(11) Número de Publicação: **PT 1578903 E**

(51) Classificação Internacional:
C12N 15/00 (2011.01) ***C12P 21/06*** (2011.01)
C12N 15/09 (2011.01) ***C12N 15/63*** (2011.01)
C12N 15/70 (2011.01) ***C12N 15/74*** (2011.01)
C12P 21/04 (2011.01) ***C12Q 3/00*** (2011.01)
C12Q 1/70 (2011.01) ***G01N 33/53*** (2011.01)

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

(22) Data de pedido: **2002.04.17**

(30) Prioridade(s): **2001.04.17 US 837306**
2001.10.24 US 516
2001.10.25 US 45674

(43) Data de publicação do pedido: **2005.09.28**

(45) Data e BPI da concessão: **2011.11.23**
009/2012

(73) Titular(es):
DYAX CORP.
300 TECHNOLOGY SQUARE, 8TH FLOOR
CAMBRIDGE, MA 02139 **US**

(72) Inventor(es):
EDWARD H. COHEN **US**
KRISTIN L. ROOKEY **US**
RENE HOET **NL**
HENDRICUS R.J.M. HOOGENBOOM **NL**
ROBERT C. LADNER **US**

(74) Mandatário:
PEDRO DA SILVA ALVES MOREIRA
RUA DO PATROCÍNIO, N.º 94 1399-019 LISBOA **PT**

(54) Epígrafe: **NOVOS MÉTODOS DE CONSTRUÇÃO DE BIBLIOTECAS COMPREENDENDO MEMBROS APRESENTADOS E/OU EXPRESSOS DE UMA FAMÍLIA DIVERSA DE PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS E NOVAS BIBLIOTECAS**

(57) Resumo:
MÉTODOS ÚTEIS NA CONSTRUÇÃO DE BIBLIOTECAS QUE APRESENTAM E/OU EXPRESSAM COLECTIVAMENTE MEMBROS DE FAMÍLIAS DIVERSAS DE PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS E AS BIBLIOTECAS PRODUZIDAS UTILIZANDO ESSES MÉTODOS. MÉTODOS DE RASTREIO DESSAS BIBLIOTECAS E PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS IDENTIFICADOS POR ESSES RASTREIOS.

RESUMO

"NOVOS MÉTODOS DE CONSTRUÇÃO DE BIBLIOTECAS COMPREENDENDO MEMBROS APRESENTADOS E/OU EXPRESSOS DE UMA FAMÍLIA DIVERSA DE PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS E NOVAS BIBLIOTECAS"

Métodos úteis na construção de bibliotecas que apresentam e/ou expressam colectivamente membros de famílias diversas de péptidos, polipéptidos ou proteínas e as bibliotecas produzidas utilizando esses métodos. Métodos de rastreio dessas bibliotecas e péptidos, polipéptidos ou proteínas identificados por esses rastreios.

DESCRIÇÃO

"NOVOS MÉTODOS DE CONSTRUÇÃO DE BIBLIOTECAS COMPREENDENDO MEMBROS APRESENTADOS E/OU EXPRESSOS DE UMA FAMÍLIA DIVERSA DE PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS E NOVAS BIBLIOTECAS"

Este pedido é, em parte, uma continuação do pedido provisório dos Estados Unidos 60/198069, apresentado a 17 de Abril de 2000, em parte, uma continuação do pedido de patente dos Estados Unidos 09/837306, apresentado a 17 de Abril de 2001, em parte, uma continuação do pedido PCT PCT/US01/12454, apresentado a 17 de Abril de 2001, em parte, uma continuação do pedido dos Estados Unidos 10/000516, apresentado a 24 de Outubro de 2001 e, em parte, uma continuação do pedido dos Estados Unidos 10/045674, apresentado a 25 de Outubro de 2001.

A presente invenção refere-se às formas de realização caracterizadas nas reivindicações.

Deste modo, esta refere-se a um método para produzir uma população ou biblioteca de genes de imunoglobina que compreende os passos de:

(i) introduzir diversidade sintética em, pelo menos, uma de CDR1 VH ou CDR2 VH desses genes; e

(ii) combinar a diversidade do passo (i) com a diversidade de CDR3 VH capturada de células B.

A presente invenção refere-se também a uma biblioteca compreendendo uma colecção de pacotes genéticos que apresentam um membro de uma família diversa de péptidos, polipéptidos ou proteínas e que, colectivamente, apresenta, pelo menos, uma porção da família, sendo os péptidos, polipéptidos ou proteínas apresentados codificados por sequências de ADN, compreendendo sequências codificando

(a) uma CDR1 VH tendo uma sequência de aminoácidos de acordo com a fórmula -X1-Y-X2-M-X3-, em que X1, X2 e X3 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y;

(b) uma CDR2 VH tendo uma sequência de aminoácidos de acordo com a fórmula X4-I-X5-X6-S-G-G-X7-T-X8-Y-A-D-S-V-K-G-, em que X4 e X5 são seleccionados independentemente do grupo consistindo em Y, R, W, V, G e S, X6 é seleccionado do grupo consistindo em P e S e X7 e X8 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; e

(c) uma sequência codificando uma CDR3 VH, em que a referida CDR3 VH é uma CDR3 VH capturada a partir da região CDR3 VH de um gene de imunoglobina de uma célula B.

Numa forma de realização preferida, os pacotes genéticos são fagos filamentosos ou fagemídeos.

A presente divulgação refere-se ainda a vectores para apresentar e/ou expressar uma família diversa de péptidos, polipéptidos ou proteínas.

A presente divulgação refere-se ainda a métodos de rastreio das bibliotecas da invenção e a péptidos, polipéptidos e proteínas identificados por esse rastreio.

ANTECEDENTES DA INVENÇÃO

É agora prática comum na técnica preparar bibliotecas de pacotes genéticos que apresentam, expressam ou compreendem um membro de uma família diversa de péptidos, polipéptidos ou proteínas e, colectivamente, apresentam, expressam ou compreendem, pelo menos, uma porção da diversidade da família. Em muitas bibliotecas comuns, os péptidos, polipéptidos ou proteínas estão relacionados com anticorpos. Frequentemente, estes são Fab ou anticorpos de cadeia única.

Em geral, os ADN que codificam membros das famílias a apresentar e/ou expressar devem ser amplificados antes destes serem clonados e utilizados para apresentar e/ou expressar o membro pretendido. Tipicamente essa amplificação utiliza iniciadores directos e inversos.

Esses iniciadores podem ser complementares a sequências nativas ao ADN a amplificar ou complementares a oligonucleótidos ligados nas extremidades 5' ou 3' desse ADN. Os iniciadores que são complementares a sequências nativas ao ADN a amplificar têm a desvantagem destas enviesarem os membros das famílias a apresentar. Serão apenas amplificados os membros que contêm uma sequência no ADN nativo que é substancialmente complementar ao iniciador. Os que não tiverem ficam ausentes da família. Para esses membros que são amplificados, será suprimida qualquer diversidade dentro da região do iniciador.

