGAGTGGTCACITTTGATAGTTCAATTTATCTTGGC
TCTTATTCCTTTTATAAATTTCTAATTTCTTCTGTGTAACCTATTTCATGTATTATTTTCTTTAAATTTACATGTCAT
TATTTTGCCTCACTAATCTAATTTTGCATATAACAATGATAAGTGTATTTTGAATCACAAAATTTACATCAAAATTC
GACATCGTTTATTTATGTTCAITTTGGATGATTAAACAAATATAACAAAGTTTGAACCTAATTAACCACCACTGAATATAA
TTAACTATAACTGTGAAAGTAGTTAACTCATTTTATATTTTATAGATCAATAAGAGAAATAACGGTATATTAATCC
CTCCAAAAAAGGATATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGG
ATAACATCCAATCCAACCAATCACAAACATCCTGATGAGATAACCCACTTTAAGCCACGATCTGTGGCAGATCT
ACATTATCTAATCACACATCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTAT
AAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAACACATACAAAGAGAGAGACTAATTAATTAATTA
ATCATCTTGAGAGAAATGTTTGGGCGCGGACCAACAAGGAAGTGATAACACCAAAATATTACGATATTTCTTGG
TGTTTCAAAAAGTGCTAGTGAAGATGAAATCAAGAAAGCCATATAGAAAGGCAGCCATGAAGAACCATCCAGATAA
GGGTGGGGATCTGAGAAAGTTCAAGGAGTTGGGCCAAGCATATGAAGTGTGAGCGATCTGAAAAGAAAGAAC
TGTATGATCAATATGGTGAAGATGCCCTTAAAGAAGGAATGGGGGAGGCGCAGGAAGCTCATTTCATAATCCGT
TTGATATTTTCAATCATTTTGGTGCAGGCTTTGGTGGTGGTGGTCCCTTACGCGCCAGAAAGACAGAAAGCAAG
GAGAAGATGTGGTGCATTCTATAAAGGTTTCCCTGGAGGATGTGTATAACGGCACTACAAAGAAGCTATCACITTC
TAGGAATGCAGTGTCTCAAAATGTAAAGGCAAGGTTCAAAAAGTGAAGTGGTGGAAAGGTTTGGATGCCA
GGGCACAGCTATGAAGATTACGAGAAGGCAAAATGGAGTGGGCATGATTCAACAAATGCAACAGCTGTGTCTGA
CTGCAAGGGAACAGGCGAGGTCATTAGTGAGAGAGATAGATGCCCTCAATGCAAGGGGAAACAAGATTACTCAAG
AAAAGAAGGTGCTGGAGGTGCATGTGAAAAAGGGGATCGAGCAGGGTCACAAGATTGATTCGAAGGACAAGCT
GATGAAGCTCCTGATACAATCACAGGAGACATAGTTTTGTCTTGAAGTAAAGGGACATCCGAAGTTTCCGAGG
GAGCGTATGACCTCCACATTGAACACAATTTGAGCTTAACTGAGGCTCTCTGTGGCTTCCAGTTAATGTGACA
CATCTTGATGGAAGGCAACTATTGGTCAAATCGAACCCTCGGCGAAGTCATCAAGCCAGGTCAACATAAGCTATA
AATGATGAGGGAATGCCACAACATGGTAGGCCGTTTCATGAAGGGAGGCTATACATCAAGTTTAGTGTGATTTC
CCGGATTCCGGTTTCTTTTCCCAAGGCAAGGCTTGAATTTAGAAAAGATATTACCTCAAAAGACAAGCAAGAAC
TTGTCCCAAAAGGAGGTAGATGATTGTGAGGAGACCACTTGCATGATGATCAATATTGAGAGGAGATGAGTGG
AAGAAGCAACAATACCCTGAGGCATATGATGAGGATGATGATGAAGATGATGAGCACTCCGAGCCTCGGGTGG
AATCGGCTCAACAGTAGGAGCTCAGCTCGAATTTCCCGGATCGTTCAAACTTTGCCAATAAAGTTTCTTAAGATT
GAATCCTGTGCGGGCTTGGGATGATTATCATATAATTTCTGTGAATACGTTAAGCATGAATAATTAACATGT
AATGATGACGCTTATTTATGAGATGGGTTTTATGATTAGAGTCCGCAATTATACATTTAATACGATAGAAAAAC
AAAATATAGCGCGCAAACTAGGATAAATATCGCGCGCGGTGTCATCTATGTACTAGATCGAATTC

Figura 42

SEQ ID NO: 78

Construção número R860 a partir de HindIII (no sítio de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). Sequência de codificação HSP70 está sublinhada