Por exemplo, na patente Europeia 368684 B1, o iniciador que é utilizado está na extremidade 5' da região V_H de um gene de anticorpo. Este emparelha com uma região da sequência de ADN nativo que é referida como sendo "suficientemente bem conservada" dentro de uma única espécie. Esse iniciador irá enviesar os membros amplificados para os que têm esta região "conservada". Qualquer diversidade dentro desta região é extinta.

É geralmente aceite que os genes de anticorpos humanos surgem através de um processo que envolve uma selecção combinatória de V e J ou V, D e J seguida por mutações somáticas. Embora a maioria da diversidade ocorra nas Regiões Determinantes da Complementaridade (CDR), ocorre também diversidade nas Regiões Estruturais mais conservadas (FR) e, pelo menos, alguma desta diversidade confere ou intensifica a ligação específica a antigénios (Ag). Como consequência, as bibliotecas devem conter tanta diversidade das CDR e FR quanto possível.

Para clonar os ADN amplificados dos péptidos, polipéptidos ou proteínas que estes codificam para apresentação num pacote genético e/ou para expressão, os ADN devem ser clivados para produzir extremidades adequadas para ligação a um vector. Essa clivagem, geralmente, é realizada utilizando locais de reconhecimento de endonucleases de restrição contidos nos iniciadores. Quando os iniciadores estão na extremidade 5' do ADN produzido a partir da transcrição reversa do ARN, essa restrição deixa regiões 5' não traduzidas nocivas no ADN amplificado. Estas regiões interferem com a expressão dos genes clonados e, deste modo, com a apresentação dos péptidos, polipéptidos e proteínas codificadas por estes.

SUMÁRIO DA INTENÇÃO

É um objectivo desta invenção proporcionar um método para produzir uma população ou biblioteca de genes de imunoglobina que compreende os passos de:

(i) introduzir diversidade sintética em, pelo menos, uma das CDR1 VH ou CDR2 VH desses genes; e

(ii) combinar a diversidade do passo (i) com a diversidade da CDR3 VH capturada a partir de células B.

Estes métodos não são enviesados para ADN que contêm sequências nativas que são complementares aos iniciadores utilizados para amplificação. Estes permitem também remover qualquer sequência que possa ser nociva à expressão do ADN amplificado antes da clonagem e apresentação e/ou expressão.

Os objectivos adicionais da invenção estão reflectidos nas reivindicações 3 a 14. Cada uma destas reivindicações é especificamente incorporada por referência nesta especificação.

BREVE DESCRIÇÃO DOS DESENHOS

A FIG. 1 é um esquema de vários métodos que podem ser empregues para amplificar genes VH sem utilizar iniciadores específicos para sequências VH.

A FIG. 2 é um esquema de vários métodos que podem ser utilizados para amplificar genes VL sem utilizar iniciadores específicos para sequências VL.

A FIG. 3 é um esquema da amplificação RACE da cadeias pesadas e leves do anticorpo.

A FIG. 4 representa a análise em gel de produtos de amplificação obtidos após a reacção de PCR primária de 4 amostras de doentes diferentes.

A FIG. 5 representa a análise em gel do ADN kappa clivado do Exemplo 2.

A FIG. 6 representa a análise em gel do ADN de extensor-kappa clivado do Exemplo 2.

A FIG. 7 representa a análise em gel do produto de PCR da amplificação de extensor-kappa do Exemplo 2.

A FIG. 8 representa a análise em gel do produto de PCR purificado a partir da amplificação de extensor-kappa do Exemplo 2.

A FIG. 9 representa a análise em gel de cadeias leves kappa clivadas e ligadas do Exemplo 2.

A FIG. 10 é um esquema para concepção da diversidade sintética de CDR1 e CDR2.

A FIG. 11 é um esquema do planeamento de clonagem para a construção do repertório da cadeia pesada.

A FIG. 12 é um esquema da clivagem e ligação da cadeia leve do anticorpo.

A FIG. 13 representa a análise em gel de cadeias leves lambda clivadas e ligadas do Exemplo 4.

A FIG. 14 é um esquema da clivagem e ligação da cadeia pesada do anticorpo.

A FIG. 15 representa a análise em gel de cadeias leves lambda clivadas e ligadas do Exemplo 5.

A FIG. 16 é um esquema de um vector de apresentação de fagos.

A FIG. 17 é um esquema de um invólucro Fab.

A FIG. 18 é um esquema de um processo para incorporação de resíduos FR1 fixos numa sequência lambda de anticorpo.

A FIG. 19 é um esquema de um processo para incorporação de resíduos FR1 fixos numa sequência kappa de anticorpo.

A FIG. 20 é um esquema de um processo para incorporação de resíduos FR1 fixos numa sequência de cadeia pesada de anticorpo.

TERMOS

Neste pedido, são utilizados os termos e as abreviaturas seguintes:

Cadeia sentido	A cadeia superior de ADN cd como normalmente escrita. Na cadeia sentido, 5'-ATG-3' codifica para Met.
Cadeia anti-sentido	A cadeia inferior de ADN cd como normalmente escrita. Na cadeia anti-sentido, 3'-TAC-5' irá corresponder a um codão Met na cadeia sentido.
Iniciador Directo	Um iniciador "directo" é complementar a uma parte da cadeia sentido e inicia a síntese de uma nova molécula de cadeia anti-sentido. "Iniciador directo" e "iniciador da cadeia inferior" são equivalentes.
Iniciador inverso	Um iniciador "inverso" é complementar a uma parte da cadeia anti-sentido e inicia a síntese de uma nova molécula de cadeia sentido. "Iniciador inverso" e "iniciador da cadeia superior" são equivalentes.
Bases	As bases são especificadas pela sua posição quer num vector quer num gene como a sua posição dentro de um gene por codão e base. Por exemplo, "89.1" é a primeira base do codão 89, 89.2 é a segunda base do codão 89.
Sv	Estreptavidina
Ap	Ampicilina
Ap^R	Um gene que confere resistência à ampicilina.

RERS	Local de reconhecimento de endonuclease de restrição
RE	Endonuclease de restrição - cliva, de um modo preferido, no RERS
URE	Endonuclease de restrição universal
Complementar funcionalmente	Duas sequências são suficientemente complementares se emparelharem sob as condições seleccionadas.
AA	Aminoácido
PCR	Reacção em cadeia da polimerase
GLG	Genes da linha germinal
Ab	Anticorpo: uma imunoglobina. O termo também abrange qualquer proteína tendo um domínio de ligação que é homólogo a um domínio de ligação de imunoglobina. Alguns exemplos de anticorpos dentro desta definição são, *inter alia*, isotipos de imunoglobina e os fragmentos Fab, $F(ab^1)_2$, scfv, Fv, dAbe e de Fd.
Fab	Molécula de duas cadeias compreendendo uma cadeia leve do Ab e parte de uma cadeia pesada.
scFv	Um Ab de cadeia simples compreendendo VH::elemento de ligação::VL ou VL::elemento de ligação::VH

w. t.	Tipo selvagem
HC	Cadeia pesada
LC	Cadeia leve
VK	Um domínio variável de uma cadeia leve Kappa.
VH	Um domínio variável de uma cadeia pesada.
VL	Um domínio variável de uma cadeia leve lambda.

Neste pedido quando é referido que os ácidos nucleicos são clivados apenas no local de quebra de uma endonuclease de restrição, deve ser entendido que pode ocorrer uma menor clivagem aleatória, *e. g.*, em locais não-específicos para além do local de quebra específico que é característico da endonuclease de restrição. O especialista na técnica irá reconhecer que essa clivagem não-específica, aleatória é de ocorrência normal. Como consequência, "apenas no local de clivagem" de uma endonuclease de restrição significa que a clivagem ocorre, de um modo preferido, no local característico dessa endonuclease.

Como utilizado neste pedido e reivindicações, o termo "local de clivagem formado pela complementação do ácido nucleico e da região em cadeia simples do oligonucleótido" inclui locais de clivagem formados pela porção em cadeia simples do oligonucleótido parcialmente em cadeia dupla formando um duplex com o ADN em cadeia simples, locais de quebra na porção em cadeia dupla do oligonucleótido parcialmente em cadeia dupla e locais de clivagem introduzidos pelo iniciador de amplificação utilizado para

amplificar a combinação de oligonucleótido de ADN em cadeia simples parcialmente em cadeia dupla.

Nos dois métodos desta invenção para preparar sequências de ácidos nucleicos em cadeia simples, é preferido o primeiro desses locais de clivagem. Nos métodos desta invenção para capturar diversidade e clonar uma família de sequências de ácidos nucleicos diversa, são preferidos os últimos dois locais de clivagem.

Neste pedido, todas as referências referidas são especificamente incorporadas por referência.

DESCRIÇÃO DETALHADA DAS FORMAS DE REALIZAÇÃO PREFERIDAS

As sequências de ácidos nucleicos que são úteis nos métodos desta invenção, *i. e.*, as que codificam, pelo menos, parcialmente os péptidos, polipéptidos e proteínas individuais apresentados ou expressos nas, ou compreendendo, as bibliotecas desta invenção, podem ser nativas, sintéticas ou uma sua combinação. Estas podem ser ARNm, ADN ou ADNc. Na forma de realização preferida, os ácidos nucleicos codificam anticorpos. De um modo muito preferido, estes codificam Fab.

Os ácidos nucleicos úteis nesta invenção podem ser naturalmente diversos, pode ser introduzida diversidade sintética nos membros naturalmente diversos ou a diversidade pode ser totalmente sintética. Por exemplo, pode ser introduzida diversidade sintética em uma ou mais CDR de genes de anticorpo. De um modo preferido, esta é introduzida em CDR1 e CDR2 de imunoglobulinas. De um modo preferido, é capturada diversidade natural nas regiões CDR3 dos genes de imunoglobina desta invenção a partir de células B. De

um modo muito preferido, os ácidos nucleicos desta invenção compreendem uma população de genes de imunoglobina que compreendem diversidade sintética em, pelo menos, uma e, de um modo mais preferido, ambas as CDR1 e CDR2 e diversidade em CDR3 capturada a partir de células B.

A diversidade sintética pode ser criada, por exemplo, por utilização da tecnologia TRIM (documento U.S. 5869644). A tecnologia TRIM permite controlar exactamente quais os tipos de aminoácidos que são permitidos em posições variegadas e em que proporções. Na tecnologia TRIM, os codões a diversificar são sintetizados utilizando misturas de trinucleótidos. Isto permite que seja incluído qualquer conjunto de tipos de aminoácidos em qualquer proporção.

Uma outra alternativa que pode ser utilizada para produzir ADN diversificado é síntese de oligonucleótidos mistos. Com a tecnologia TRIM, pode-se permitir Ala e Trp. Com a síntese de oligonucleótidos mistos, uma mistura que incluiu Ala e Trp também irá necessariamente incluir Ser e Gly. Os tipos de aminoácidos permitidos nas posições variegadas são seleccionados com referência à estrutura de anticorpos ou de outros péptidos, polipéptidos ou proteínas da família, à diversidade observada em genes da linha germinal, às mutações somáticas observadas frequentemente observadas e às áreas e aos tipos de variegação pretendidos.

Numa forma de realização preferida desta invenção, as sequências de ácidos nucleicos para pelo menos uma CDR ou outra região dos péptidos, polipéptidos ou proteínas da família são ADNc produzidos por transcrição reversa de ARNm. De um modo mais preferido, os ARNm são obtidos a partir de células sanguíneas periféricas, células de medula óssea, células do baço ou células de

nódulos linfáticos (tais como, linfócitos B ou células plasmáticas) que expressam membros de conjuntos naturalmente diversos de genes relacionados. De um modo mais preferido, os ARNm codificam uma família diversa de anticorpos. De um modo muito preferido, os ARNm são obtidos a partir de doentes sofrendo, pelo menos, de um distúrbio auto-imunitário ou cancro. De um modo preferido, são utilizados ARNm contendo uma elevada diversidade de doenças auto-imunitárias, tais como lupus eritematoso sistémico, esclerose sistémica, artrite reumatóide, síndrome antifosfolípidos e vasculite.

Numa forma de realização preferida desta invenção, os ADNc são produzidos a partir de ARNm utilizando transcrição reversa. Nesta forma de realização preferida, os ARNm são separados da célula e degradados utilizando métodos convencionais, de modo que apenas permanecem os ARNm completos (*i. e.*, rematados). O remate é, de seguida, removido e é utilizada transcrição reversa para produzir os ADNc.

A transcrição reversa da primeira cadeia (anti-sentido) pode ser realizada de qualquer modo com qualquer iniciador adequado. Ver, *e. g.*, HJ de Haard *et al.*, *Journal of Biological Chemistry*, **279**(26):18218-30 (1999). Na forma de realização preferida desta invenção, em que os ARNm codificam anticorpos, podem ser utilizados iniciadores que são complementares às regiões constantes de genes de anticorpo. Esses iniciadores são úteis porque não produzem enviesamento para subclasses de anticorpos. Numa outra forma de realização, podem ser utilizados iniciadores poli-dT (e podem ser preferidos para os genes de cadeia pesada). Alternativamente, podem ser ligadas sequências complementares ao iniciador aos terminais da cadeia anti-sentido.

Numa forma de realização preferida desta invenção, o iniciador da transcritase reversa pode ser biotinilado, permitindo deste modo que o produto de ADNc seja imobilizado em esferas com estreptavidina (Sv). A imobilização pode ser também realizada utilizando um iniciador marcado na extremidade 5' com um de a) grupo amina livre, b) tiol, c) ácido carboxílico ou d) outro grupo não encontrado no ADN que pode reagir para formar uma ligação forte com um parceiro conhecido num meio insolúvel. Se, por exemplo, for proporcionada uma amina livre (de um modo preferido, uma amina primária) na extremidade 5' de um iniciador de ADN, esta amina pode ser feita reagir com grupos de ácido carboxílico numa esfera de polímero utilizando química de formação de amida convencional. Se essa imobilização preferida for utilizada durante a transcrição reversa, a cadeia superior do ARN é degradada utilizando enzimas bem conhecidas, tais como uma combinação de ARNase H e ARNase A, antes ou após a imobilização.