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCGGGCTGGTCTGTACATTCATCTTGCCGCCCTTGCAATTCACCTGGCCAC
AAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGCAAGTGCACGAGGGTCAGAACTGTATATCATA
TCTATGTGAGAGAAAGGGGAACATTTGAGATGGAGTCCATTTACCTTGAGGTATACCTTATTATTTTGATCAATAAATTTGTATACCTCTT
ATTTAGATCAATAAATTTGTCATTAAGCTATAATCCAAAATAAATTACGATCAAAATGCAAATGTTAGCCAGTACTTGTGTTAACTGTG
ATGGCATCTCTTGGTTTCTTTGGCAATCACATGCCATAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGAT
GACTGTGTTTGGATAAACAGCTTAATTAAGCGCTTATAGAAATATCATATGATTGTGTTTGGTCAGACTTCAGAGTCATCTTGGTTTCT
CTGGCAATCATATGCCATAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACCTGTTTGGGTAACA
GCTTAATTAAGTCTTATAGAAATAAGCGCTTATCATATAAGTGCTTTTGTACAGTTATTTCTATGAAAGTAGAAGAAATAGTCATATTG
TTTAATATAAGCTATCTGGAGAGCTTGTGGAAATAACAGAAAAGAAGCTTATGGACAGCTCATGAGCTGTTTACATAAGATCTCCC
TAACAGTCTCAAAAGTGTATGCGAGTAGATAAAATCAAATAAGTCAATCTAAACAGACCTTAAATCCATTATGGTACCTATCATTTT
AGCTTATTCATCTTTATTAAGAATGTCATGAGATAACATAAGATAACACATTATTTGACACAAATGGCCAGATCTAGCAATTTAAC
TCTGGAGTCCCTCAAGACTGCTGTTCTTACGAAGTTCACGTCCCTGAATCATGTTTCTGTATGGAAGCGTGAAGACCTCAAATCTA
AAAGGTGGCGATAAATGAAGGTTTACAAATATACCTGCGGGCTTGACACAGAGGCAAGCTCTTATACCTCCAGTTCAACGGG
GATGTTGATTTCAGAAGTCACTTGGAGAGCAATCCTTGTGCGAAGTTTGAAGTAATTTTGTGTAGCATATGTTGAGCTACCTACAAT
TTACATGATCACCTAGCATTAGCTCTTCACTTAAGTGAAGATGAAGTTTGAAGATGAGTATGACCATGGAGTCCGATGGCTT
GTAAATGCTACCTACTTTTGGCAACTCATCGGGATTACATTCAAGAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTGTA
AGATAACTGAATGTTTATTTAATGTTGGGTTGTAGTGTTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTG
GGTCTGTACTGTTATTTATTTATTTTATTTAGCAGAAACACCTTATCTTTGTTTCTGTTGAATGTAGAATGAAATAAAGAAAGAA
AATATAACATCATCGGCGCGCTTGTCTAATTCGGGCGAGTTAGGATCCTCTCGGTCACCGGAAAGTTTCAGTAGAAGAAACAAAA
AAATATAACATCATCGGCGCGCTTGTCTAATTCGGGCGAGTTAGGATCCTCTCGGTCACCGGAAAGTTTCAGTAGAAGAAACAAAA
GGGCGCTACAACCTGAAACAAAAAGAGATATTTCTCTCTTTCAGAAATGTAAGTTTCTTTACAGATACCCATTCCACATTTGATT
CAGATGTGGTGACTAGAGATAAAGCATACTAATTTGACTCTTGGAAACCCATAAAGTTTATGTTATCCGTGTTCTGGACCAATCCACT
TGGGGGCGATAACCTGTGTCTATGTGTGTTTGGTTTCCATTCTGATTATGCGGCCACTTGTAAATTTAAATCTAGGAGGGGCGAGAC
ATTGAACAATCCCAATTTTAAATAATGCAAGATTTTCTTATTAATGAGATGATGTTTGTGATGAGATGCTTCAACCTGATCATCATTT
CACCTAAGAAATGGTTCCAAGTACCAACTATCATGACCCAGTTGCAACATGACGTTCCGGGAGTGGTCACTTTGATAGTTCAATTTCT
ATCTTGGCTTCTTATTCCTTTTATAATTTCTAATTTCTTGTGTAAACTATTTTCATGATTATTTTCTTTAAATTTTACATGTCATTTATT
TTGCTCTACTAAGTCAATTTTGCATATAAATGATAAGTGATATTTTACTCACAAATTTACATCAATTTCCAGATCGTTTATTATG
TTGATTGGATGAATTAACAAATATAACAACTTTGCAACTAATTAACCAACCAAGTGAATATAATTAACATATACTGAAAGTGAATTAAT
TCATTTTATATTTTATAGATCAAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGGGTATATTTACTAAAAATCTA
AGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCCAATCCCAACATCACAACATCCTGATGAGATAACCCACTT
TAAGCCACGCATCTGTGGGCATCTACATTTATCTAAATCACACATTTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTT
TATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAACACATACAAAGAGAGAGACTAATTAATTAAT
TTAATCATCTTGAGAGAAATGTCGGGTAAGGAGAGAGGACGAGCTATCGGTATCGATCTTGGTACCCTTACTCTTGGCTCGGAGT
ATGGCAACAGGACCTGTTGAGATCATTTGCTAATGATCAAGGAACAGAACCCAGCCATCTTACGTTGCTTTACCGGACTCCGAGA
GGTTGATCGGTGACCGAGCTAAGAAATCAGGTCCCATGAACCCGTTTAAACCCGTTTTCGACCTAAGAGGTTGATCGGTGCTGCT
TTCTCTGAGAGCTGTTTTCAGAGTGACATGAATTTGTGCCATTCAGGATTCAGGCGGACCTGCGGATAAGCCAAATGCTACGTC
GAATACAAGGGTGAAGAGAAAGAGTTCCGAGCTGAGGAGATTTCTTCCATGGTTCTTATTAAAGATGGGTGAGATTGCTGAGGCTTA
CTTGGTGTCAATCAAGAAGCGCTTGTACCGGTTCCAGCTTACTTCAACGACTCTCAGCGTCAAGGCTACAAAGGATGCTGGTGT
CATCGCTGGTTTGAAGCTTATGGGAATCATCAACGAGCCTACAGCCCGGCTATTGCTACGGTCTTTCAGCAAAAGGCTACCCAGCG
TTGGAGAGAAGAATGTTCTTATCTTCGATCTTGGTGGTGGCATTGATGCTCTCTTCTTACCATTGAAGAGGGTATCTTTGAGGT
GAAGGCAACTGCTGGTGACACCCATCTTGGTGGGGAAGATTTTGCACACAGAAATGGTTAACCACTTTGTCCAAGAGTTCAAGAGGA
AGAGTAAGAAGGATATGACCGGTAACCCAGAGCTCTTAGGAGGTTGAGAATCTGCTGTGAGAGAGCGAAGAGGACTCTTCTTCC
ACTGCTCAGACCCATCGAGATTGACTCTCTATACGAGGGTATCGACTTCTACTCCACCATCAGCGGTGCTAGATTGAGGAGCTC
AACATGGATCTCTCAGGAAGTGTATGGAGCCAGTTGAGAAGTGTCTTGGTATGCTAAGATGGACAAGAGGACTGTTCTATGATGTT
GTCTTGTGTTGGTGGTTCTACCCGATCCCTAAGGTTCAAGCAATTTGATCCAGGACTTCTTCAACGGCAAGAGCTTTGCAAGTCTATT
AACGCTGATGAGGCTGTTGCTACGGTGTCTGCTGTCCAGGGAGCTATTCTCAGCGGTGAAGGAACGAGAAGGTTCAAGATCTTCT
ATTGCTCGATGTACT
AACCAAGGAAGGAACAGTCTTCTCCACCTACTCAGACAACCAACCGGTTGTTGATCCAGGTTGACGAAGGAGAGAGAGGCGAGAA
CCAGGACAACAACCTTCTTGGTAAATTTGAGCTCTCCGGAATCTCCAGGCTCTGCTGGTGTCTCCCGGAGATCAGAGTCTGCTTGTG
ACATTGATGCCAATGGTATCTCTCAATGTCTCTGCTGAGGACAAGACCCGAGACAGAAGAACAGATCAACATCAGCAATGACAAG
GCTCGTCTCTCCAAGGATGAGATTGAGAAGATGGTTCAAGAGGCTGAAGAGTACAAGTCCGAAGAGAGGAGGACACAAGAGAAGG
TTGAAGCCAAGAACGCTCTCGAGAATACGCTTACAACATGAGGAACACCAACCAAGACGAGAAGATTGGTGAGAAGCTCCCGCT
GCAGACAAGAGAAGATCGAGGATTCTATTGAGCAGGGGATTCATGCTCCAGGCTGAGGAGGTAACCAAGTTGAGGCTGAGGCTGATGAGTCC
AAGACAAGATGAAGGAATTTGGAGAGCATCTGCAACCAATCATTTGCCAAGATGTACCAAGGAGCTGGTGGTGAAGCGGCTGGTCC
AGGTGCTGTGGTATGGACGATGATGCTCCCGCTGCTTCAAGCGGCTGCTGACCTAAGATCGAGGAGGTCGACTAAGAGCTCAGC
TCCAATTTCCCGATCGTTCAACATTTGGCAATAAGTTTCTTAAGATGAATCTGTTGCCGGTCTTGGATGATTATCATATAATT
TCTGTTGAATTACGTTAATAATTAACATGAATGCAATGACGTTATTTATGAGATGGGTTTTATGATTAGGATGAGGATGAGGATGAGG
TATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCG
AATTG