As sequências de ácidos nucleicos úteis nos métodos desta invenção são geralmente amplificadas antes de serem utilizadas para apresentar e/ou expressar os péptidos, polipéptidos ou proteínas que estas codificam. Antes da amplificação, os ADN em cadeia simples podem ser clivados utilizando qualquer dos métodos descritos anteriormente. Alternativamente, os ADN em cadeia simples podem ser amplificados e depois clivados utilizando um desses métodos.

Pode ser utilizado nessa amplificação qualquer dos métodos bem conhecidos para amplificar sequências de ácidos nucleicos. São preferidos métodos que maximizam e não enviesam a diversidade. Numa forma de realização preferida desta invenção, em que as sequências de ácidos nucleicos são derivadas de genes de anticorpo, a presente invenção utiliza, de um modo preferido, iniciadores nas regiões

constantes dos genes da cadeia pesada e leve e iniciadores para uma sequência sintética que são ligados na extremidade 5' da cadeia sentido. A iniciação nessas sequências sintéticas evita a utilização de sequências dentro das regiões variáveis dos genes de anticorpo. Esses locais de iniciação nas regiões variáveis produzem enviesamento contra genes V que são quer de subclasses raras quer foram mutados nos locais de iniciação. Este enviesamento é parcialmente devido à supressão da diversidade dentro da região do iniciador e parcialmente devido à ausência de iniciação quando estão presentes muitas mutações na região complementar ao iniciador. Os métodos divulgados nesta invenção têm a vantagem de não enviesarem a população de genes de anticorpo amplificados para tipos de genes V particulares.

As sequências sintéticas podem ser ligadas à extremidade 5' da cadeia de ADN por vários métodos bem conhecidos para a ligação de sequências de ADN em conjunto. O RT CapExtention é um método preferido.

Em RT CapExtention (derivado de Smart PCR(™)), são utilizadas uma sobreposição curta (5'-...GGG-3' no iniciador da cadeia superior (USP-GGG) complementa 3'-CCC....5' na cadeia inferior) e transcritases reversas de modo que o complemento reverso do iniciador da cadeia superior é ligado à cadeia inferior.

As FIG. 1 e 2 mostram esquemas para amplificar genes VH e VL utilizando RT CapExtention. A FIG. 1 mostra um esquema da amplificação de genes VH. A FIG. 1, Painel A mostra um iniciador específico para a região poli-dT da UTR 3' iniciando a síntese da primeira cadeia inferior. São também adequados iniciadores que se ligam à região constante. O Painel B mostra a cadeia inferior prolongada na sua extremidade 3' por três C que não são

complementares ao ARNm. O Painel C mostra o resultado do emparelhamento de um iniciador sintético da cadeia superior terminando em três GGG que hibridam com CCC do terminal 3′ e prolongando a transcrição reversa prolongando a cadeia inferior através do complemento reverso da sequência do iniciador sintético. O Painel D mostra o resultado da amplificação por PCR utilizando um iniciador sintético da cadeia superior biotinilado em 5′ que replica a extremidade 5′ do iniciador sintético do painel C e um iniciador da cadeia inferior complementar a parte do domínio constante. O Painel E mostra ADNc em cadeia dupla (cd) imobilizado obtido utilizando um iniciador da cadeia superior biotinilado em 5′.

A FIG. 2 mostra um esquema semelhante para amplificação de genes VL. A FIG. 2, Painel A mostra um iniciador específico para a região constante na extremidade 3′ ou próximo iniciando a síntese da primeira cadeia inferior. São também adequados iniciadores que se ligam à região poli-dT. O Painel B mostra a cadeia inferior prolongada na sua extremidade 3′ em três C que não são complementares ao ARNm. O Painel C mostra o resultado do emparelhamento de um iniciador sintético da cadeia superior que termina em três GGG que hibridam os CCC do terminal 3′ e prolonga a transcrição reversa prolongando a cadeia inferior pelo complemento reverso da sequência do iniciador sintético. O Painel D mostra o resultado da amplificação por PCR utilizando um iniciador sintético da cadeia superior biotinilado em 5′ que replica a extremidade 5′ do iniciador sintético do painel C e um iniciador da cadeia inferior complementar a parte do domínio constante. O iniciador da cadeia inferior contém também um local de endonuclease de restrição útil, tal como AscI. O Painel E mostra ADNc cd imobilizado obtido utilizando um iniciador da cadeia superior biotinilado em 5′.

Nas FIG. 1 e 2, cada gene V consiste numa região 5' não traduzida (UTR) e um sinal de secreção, seguido pela região variável, seguida por uma região constante, seguida por uma região não traduzida 3' (que tipicamente termina em poli-A). Um iniciador inicial para transcrição reversa pode ser complementar à região constante ou ao segmento poli A da 3'-UTR. Para genes da cadeia pesada humana, é preferido um iniciador com 15 T. As transcritases reversas ligam vários resíduos C à extremidade 3' do ADN recentemente sintetizado. A RT CapExtention explora esta característica. A reacção de transcrição reversa é inicialmente realizada com apenas um iniciador da cadeia inferior. Após cerca de 1 hora, são adicionados um iniciador terminando em GGG (USP-GGG) e mais RTase. Isto faz com que o ADNc da cadeia inferior seja estendido pelo complemento reverso do USP-GGG até ao GGG final. Todos os genes V são amplificados independentemente da sua subclasse de gene V, utilizando um iniciador idêntico à parte da sequência sintética ligada e um segundo iniciador complementar a uma região de sequência conhecida na extremidade 3' da cadeia sentido,.

Numa outra forma de realização preferida, podem ser adicionadas sequências sintéticas por Amplificação Rápida das Extremidades de ADNc (RACE) (ver Frohman, M. A., Dush, M. K. e Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* (**85**):8998-9002).

A FIG. 1 mostra um esquema de amplificação por RACE das cadeias pesadas e leves do anticorpo. Primeiro, o ARNm é seleccionado por tratamento de ARN total ou poli(A+) com fosfatase intestinal de vitelo (CIP) para remover os fosfatos 5' de todas as moléculas que os possuem tais como ARN ribossómico, ARNm fragmentado, ARNt e ADN genómico. O ARNm completo (contendo uma estrutura de remate protectora de 7 metilos) não é afectado. O ARN

é de seguida tratado com pirofosfatase ácida de tabaco (TAP) para remover a estrutura de remate dos ARNm completos deixando um grupo monofosfato 5'. De seguida, é ligado um adaptador de ARN sintético à população de ARN, apenas as moléculas que têm um 5-fosfato (ARNm completos, não-rematados) irão aceitar o adaptador. São de seguida utilizadas reacções de transcritase reversa utilizando um iniciador oligodT e PCR com iniciadores internos (utilizando um iniciador do adaptador (localizado no adaptador sintético 5') e um iniciador para o gene) para amplificar o transcrito pretendido.

Numa forma de realização preferida desta invenção, o iniciador da cadeia superior ou da cadeia inferior pode ser também biotinilado ou marcado na extremidade 5' com um de a) grupo amino livre, b) tiol, c) ácido carboxílico e d) outro grupo não encontrado no ADN que pode reagir para formar uma ligação forte com um parceiro conhecido como um meio insolúvel. Estes podem ser de seguida utilizados para imobilizar a cadeia marcada após amplificação. O ADN imobilizado pode ser de cadeia simples ou dupla.

Após amplificação (utilizando *e. g.*, RT CapExtension ou RACE), os ADN desta invenção são tornados em cadeia simples. Por exemplo, as cadeias podem ser separadas utilizando um iniciador biotinilado, capturando o produto biotinilado em esferas com estreptavidina, desnaturando o ADN e removendo a cadeia complementar por lavagem. Dependendo de qual extremidade do ADN capturado é pretendida, escolher-se-á imobilizar a cadeia superior (sentido) ou a cadeia inferior (anti-sentido).

Para preparar os ADN amplificados em cadeia simples para clonagem em pacotes genéticos de modo a realizar a apresentação, ou para expressão, dos péptidos, polipéptidos ou proteínas

codificados, pelo menos parcialmente, por esses ADN, estes devem ser manipulados para proporcionar extremidades adequadas para clonagem e apresentação e/ou expressão. Em particular, devem ser removidas e substituídas quaisquer regiões 5' não traduzidas e sequências sinal de mamífero, na grelha de leitura correcta, por uma sequência sinal adequada que funciona no hospedeiro de apresentação ou de expressão. Além disso, podem ser removidas partes dos domínios variáveis (em genes de anticorpo) e substituídas por segmentos sintéticos contendo diversidade sintética. A diversidade de outras famílias genéticas pode ser, de mesmo modo, expandida com diversidade sintética.

De acordo com os métodos desta invenção, existem duas formas de manipular os ADN em cadeia simples para apresentação e/ou expressão. O primeiro método compreende os passos de:

(i) fazer contactar o ácido nucleico com um oligonucleótido em cadeia simples, sendo o oligonucleótido funcionalmente complementar ao ácido nucleico na região em que a quebra é pretendida e incluindo uma sequência que com o seu complemento no ácido nucleico forma um local de reconhecimento de uma endonuclease de restrição que em restrição resulta na quebra do ácido nucleico na posição pretendida; e

(ii) clivar o ácido nucleico apenas no local de reconhecimento formado pela complementação do ácido nucleico e do oligonucleótido;

sendo os passos de fazer contactar e de clivagem realizados a uma temperatura suficiente para manter o ácido nucleico numa forma substancialmente em cadeia simples, sendo o oligonucleótido funcionalmente complementar ao ácido nucleico ao longo de uma região suficientemente grande para permitir que as duas cadeias se

associem de modo que a clivagem possa ocorrer à temperatura seleccionada e na posição pretendida e sendo a clivagem realizada utilizando de uma endonuclease de restrição que é activa à temperatura seleccionada.

Neste primeiro método, são emparelhados oligonucleótidos curtos com o ADN em cadeia simples de modo que locais de reconhecimento da endonuclease de restrição formados dentro das regiões do ADN agora localmente em cadeia dupla possam ser clivados. Em particular, é idêntico um local de reconhecimento que ocorre na mesma posição numa fracção substancial dos ADN em cadeia simples.

Para genes de anticorpo, isto pode ser feito utilizando um catálogo de sequências da linha germinal. Ver, *e. g.*, "http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/ok.htm 1." Podem ser obtidas actualizações deste local sob o título "*Amino acid and nucleotide sequence alignments*." Existem comparações semelhantes para outras famílias e podem ser utilizadas para seleccionar regiões adequadas para clivagem e manter a diversidade.

Por exemplo, a Tabela 1 representa as sequências de ADN das regiões FR3 de 51 genes VH da linha germinal humana conhecidos. Nesta região, os genes contêm locais de reconhecimento da endonuclease de restrição mostrados na Tabela 2. São preferidas endonucleases de restrição que clivam uma grande fracção de genes da linha germinal no mesmo local em relação a endonucleases que cortam em vários locais. Além disso, é preferido que exista apenas um local para as endonucleases de restrição dentro da região a que o oligonucleótido curto se liga ao ADN em cadeia simples, *e. g.*, cerca de 10 bases de cada lado do local de reconhecimento da endonuclease de restrição.

É também mais preferida uma enzima que cliva a jusante em FR3 porque esta captura menos mutações na estrutura. Isto pode ser vantajoso em alguns casos. No entanto, é bem conhecido que existem mutações de estrutura e conferem e intensificam a ligação do anticorpo. A presente invenção permite que seja capturada toda ou parte de diversidade de FR3, por selecção do local de restrição adequado. Deste modo, o método permite também que seja capturada uma diversidade extensa.

Finalmente, são utilizados nos métodos desta invenção endonucleases de restrição que são activas entre cerca de 37 °C e cerca de 75 °C. De um modo preferido, podem ser utilizadas endonucleases de restrição que são activas entre cerca de 45 °C e cerca de 75 °C. De um modo mais preferido, são utilizadas enzimas que são activas acima de 50 °C e, de um modo muito preferido, activas a cerca de 55 °C. Essas temperaturas mantêm a sequência de ácidos nucleicos a clivar substancialmente na forma de cadeia simples.

As enzimas mostradas na Tabela 2 que cortam muitas das cadeias pesadas de genes FR3 da linha germinal numa posição única incluem: *Mae*III (24@4), *Tsp45*I (21@4), *Hph*I (44@5), *Bsa*JI (23@65), *Alu*I (23@47), *Blp*I (21@48), *Dde*I (29@58), *Bgl*II (10@61), *Msl*I (44@72), *Bsi*EI (23@74), *Eae*I (23@74), *Eag*I (23@74), *Hae*III (25@75), *Bst*4CI (51@86), *Hpy*CH4III (51@86), *Hin*fI(38@2), *Mly*I (18@2), *Ple*I (18@2), *Mnl*I (31@67), *Hpy*CH4V (21@44), *Bsm*AI (16@11), *Bpm*I (19@12), *Xmn*I (12@30) e *Sac*I (11@51). (A notação utilizada significa, por exemplo, que *Bsm*AI corta 16 dos genes FR3 da linha germinal com um local de reconhecimento de endonuclease de restrição que inicia na base 11 de FR3.)