Figura 43

SEQ ID NO: 79

Construção número R870 a partir de HindIII (no sítio de clonagem múltipla, a montante do promotor) para EcoRI (imediatamente a jusante do terminador NOS). Sequência de codificação HSP40 está sublinhada em *itálico* e a sequência de codificação HSP70 está sublinhada

AAGCTTGCATGCTGCAGGTCGACTCTAGAGGATCCCGGGCTGGTCTGTACATTTCATCTTGCCTGCTTGCATTCACTTGGC
CACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGCAGTGCACTCGAGGGTCAGAACTGT
ATATCATATCTATGTGAGAGAAAGGGGAACATTTGAGATGGAGTCCATTTACTTGAGGTATACCTTATTTTGTATCAATAAAT
TGATACTTCTTATTTAGATCAATAAATTTGTCATTAAAGCTATAATCCAAAATAAATTACGATCAAAATATGCAAAATGTTAGCCAGT
ACTTGTGTTAACTTGATGGCATCTCTTGGTTCTTTGGCAATCACATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGG
TCAGATTCAGAGTCAGATGACTCTGTTTGGATAAACAGCTTAATTAAGCGCTTATAGAATATCATATGATTGTGTTTGGTCAG
ACTTCAGAGCATCTCTTGGTTCTCTGGCAATCATATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTCAG
AGTCAGATGACCTCTGTTTGGGTAACAGCTTAATTAAGTCTTATAGAATAAGCGCTTATCATATAAGTGCCTTTGTACAGTTAT
TTGTATGAAAGTAGAAGAAATAGTATATTTGTTTAAATATAAGCTATCCTGGAGAGCTTGTGGAATAACAGAAAGAACTTAT
GGACAGGTCATGAGCTGTTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTATTATGCCAGTAGATAAATCAAATAAGTCAAT
CTAAACAGAGCCCTAAATCCATTATGGTACCTATCATTTTAGCTTATCCATCTTTATTAAGAATGTCATGAGATAACATAATGATA
ACACATTATTTTACACAAATGGGCAGATCTAGCAATTTAACTCTGGAGTCTTCAAGACTGCTGTTCTTACGAAGTTCAGCTC
CCTGAATCATGTTCCGTGTATGGAAGCCCTGAAAGAGCTCAAAATCTAAAAGGTGGCGATAAATGAAGGTTTACAAATATACCC
TGCGGGCTTGACACAGAGGCAAGCTCTTTATACCTTCCAGTTCAACGGGGATGTTGATTTCAGAAGTCACTTGGAGAGCAATC
CTTGTGCCAAGTTGAAGTAATTTTGTGTAGCATATGTTGAGCTACCTACAATTTACATGATCAGCTAGCATTAGCTCTTTCAC
TTAAGTGAAGAAATGAAGTTTATGGAATGAGTATGACCATGGAGTGGCATGGCTTTGTAATGCCCTACCTACTTTGGCCAC
TCATCGGGGATTTACATTCAGAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTAAAGATAACTGAATGTTTCATTTTA
ATGTTGGGTTTGTAGTGTGTTTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTGGGCTGTGACTGTTATTTAT
TTATTTTTTTTATAGCAGAAACACCTTATCTTTGTTTCTGTTGAATGTAGAATGAAAATAAAGAAAGAAATATAACATCATCG
GCGCGCTTGTCTAAATTCGGGCAAGTAGGATCCTCTCGGCTACCCGAAAGTTTCAGTAGAAGAAACAAACACCGGTGACTA
AAATGACACTATTATTTATTTATTTGTTTCTTTTCTACCGGAACCTTTTGAACCGGATCCCACTGTTCCGGGGCGCT
ACAAGTGAACAAAGAGATATTTCTCTCTCTCAGAAATGTAAGTTTCTTTACAGATACCCATTCAACATTGATTGAGAT
GTGGTGACTAGAGATAAAGCATACATAATTTGACTCTTGGAAACCCATAAAGTTTATGTTATCCGTGTTCTGGACCAATCCACTT
GGGGGCAAAACCTGTGTCTATGTGTGTTGGTTTCCATTCTGATTATGCGGGCACTTGAATTTAAATCTAGGAGGGGCA
GACATTGAACAAATCCCAATAATTTAATAACTTATGCAAGATTTTATTTAATGAGATGATGTTTGTGACTGAGATTGAGTCA
TACATTTCACTAAGAAATGGTTCCAAGTACCAACTATCATGACCCAGTTGCAAAACATGACGTTCCGGAGTGGTCACTTTGATA
GATTCATTTTCATCTTGGCTTCTTATCTTTTATAATCTAATTTCTTGTGTAAACTATTTTCATGTATTATTTTCTTAAATTT