Para clivagem de cadeias pesadas humanas em FR3, as endonucleases de restrição preferidas são: *Bst*4CI (ou *Taa*I ou *Hpy*CH4III), *Blp*I, *Hpy*CH4V e *Msl*I. Dado que ACNGT (o local de reconhecimento da endonuclease de restrição para *Bst*4CI, *Taa*I e *Hpy*CH4III) é encontrado num local consistente em todos os genes FR3 da linha germinal humana, é muito preferida uma dessas enzimas para a captura da diversidade da CDR3 da cadeia pesada. *Blp*I e *Hpy*CH4V são complementares. *Blp*I corta a maioria dos membros das famílias VH1 e VH4 enquanto *Hpy*CH4V corta a maioria dos membros das famílias VH3, VH5, VH6 e VH7. Nenhuma enzima corta VH2, mas esta é uma família muito pequena, contendo apenas três membros. Deste modo, estas enzimas também podem ser utilizadas em formas de realização preferidas dos métodos desta invenção.

Todas as endonucleases de restrição *Hpy*CH4III, *Bst*4CI e *Taa*I reconhecem 5'-ACnGT-3' e cortam o ADN em cadeia superior após n e o ADN da cadeia inferior antes da base complementar a n. Este é o local de reconhecimento de endonuclease de restrição muito preferido para este método em cadeias pesadas humanas porque este é encontrado em todos os genes da linha germinal. Além disso, a região de reconhecimento de endonuclease de restrição (ACnGT) combina as segundas e terceiras bases de um codão de tirosina (tay) e o codão cisteína seguinte (tgy) como mostrado na Tabela 3. Estes codões são altamente conservados, especialmente a cisteína em genes de anticorpo maduros.

A tabela que 4 E mostra os oligonucleótidos distintos com 22 bases de comprimento (excepto o último que tem 20 de comprimento). A tabela 5 C mostra a análise de 1617 genes de cadeia pesada de anticorpos reais. Destes, 1511 têm o local e correspondem a um dos oligonucleótidos candidatos dentro de 4 emparelhamentos incorrectos. Oito oligonucleótidos correspondem à maioria das

correspondências e são dadas na Tabela 4 F.1. Os 8 oligonucleótidos são muito semelhantes de modo que seja provável que seja obtida uma clivagem satisfatória com apenas um oligonucleótido (tal como H43.77.97.1-02#1) por ajustamento da temperatura, pH, salinidade e semelhantes. Do mesmo modo, podem ser suficientes um ou dois oligonucleótidos sempre que as sequências génicas da linha germinal diferirem muito pouco e especialmente se estas diferirem muito pouco próximo à região de reconhecimento da endonuclease de restrição a clivar. A tabela 5 D mostra uma análise de repetição de 1617 genes de cadeia pesada de anticorpo reais utilizando apenas 8 oligonucleótidos seleccionados. Isto mostra que 1463 das sequências correspondem a, pelo menos, um dos oligonucleótidos dentro de 4 emparelhamentos incorrectos e têm o local como esperado. Apenas 7 sequências têm uma segunda região de reconhecimento de endonuclease de restrição *Hpy*CH4III nesta região.

Uma outra ilustração da selecção de um local de reconhecimento de endonuclease de restrição adequado envolve clivagem em FR1 de cadeias pesadas humanas. A clivagem em FR1 permite a captura da totalidade da diversidade de CDR da cadeia pesada.

Na Tabela 6 são mostrados os genes da linha germinal para a FR1 da cadeia pesada humana. A tabela 7 mostra os locais de reconhecimento de endonuclease de restrição encontrados em genes da FR1 da linha germinal humana. Os locais preferidos são *Bsg*I (GTGCAG;39@4), *Bso*FI(GCngc:43@6,11@9,2@3,1@12), *Tse*I (Gcwgc; 43@6,11@9,2@3,1@12), *MspA1*I (CMGckg; 46@7,2@1), *Pvu*II (CAGctg; 46@7,2@1), *Alu*I (AGct; 48@82@2), *Dde*I (Ctnag; 22@52,9@48), *Hph*I (tcacc;22@80), *BssK*I (Nccngg; 35@39,2@40), *Bsa*JI (Ccnngg; 32@90,2@41), *BstN*I (CCwgg;33@40), *Scr*FI(CCngg:35@40,2@41), *EcoO109*I (RGgnccy;22@46, 11@43), *Sau96*I (Ggncc; 23@47,11@44), AvaII (Ggwcc; 23@47,4@44), *PpuM*I (RGgwccy; 22@96,9@43), *Bsm*FI(gtccc;20@48),

*Hin*fI(Gantc:34@16,21@56,21@77), *Tfi*I (21@77), *Mly*I (GAGTC;34@16), *Mly*I (gactc;21@56) e *AlwN*I (CAGnnnctg:22@68). Os locais mais preferidos são *Msp*AI e *Pvu*II. *Msp*AI e *Pvu*II têm 46 locais em 7-12 e 2 em 1-6. Para evitar a ´clivagem em ambos os locais, são utilizados oligonucleótidos que não abrangem totalmente o local em 1-6. Deste modo, o ADN não irá ser clivado nesse local. Os requerentes mostraram que o ADN que se estende 3, 4 ou 5 bases para além de um local PvuII pode ser eficientemente clivado.

Uma outra ilustração da selecção de um local de reconhecimento de endonuclease de restrição adequado envolve a clivagem em FR1 de cadeias leves kappa humanas. A tabela 8 mostra genes da FR1 de kappa da linha germinal humana e a Tabela 9 mostra locais de reconhecimento de endonuclease de restrição que são encontrados num número substancial de genes da FR1 de kappa da linha germinal humana em posições consistentes. Dos locais de reconhecimento da endonuclease de restrição listados, *Bsm*AI e *Pfl*FI são enzimas muito preferidas. Os locais *Bsm*AI são encontrados em 35 de 40 genes na base 18 da linha germinal. Os locais *Pfl*FI são encontrados em 35 de 40 genes da linha germinal na base 12.

Um outro exemplo da selecção de um local de reconhecimento de endonuclease de restrição adequado envolve a clivagem em FR1 da cadeia leve lambda humana. A Tabela 10 mostra as 31 sequências de genes da linha germinal dr FR1 da lambda humana conhecidas. A tabela 11 mostra locais de reconhecimento de endonuclease de restrição encontrados nos genes da linha germinal de FR1' da lambda humana. *Hin*FI e *Dde*I são as endonucleases de restrição mais preferidas para cortar cadeias lambda humanas na FR1.

Após serem seleccionados o local ou os locais para clivagem adequados, são preparados um ou mais oligonucleótidos curtos de

modo a complementar funcionalmente, isoladamente ou em combinação, o local de reconhecimento seleccionado. Os oligonucleótidos incluem também sequências que flanqueiam o local de reconhecimento na maioria dos genes amplificados. Esta região que flanqueia permite que a sequência emparelhe com o ADN em cadeia simples suficientemente para permitir a clivagem pela endonuclease de restrição específica para o local seleccionado.

O comprimento e a sequência reais do oligonucleótido dependem do local de reconhecimento e das condições a utilizar para fazer contactar e clivar. O comprimento deve ser suficiente para que o oligonucleótido seja funcionalmente complementar ao ADN em cadeia simples ao longo de uma região suficientemente grande para permitir que as duas cadeias se associem de modo que a quebra possa ocorrer à temperatura seleccionada e na posição pretendida.

Tipicamente, os oligonucleótidos deste método preferido da invenção têm cerca de 17 a cerca de 30 nucleótidos de comprimento. Abaixo de cerca de 17 bases, o emparelhamento é demasiado fraco e acima de 30 bases pode haver uma perda de especificidade. Um comprimento preferido é 18 a 24 bases.

Os oligonucleótidos deste comprimento não têm de ser complementos idênticos dos genes da linha germinal. Pelo contrário, podem ser tolerada a ocorrência de alguns emparelhamentos incorrectos. No entanto, de um modo preferido, não são permitidos mais de 1-3 emparelhamentos incorrectos. Esses emparelhamentos incorrectos não afectam adversamente o emparelhamento do oligonucleótido com o ADN em cadeia simples. Deste modo, diz-se que os dois ADN são funcionalmente complementares.

O segundo método para manipular os ADN em cadeia simples desta invenção para apresentação e/ou expressão compreende os passos de:

(i) fazer contactar o ácido nucleico com um oligonucleótido parcialmente em cadeia dupla, sendo a região em cadeia simples do oligonucleótido funcionalmente complementar ao ácido nucleico na região em que a clivagem é pretendida e tendo a região em cadeia dupla do oligonucleótido um local de reconhecimento de endonuclease de restrição; e

(ii) clivagem do ácido nucleico apenas no local de clivagem formado pela complementação do ácido nucleico e da região em cadeia simples do oligonucleótido;

sendo os passos de fazer contactar e de clivagem realizados a uma temperatura suficiente para manter o ácido nucleico na forma substancialmente em cadeia simples, sendo o oligonucleótido funcionalmente complementar ao ácido nucleico ao longo de uma região suficientemente grande para permitir a associação das duas cadeias de modo que a clivagem possa ocorrer à temperatura seleccionada e na posição pretendida e sendo a clivagem realizada utilizando uma endonuclease de restrição que é activa à temperatura seleccionada.

Como exposto acima, o local de clivagem pode ser formado pela porção em cadeia simples do oligonucleótido parcialmente em cadeia dupla formando um duplex com o ADN em cadeia simples, o local de clivagem pode ser contido na porção em cadeia dupla do oligonucleótido parcialmente em cadeia dupla ou o local de quebra pode ser introduzido pelo iniciador de amplificação utilizado para amplificar a combinação ADN em cadeia simples-oligonucleótido parcialmente em cadeia dupla. Nesta forma de realização, é

preferido o primeiro. E o local de reconhecimento da endonuclease de restrição pode estar localizado na porção em cadeia dupla do oligonucleótido ou introduzido pelo iniciador de amplificação, que é complementar à região em cadeia dupla, como utilizado amplificar a combinação.

De um modo preferido, o local de endonuclease de restrição é de um Tipo II-S de endonuclease restrição, cujo local de quebra está localizado a uma distância conhecida do seu local de reconhecimento.

Este segundo método, de um modo preferido, emprega a *Universal Restriction Endonucleases* ("URE"). Os URE são oligonucleótidos parcialmente em cadeia dupla. A porção em cadeia simples ou a sobreposição dos URE consiste de um adaptador de ADN que é funcionalmente complementar à sequência a clivar no ADN em cadeia simples. A porção em cadeia dupla consiste de um local de reconhecimento de uma endonuclease de restrição, de um modo preferido, do tipo II-S.

O método URE desta invenção é específico e exacto e pode tolerar alguns (*e. g.*, 1-3) emparelhamentos incorrectos nas regiões complementares, *i. e.*, é funcionalmente complementar a essa região. Além disso, as condições sob as quais a URE é utilizada podem ser ajustadas de modo que a maioria dos genes que são amplificados possam ser cortados, reduzindo o enviesamento na biblioteca produzida a partir desses genes.

A sequência do adaptador de ADN em cadeia simples ou a porção de sobreposição do URE consiste tipicamente em cerca de 14-22 bases. No entanto, podem ser utilizados adaptadores mais longos ou mais curtos. O tamanho depende da capacidade do adaptador se

associar ao seu complemento funcional no ADN em cadeia simples e da temperatura utilizada para fazer contactar o URE e o ADN em cadeia simples à temperatura utilizada para clivar o ADN com a enzima de restrição. O adaptador deve ser funcionalmente complementar ao ADN em cadeia simples ao longo de uma região suficientemente grande para permitir que as duas cadeias se associem de modo que a clivagem possa ocorrer à temperatura seleccionada e na posição pretendida. Os requerentes preferem porções em cadeia simples ou sobrepostas com 14-17 bases de comprimento e, de um modo mais preferido, com 18-20 bases de comprimento.

O local seleccionado para clivagem utilizando URE é, de um modo preferido, um que esteja substancialmente conservado na família dos ADN amplificados. Em comparação com o primeiro método de quebra desta invenção, estes locais não têm de ser locais de reconhecimento de endonuclease. No entanto, como o primeiro método, os locais seleccionados podem ser sintéticos ao contrário de existirem no ADN nativo. Esses locais podem ser seleccionados por referências às sequências de anticorpos conhecidos ou outras famílias de genes. Por exemplo, em http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/ok.html são descritas sequências de muitos genes da linha germinal. Por exemplo, um local preferido ocorre próximo da extremidade de FR3 - codão 89 até à segunda base do codão 93. A CDR3 inicia no codão 95.

Em http://www.ncbi.nlm.nih.gov/entre2/nucleótido.html estão também disponíveis sequências de 79 genes de cadeia pesada humanos. Este local pode ser utilizado para identificar sequências adequadas para a clivagem de URE de acordo com os métodos desta invenção. *Ver, e. g.*, Tabela 12B.

De um modo muito preferido, são identificadas uma ou mais sequências utilizando estes locais ou outra informação de sequência disponível. Estas sequências em conjunto estão presentes numa fracção substancial dos ADN amplificados. Por exemplo, podem ser utilizadas sequências múltiplas para levar em consideração a diversidade conhecida em genes da linha germinal ou para mutações somáticas frequentes. Podem ser também utilizadas sequências degeneradas sintéticas. De um modo preferido, é(são) seleccionada(s) uma sequência(s) que ocorre(m) em, pelo menos, 65% de genes examinados com não mais de 2-3 emparelhamentos incorrectos

São de seguida feitos adaptadores URE em cadeia simples ou sobreposições para serem complementares às regiões seleccionadas. As condições para utilizar os URE são determinadas empiricamente. Estas condições devem permitir a clivagem do ADN que contém as sequências funcionalmente complementares com não mais de 2 ou 3 emparelhamentos incorrectos mas que não permitam a clivagem do ADN isento dessas sequências.

Como descrito acima, a porção em cadeia dupla do URE inclui um local de reconhecimento de endonuclease, de um modo preferido, um local de reconhecimento do Tipo II-S. Pode ser utilizada qualquer enzima que seja activa a uma temperatura necessária para manter o ADN em cadeia simples substancialmente nessa forma e permitir que a porção do adaptador de ADN em cadeia simples do URE emparelhe durante tempo suficiente com o ADN em cadeia simples para permitir a clivagem no local pretendido.