ACATGTCATTTATTTTGGCTCACTAACTCAATTTTGCATATAACAATGATAAGTGATTTTGAAGTCAACAATAATTTGAGTCAACA
CGACATCGTTTATTATGTTTCAATTGGATGATTAACAAATATAACAACTTTGCAACTAATTAACCACTGAATATAATTAACCTA
TAAGTGTGAAAGTAGTTAATCTATTTTATATTTTCATAGATCAAAATAGAGAAATAACGGTATATTAATCCCTCCAAAAA
ACGGTATATTTACTAAAAATCTAAGCCAGCTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATCCAAATCCAAACATCA
CAACAACTCCTGATGAGATAACCCACTTTAAGCCCAAGCATGTTGGCCACATCTACATTATCTAAATCACACATTTCTCCACACA
TCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCTATCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTT
CAACACATACATAAGAGAGAGACTAATTAATTAATCATCTTGAGAGAAATGTGGGGTAAAGGAGAGGAGCCAGCTATC
GGATCCGATCTGGTACCCTTACTCTTGGCTGGAGTATGGCAACGACCGCTGTTGAGATCATTGCTAATGAAGGAA
GAGAACCCAGCCCATCTTACGTTGCTTTCACCGAGCTCCGAGAGGTTGATGCGGTGACCGCAGCTAAGAAATCAGGTCGCCATGAAC
CCGTTTAAACACGCTTTTCGACGCTAAGAGGTTGATCGGTGCTGCTTCTCTGACAGCTCTGTTTCAGAGTGAACATGAATTTGTG
GCCATTCAGATTCAAGCTCAAGCCGACCTGCCGATAAGCCAAATGATCTACGTCGAATACAAGGGTGAAGAGAAAGAGTTCCGACCT
GAGGAGATTCTTCCATGGTCTTATTAAGATGGCTGAGATTGCTGAGGCTTACCCTTGGTGTCACAATCAAGAAGCCGCTTGT
ACCGTTCCAGCTTACTTCAACGACTCTCAGGTCAGGCTACAAAGGATGCTGCTGTCATCGCTGGTTTGAAGCTTATCGGAAT
CATCAAGAGGCTACAGCCCGCGCTATTGCTTACGGTCTGACAAAAAGGCTACCAGGCTTGGAGAGAAAGATGTTCTTATCT
TCGATCTTGGTGGTGGGCAAGATTTTGAACAACGAATGTTTAAACCACTTTGTCGAAGAGTTCAAGAGGAGAGTAAAGAGGATAT
ACCCATCTTGGTGGGGAAGATTTTGAACAACGAATGTTTAAACCACTTTGTCGAAGAGTTCAAGAGGAGAGTAAAGAGGATAT
CACCGGTAACCCAAAGAGCTCTTAGGAGGTTGAGAACCTCTGTTGAGAGAGGCAAGAGGAGCTTCTTCTTCCACTGCTCAGACC
ACCATCGAGATTGACTCTCTATACGAGGATCGACTCTTACTCCACCATCACCGGCTCTAGATTGAGGAGCTCAACATGGA
TCTCTTCAAGAGTGTATGGAGCCAGTTGAGAACTGCTTCTGATGCTAAGATGGACAAGAGCACTGTTTCATGATGTTGTCT
TTGTTGGTGGTCTACCGGATCCCTAAGGTTCAAGCAATGCTCCAGGACTTCTTCAACGGCAAGAGCTTTGCAAGTCTATTA
ACCTGTATGAGGCTGTTTCCCTACCGGCTGCTGCTGTCAGGAGGCTATTCTCAGCGGTGAAGGAAACGAGAGGTTCAAGATCT
TCTATGCTCGATGTCACCTCTCTCTCTGTTGGTTTGAAGCTCCGGTGGTGTGATGACCACTTTGATCCGAAGGAACACAA
CCATCCCAACCAAGAAAGGAACCAAGTCTTCTCCACCTACTCAGACAACCAACCCGGTGTGTTGATCCAGGCTGATCAAGAGGA
GAGAGCCAGAACCAAGGACAACACCTTCTTGGTAAATTTGAGCTCTCCGGAATTCCTCCAGCTCTCTGTTGGTGTCCCGCAG
ATCACAGTCTGCTTTGACATTGATGCCAATGGTATCTCAATGCTCTGCTGAGGACAAGACACCGGACAGAAAGAACAGAT
CAACATCAACCAATGACAAGGGTCTGCTCTCCAAGGATGAGATTGAGAAGATGGTTCAAGAGGCTGAGAAGTACAAGTCCGAA
GACGAGGAGCACAAGAAAGAGGTTGAAGCCAAGAGCTCTCGAGAACTACGCTTACAACATGAGGAACACCATCCAGAGC
AGAAGATTGGTGAGAAGCTCCCGGCTGACAGACAAGAAAGATCGAGGATTCTATTGAGGAGGCGATTCAATGGCTCGAGG
GTAACAGTTGGCTGAGGCTGATGAGTTGGAAGAGATGAGTAAGGAATTTGAGAGCATCTGCAACCAATCATGCAAGAT
GTACCAAGGAGCTGGTGGTGAAGCCGGTCCAGCTGCTCTGCTATGAGGATGATGCTCCCTCTCTCAGGCGGCTGCTCAGGCGGCTG
TGGACCTAAGATCGAGGAGGTCGACTAAGAGCTCAGCTCGAATTTCCCGATGCTTCAACATTTGGCAATAAAGTTCTTAA
GATTGAATCCTGTTGCCGCTTGGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG

Figura 44a

CATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCA
AACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG
TTATCGCGGGCTGGTCTGTACATTCATCTTCCCGCTTTGCATTCACTTGGCCACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCCA
GACTTCAAGAAGCGACCTTGCAAGTGCACCTCGAGGGTCAGAAACTGTATATCATATCTATGTGAGAGAAAGGGAACATTTGAGATG
GAGTCCATTTACTTGAGGTATACCTATTATTTTGTATCAATAAATTTGTATACCTCTTATTTAGATCAATAAATTTGTCAATTAAGCTATAAT
CCAAAATAAATTACGATCAAAATATGCAAAATGTTAGCCAGTACTTGTGTAAACTTGATGGCATCTCTGGTTCCTTTGGCAATCACATG
CCTAAGAAAATAAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGATAAACAGCTTAATTAAGCGC
TTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGCAATCATATGCCATAAGAAATAAATAGTAT
CATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGGTAACAGCTTAATTAAGTGTATAGAATAAGCGCTTAT
CATATAAGTGTCTTTGTACAGTTATTTCTATGAAAGTAGAAGAAATAGTCATATTGTTTAAATAAAGCTATCCTGGAGAGCTTGTGGA
AATAACCAGAAAAGAACTTATGGACAGCTCATGAGCTGTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTATGCCAGTAGATA
AATTCACAACTATCAATCTAAACAGAGCCCTAAATCCATTATGGTACCTATCATTTTATAGCTTATTCATCTTTATTAAGAAATGATGAGA
TAACATAATGATAACACATTATTTTACACAAATGGGCAGATCTAGCAATTTAACTCTGGAGTCTTCAAGACTGCTGTTCTTACGAA
GTTTACGTCCTCTGAATCATGTTCTGTATGGAAAGCCTGAAAGACCTCAAAATCTAAAGGTGGCGATAAATTGAAGGTTTACAAAATA
TACCTCGCGGGCTTGACACAGAGGCAAGCTCTTATACCTTCCAGTTCACCGGGGATGTTGATTTCAGAAGTCACCTTGGAGAGCAAT
CCTTGTGCCAATGATCAATGTTTGTGTAGCATATGTTGATGACCAATCACTTTGAGTATCATGACCTTACCTGATGACCTTACCTTACCT
ACTGAGAGAATGAAGTTTATGGAATGAGTATGACCATGGAGTCGGCATGGCTTTGTAATGCCTACCTACTTTGGCCAACTCATCGG
GGATTACATTCAGAAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTAAAGATAACTGAATGTTTCAATTTAATGTTGGGTTG
TAGTGTTTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTGGGTCTGACTGTTATTTATTTATTTTTTTTATG
AGAAACACCTTATCTTTGTTTGTGTTGAATGTAGAATAAAGAAAGAAATAAATACATCATCGGCCCGCTTGTCTAATTTT
CGGCAGTTAGGATCCTCTCCGGTCACCGGAAAGTTTCACTAGAGAAGAAACAAACACCGTGACTAAAATGATACTATTATTTATTTAT
TGTGTTTTCTTTTTGTACCGGAACCTTTTGAACGGATCCCAACTCGTTCGGGGGGCGCTACAACCTGAAACAAAAGAGATATTT
CTCTCTCTTCAAGAAATGTAAGTTTCTCTTACAGATACCCATTCACCATTTGATTGAGATGTGGTGACTAGAGATAAAGCATACTAAT
TGACTCTTGGAAACCCATAAAGTTTATGTTATCCGTGTTCTGGACCAATCACTTTGGGGGCATAACCTGTGTCTATGTGTGGTTGGT
TTCCATTCTGATTTATGCGGCGACTTGTAAATTTAAATCTAGGAGGGGCAGACATTGAACAATCCCAATATTTAATACTTATGCAAG
ATTTTTTTTATTAATGAGATGATGTGTTTGTGACTGAGATTGAGTCATACATTTCACTAAGAAATGGTCCAAAGTACCAAACTATCATG
ACCCAGTTGCCAACATGACGTTCCGGAGTGGTCACCTTGTATGTTCAATTCATCTTGGCTTCTTATTCCTTTTATAATTTCTAATTTCTT
CTTGTGTAAACTATTTCATGTATTATTTTCTTTAAATTTTACATGTCTATTTTGGCTCACTAAGTCAATTTTGCATATAACAATGAT
AAGTGATATTTTGACTCACAAAATTTACATCAATTTGACATCGTTTATTATGTTTATTGGATGATTAACAAATATAACAACTTTGCA
ACTAATTAACCACCACTGAATATAATTAACATAACTGTGAAAGTAGTTAACTCATTTTATATTTTATAGATCAAAATAAGAGAAATAA
CGGTATATTAATCCCTCCAAAAAAGAAACCGGTATATTTACTAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGGTAGGA
GGATAACATCCCAATCCAACTCACAAACATCCCTGATGAGATAACCCCACTTTAAGCCACGCTATGTGGCAGATCTACATTATCTA
AATCACACATTTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTC
TACACTTTGATTCCCTTCAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAAATGTTTGGGGCGCGGACC
AACAAGGAAGAGTGATAACACCAATATTACGATATTCTTGGTGTTCAAAAAGTGCTAGTGAAGATGAATGAAGAAAGCCATAGA
AAGGCAGGATGAAGAACCATCCAGATAAGGGTGGGATCTGAGAAGTTCAAGGAGTTGGGCCAAGCATATGAAGTGTGTAGGG
ATCTTGAAAAAGAAAGAACTGTATGATCAATATGGTGAAGATGSCCTTAAAGAAAGGAATGGGGGGAGGGCGAGGAAGCTCATTTTCAT
AATCCGTTTGATATTTTGAATCATTTTGGTGCAGGCTTTGGTGGTGGTGGTCCCTTACGCGCAAGAGACAGAAAGCAAGGAGAA
GATGTGGTGCATTTCTATAAAGGTTTCTTGGAGGATGTGTATAACGGCACTACAAAGAAGCTATCACTTTCTAGGAATGCACTGTGC
TCAAAATGTAAAGGGAAGGTTCAAAAAGTGAACCTGCTGGAAGGTGTTTGGATGCCAGGGCACAGGTATGAAGATTACCAGAAG
GCAAAATGGACTGGGCATGATTCAACAAATGCAACAGCTCTGTCTGACTGCAAAAGGAACAGGGCGAGGTGATTAGTGAGAGAGATA
GATGCCCTCAATGCAAGGGAACAAAGATTACTCAAGAAAGAAAGGTGCTGGAGGTGCATGTGGAAGAAAGGGATGCAGCAGGGTCA
CAAGATTGTATTGCAAGGACAAGCTGATGAAGCTCCTGTATACAATCACAGGAGACATAGTTTTGTCTTGCAAGTAAAGGGACATCC
GAAGTTTCCGAGGGAGCGTGATGACCTCCACATTGAACACAAATTTGAGCTTAACTGAGGGCTCTCTGTGGCTTCCAGTTTAAATGTAC
ACATCTTGATGGAAGGCAACTATTGGTCAAAATCGAACCCCGGCAAGTCATCAAGCCAGGTCAACATAAAGCTATAAATGATGAGGG
AATGCCACAACATGGTAGGCCGTTTATGAAGGAGCGCTATACATCAAGTTAGTGTGATTTCGCCGATTCCGGTTTTCTTCCCG
AAGCCAAAGCCTGGAATTAGAAAAGATATTACCTCAAAAGACAAGCAAGAAGTTGTCCCAAAAGGAGGTAGATGATTGTGAGGAGAC
CACCTGTCATGATGTCAATATTGCAGAGGAGATGAGTCGAAAGAAAGCAACATACCGTGAGGCATATGATGACGATGATGAAG
ATGATGAGCACTCGCAGCCTCGGGTGCAATGCGCTCAACAGTAGGAGCTCAGCTCGAATTTCCCGATCGTTCAACATTTGGCAA
TAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAAC
ATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATA
GCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGAATTC

Figura 44b

Padrão de ponte dissulfido de tipo A HAs

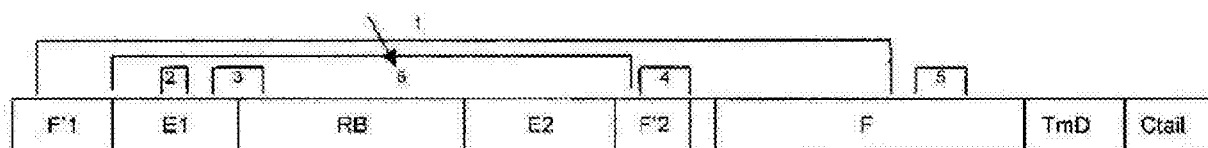


Figura 46

Padrão de ponte dissulfido de tipo B HAs

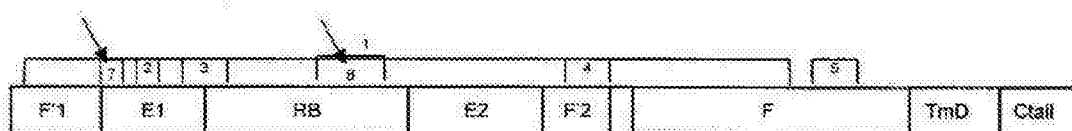
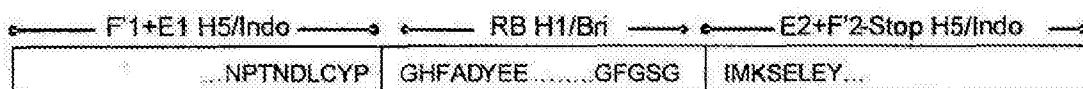


Figura 47

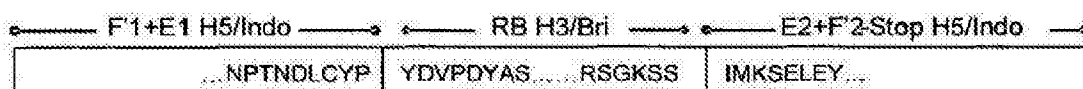
Figura 48

A

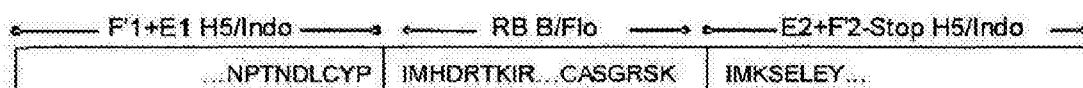
RB of H1/Brisbane on H5/Indo stem



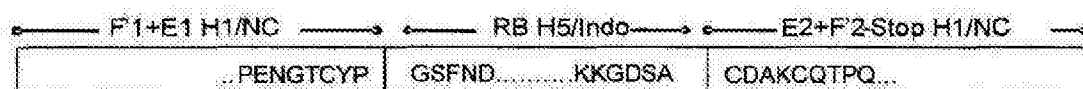
RB of H3/Brisbane on H5/Indo stem



RB of B/Florida on H5/Indo stem

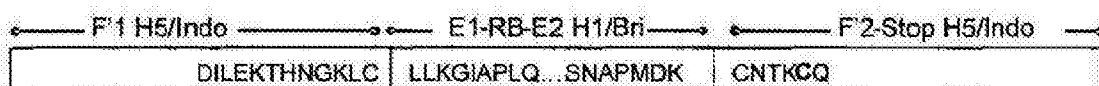


RB H5/Indo on H1/New Cal stem

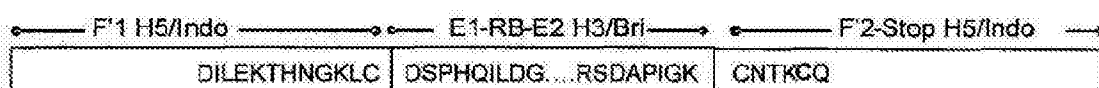


B

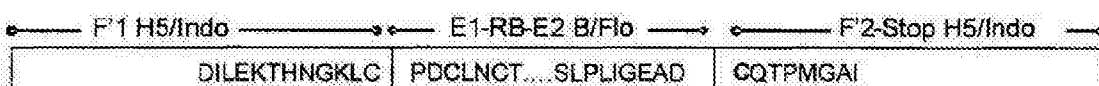
E1-RB-E2 of H1/Brisbane on H5/Indo stem



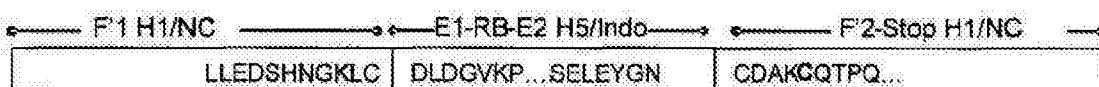
E1-RB-E2 of H3/Brisbane on H5/Indo stem



E1-RB-E2 of B/Florida on H5/Indo stem



E1-RB-E2 H5/Indo on H1/New Cal stem



SEQ ID NO:86

PDI sp - A/California/04/09

ATGGCGAAAAACGTTGCCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCCT
CTCAGATCTTCGCTGACACATTATGTATAGGTTATCATGCGAACAATTCAACAGA
CACTGTAGACACAGTACTAGAAAAGAATGTAACAGTAACACACTCTGTAAACCTT
CTAGAAGACAAGCATAACGGGAAACTATGCAAACTAAGAGGGGTAGCCCCATTG
CATTTGGGTAAATGTAACATTGCTGGCTGGATCCTGGGAAATCCAGAGTGTGAAT
CACTCTCCACAGCAAGCTCATGGTCCTACATTGTGGAAACACCTAGTTCAGACAA
TGGAACGTGTTACCCAGGAGATTTTCATCGATTATGAGGAGCTAAGAGAGCAATT
AAGCTCAGTGTTCATCATTTGAAAGGTTTGAGATATTCCCCAAGACAAGTTCATGG
CCCAATCATGACTCGAACAAGGTGTAACGGCAGCATGTCCTCATGCTGGAGCA
AAAAGCTTCTACAAAAATTTAATATGGCTAGTTAAAAAAGGAAATTCATACCCAA
AGCTCAGCAAATCCTACATTAATGATAAAGGGAAAGAAGTCCTCGTGCTATGGG
GCATTACCATCCATCTACTAGTGCTGACCAACAAAGTCTCTATCAGAATGCAGA
TACATATGTTTTTGTGGGGTCATCAAGATACAGCAAGAAGTTCAAGCCGGAAATA
GCAATAAGACCCAAAGTGAGGGATCAAGAAGGAGAATGAACATTACTGGAC
ACTAGTAGAGCCGGGAGACAAAATAACATTCGAAGCAACTGGAAATCTAGTGGT
ACCGAGATATGCATTTCGCAATGGAAAGAAATGCTGGATCTGGTATTATCATTTCA
GATACACCAGTCCACGATTGCAATACAACCTGTCAAACACCCAAGGGTGCTATA
AACACCAGCCTCCCATTTTCAGAATATACATCCGATCACAATTGGAAAATGTCCAA
AATATGTAAAAAGCACAAAATTGAGACTGGCCACAGGATTGAGGAATATCCCGT
CTATTCAATCTAGAGGACTATTTGGGGCCATTGCCGGTTTCATTGAAGGGGGGTG
GACAGGGGATGGTAGATGGATGGTACGGTTATCACCATCAAAATGAGCAGGGGTC
AGGATATGCAGCCGACCTGAAGAGCACACAGAATGCCATTGACGAGATTACTAA
CAAAGTAAATTCTGTTATTGAAAAGATGAATACACAGTTCACAGCAGTAGGTAAA
GAGTTCAACCACCTGGAAAAAAGAATAGAGAATTTAAATAAAAAAAGTTGATGAT
GGTTTCCTGGACATTTGGACTTACAATGCCGAACCTGTTGGTTCTATTGGAAAATG
AAAGAACTTTGGACTACCAGGATTCAAATGTGAAGAACTTATAGAAAAGGTAAAG
AAGCCAGCTAAAAACAATGCCAAGGAAATTGGAAACGGCTGCTTTGAATTTTA
CCACAAATGCGATAACACGTGCATGGAAAGTGTCAAAAATGGGACTTATGACTA
CCCAAAATACTCAGAGGAAGCAAAATTAACAGAGAAGAAATAGATGGGGTAAA
GCTGGAATCAACAAGGATTTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGT
TCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATG
GGTCTCTACAGTGTAGAATATGTATT

Figura 49A

SEQ ID NO:87

PDI sp - A/California/04/09

MAKNVAIFGLLFSLVLVPSQIFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRG
VAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEI
FPKTSSWPNHDSNKGVTAACPHAGAKSFYKNLIWLVKKGNSYPKLSKSYINDKGKEVLVLWGIHHPST
SADQQSLYQNADTYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVVPR
YAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINSTLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSI
QSRGLFGAIAAGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNKVNSVIEKMNTQFT
AVGKEFNHLEKRIENLNKKVDDGFLDIWYNAELLVLENERTLDYHDSNVKNLYEKVRSQKNNAKEI
GNGCFEFYHKDNTCMESVKNGTYDPKYSEEAKLNREEIDGVKLESTRIYQILAIYSTVASSLVLVSL
GAISFWMCSNGLQCRICI

Figura 49B

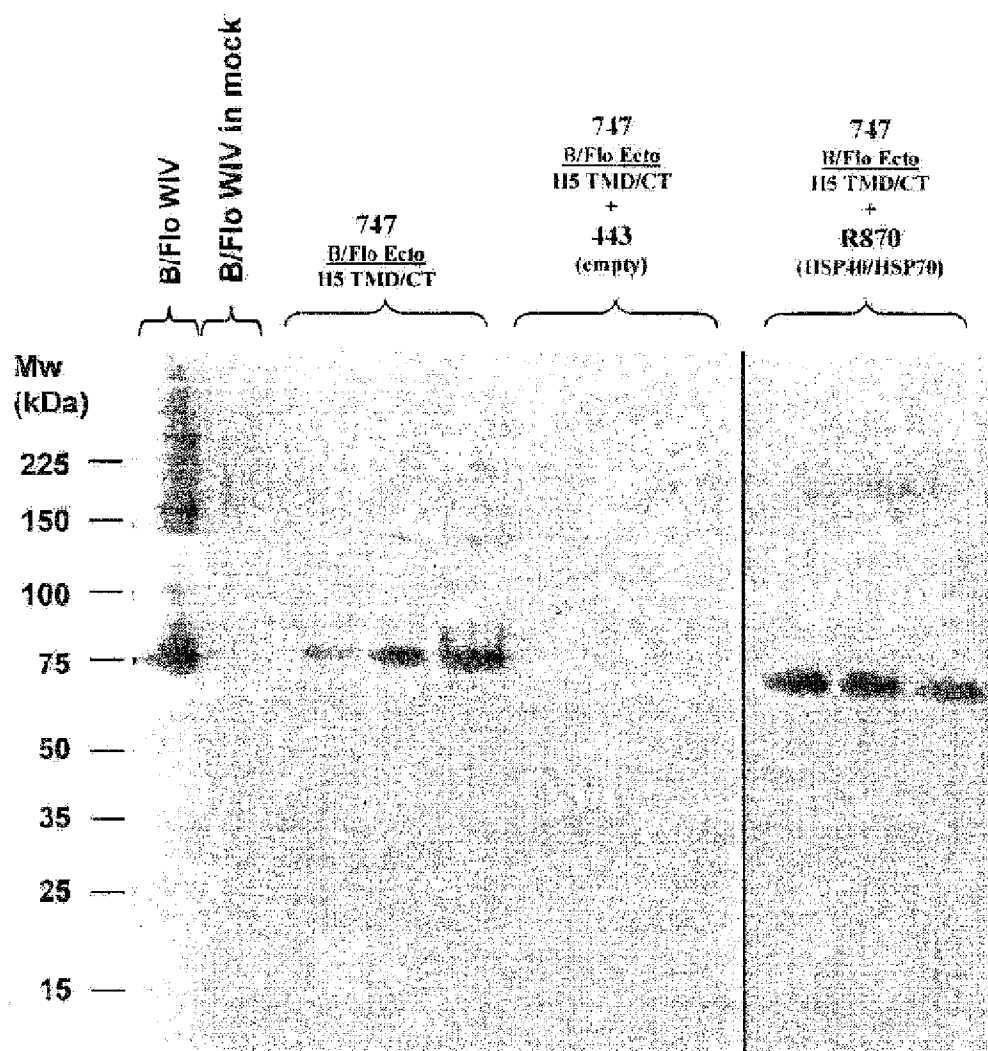


Figura 50

SEQ ID NO: 88

Sequência promotora 2X35S

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAA
GATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGG
GTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTG
GTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAA
GACGTTCCAACACGTCTTCAAAGCAAGTGGATTGATGTGATAACATG
GTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCT
CAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCG
GAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAA
GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAA
AGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAG
ATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATAT
AAGGAAGTTCATTTCAATTGGAGAGG

Figura 51A

SEQ ID NO: 93

Construção 747 a partir de Pac1 (a montante do promotor 35S) para Ascl (imediatamente a jusante do terminador NOS). Sequência de codificação de HA quimérico está sublinhada. A sequência promotora 2X35S está em *itálico*

TTAATTAAGTCGACAAGCTTGCATGCCTGCAGGTCAACATGGTGGAGCACGACACACTTGTCTACTC
CAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATC
CGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAG
GTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGGCCATCGTTGAAGATGCCTCTGCCGACAG
TGGTCCCAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCT
TCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAA
AGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCC
TCCGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTAC
AAATGCCATCATTGCGATAAAGGAAAGGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAG
ATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGT
GGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTT
CCTCTATATAAGGAAGTTCAATTTCAATTTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGCGAAG
GTGGGGAAACCCGAACCAAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTG
TCTTTCTTGCCTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTC
ACTGAAGCGAAATCAAAGATCTTTTTGTGGACACGTAGTGCGGGCGCCATTAAATAACGTGTACTTGTG
CTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTAGCCCCATACATTA
CTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGC
CTGTACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGT
TTTCCCGTGGTTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGGGC
CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTTCTTCTCAGATCT
TCGCTGATCGAATCTGCACCTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAA
GGGAGGTCATGTGACTGGTGTGATACCACTAACAAACACCAACCAAAATCTTATTTTGCAAATCT
CAAAGGAACAAGGACCAGAGGGGAAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCT
TTGGGCAGACCAATGTGTGTGGGGACCAACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAAC
CTGTTACATCCGGGTGCTTTCCATATAATGCACGACAGAACAAAAATCAGGGCAACTACCCAATCTTCTC
AGAGGATATGAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGGCACCAGGAGGAC
CCTACAGACTTGGAACTCAGGATCTTGCCTAACGCTACCGTAAGAGGGATTTTCGCAACAATCT
GCTTGGGCTGTCCCAAGGACAAACAACAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACT
TTGTACAGAAGGGGAAGACCAAAATCACTGTTTTGGGGGTTCATTTCAGATAACAAAACCCAAATGAAGA
ACCTCTATGGAGACTCAAATCCCTCAAAAGTTACCTTCATCTGTGTAATGGAGTAACCCACACACTATGTTT
CTCAGATTGGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGGCAGGATTGTTGT
TGATTACATGATGCAAAAACCTGGGAAAAACAGGAACAATTGTCTACCAAGAGGGTGTGTTGTTGCCTC
AAAAGGTGTGGTGGCGAGTGGCAGGAGCAAAAGTAATAAAAGGGTCTTGCCCTTAAATGGTGAAGC
AGATTGCCCTCATGAAAAATACGGTGGATTAAACAAGCAAGCCCTTACTACACAGGAGAACATGCAA
AAGCCATAGGAAATTCGCAATATGGGTGAAAACACCTTTGAAGCTCGGCAATGGAACCAATAIAGA
CCTCCTGCAAAACTATTAAGGAAAGGGGTTCCTTCGGAGCTATTGCTGGTTTCCCTAGAAGGAGGATG
GGAAGGAATGATTGCAGGCTGGCAGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGGC
GGACCTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCCTTGAGTGAGGTAGA
AGTAAAGAATCTTCAAAGACTAAGTGGTGGCATGGATGAACCTCCACAACGAAATACTCGAGCTGGATG
AGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAAIAGAAGCTTGCAGTCTTGCTTTCCAAC
GAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACATAAAGAAAATGCTGGG
TCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAAACACAAGTGCAACCAGACCTGCTTAG
ACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAAATTTCTCTCCCCACTTTTGAATCACTGAACATTA
CTGCTGCATCTTAAATGATGATGGATTGGATAACTACCAATACTGTCAATTTATCAACAGTGGCGA
GTTCCTTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAA
TGCAGAAATTCATTTAAAGGCCATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATG
TTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGT
TTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAGA
CCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGA
TTGAATCCTGTTGCCGGTCTTGCAGTATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAA
TTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTA
ATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGT
TACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Figura 51B