As enzimas do Tipo II-S preferidas para utilização nos métodos URE desta invenção proporcionam a clivagem assimétrica do ADN em cadeia simples. De entre estas estão as enzimas listadas na Tabela 13. A enzima muito preferida do Tipo II-S é *Fok*I.

Quando é utilizado URE contendo a *Fok*I preferida são, de um modo preferido, utilizadas várias condições para realizar a clivagem:

1) Deverá estar presente excesso de URE em relação ao ADN alvo para activar a enzima. O URE presente apenas em quantidades equimolares em relação ao ADN alvo irá produzir pouca quebra de ADNcs porque a quantidade da enzima activa disponível irá ser limitante.

2) Pode ser utilizado um activador para activar parte da enzima *Fok*I para dimerizar sem provocar clivagem. Na Tabela 14 são mostrados exemplos de activadores adequados.

3) A reacção de clivagem é realizada a uma temperatura entre 45 °-75 °C, de um modo preferido, acima de 50 °C e, de um modo muito preferido, acima de 55 °C.

Os URE utilizados na técnica anterior continham um segmento em cadeia simples com 14 bases, uma haste de 10 bases (contendo um local *Fok*I), seguido pelo palindroma da haste de 10 bases. Embora esses URE possam ser utilizados nos métodos desta invenção, os URE preferidos desta invenção incluem também um segmento com três a oito bases (uma ansa) entre os segmentos contendo o local de reconhecimento da endonuclease restrição *Fok*I. Na forma de realização preferida, a haste (contendo o local *Fok*I) e o seu palindroma são também mais longos do que 10 bases. De um modo preferido, estes têm 10-14 bases de comprimento. Na Tabela 15 são mostrados exemplos destes adaptadores URE "lollipop".

Um exemplo de utilização de um URE para clivar um ADN em cadeia simples envolve a região FR3 da cadeia pesada humana. A Tabela 16 mostra uma análise de 840 cadeias pesadas humanas maduras completas com as sequências de reconhecimento URE mostradas. A grande maioria (718/840=0,85) irá ser reconhecida com 2 ou menos emparelhamentos incorrectos utilizando cinco URE (VHS881-1.1, VHS881-1.2, VHS881-2.1, VHS881-4.1 e VHS881-9.1). Cada um tem uma sequência adaptadora com 20 bases para complementar o gene da linha germinal, um segmento em haste com dez bases contendo um local *Fok*I, uma ansa com cinco bases e o complemento reverso do primeiro segmento da haste. O emparelhamento desses adaptadores, isoladamente ou em combinação, com o ADN da cadeia pesada anti-sentido em cadeia simples e o tratamento com *Fok*I na presença de, *e. g.*, activador FOKIact, levará à clivagem da cadeia anti-sentido na posição indicada.

Outro exemplo de utilização de (um) URE para quebrar um ADN em cadeia simples envolve a região FR1 das cadeias leves Kappa humanas. A tabela 17 mostra uma análise de 182 cadeias kappa humanos completas para a correspondência com as quatro sequências de sonda com 19 bases mostradas. Noventa e seis por cento das sequências combinam com uma das sondas com 2 ou menos emparelhamentos incorrectos. Os adaptadores URE mostrados na Tabela 17 são para civagem da cadeia sentido de cadeias kappa. Deste modo, as sequências do adaptador são o complemento reverso das sequências génicas da linha germinal. O URE consiste numa haste com dez bases, uma ansa com cinco bases, o complemento reverso da haste e a sequência de complementação. A ansa aqui mostrada é TTGTT, mas podem ser utilizadas outras sequências. A sua função consiste em interromper o palindroma das hastes para que seja favorecida a formação de um "lollypop" monomérico em relação à dimerização. A tabela 17 mostra também em que sentido é clivada a cadeia.

Um outro exemplo de utilização de um URE para clivar um ADN em cadeia simples envolve a cadeia leve lambda humana. A tabela 18 mostra a análise de 128 cadeias leves lambda humanas para a correspondência com as quatro sondas de 19 bases mostradas. 88 de 128 (69%) das cadeias combinam com uma das sondas, com três ou menos emparelhamentos incorrectos. A tabela 18 mostra também adaptadores URE correspondentes a estas sondas. O emparelhamento destes adaptadores com ADNcs da cadeia superior de cadeias lambda e tratamento com *Fok*I na presença de FokIact a uma temperatura igual ou superior a 45 °C irá originar a clivagem específica e exacta das cadeias.

As condições sob as quais as sequências oligonucleotídicas curtas do primeiro método e os URE do segundo método são feitos contactar com os ADN em cadeia simples podem ser determinadas empiricamente. As condições devem ser de modo que o ADN em cadeia simples permaneça na forma substancialmente em cadeia simples. Mais particularmente, as condições devem ser de modo que o ADN em cadeia simples não forme ansas que possam interferir com a sua associação com a sequência oligonucleotídica ou com URE ou que possam, por si, proporcionar locais da clivagem pela endonuclease de restrição seleccionada.

A eficácia e a especificidade dos oligonucleótidos curtos (primeiro método) e de URE (segundo método) podem ser ajustadas pelo controlo das concentrações dos adaptadores URE/oligonucleótidos e do ADN substrato, a temperatura, o pH, a concentração de iões metálicos, a força iónica, a concentração dos caotrópicos (tais como ureia e formamida), a concentração da endonuclease de restrição (*e. g.*, *Fok*I) e o tempo de digestão. Estas condições podem ser optimizadas com oligonucleótidos

sintéticos tendo: 1) sequências génicas da linha germinal alvo, 2) sequências génicas alvo mutadas ou 3) sequências não-alvo algo relacionadas. O objectivo é clivar a maioria das sequências alvo e quantidades mínimas das não-alvo.

De acordo com esta invenção, o ADN em cadeia simples é mantido substancialmente nessa forma utilizando uma temperatura entre cerca de 37 °C e cerca de 75 °C. De um modo preferido, é utilizada uma temperatura entre cerca de 45 °C e cerca de 75 °C. De um modo mais preferido, é utilizada, uma temperatura entre 50 °C e 60 °C, de um modo muito preferido, entre 55 °C e 60 °C. Estas temperaturas são utilizadas quando se faz contactar o ADN com o oligonucleótido ou com o URE e quando se cliva o ADN utilizando os métodos desta invenção.

Os dois métodos de quebra desta invenção têm várias vantagens. O primeiro método permite que os membros individuais da família de ADN em cadeia simples sejam clivados, de um modo preferido, num local de reconhecimento de endonuclease substancialmente conservado. O método também não requer que seja incorporado um local de reconhecimento de endonuclease na transcrição reversa ou nos iniciadores da amplificação. Pode ser utilizado qualquer local nativo ou sintético na família.

O segundo método tem ambas estas vantagens. Além disso, o método URE preferido permite que os ADN em cadeia simples sejam clivados em posições em que não ocorre naturalmente ou não foi construído sinteticamente um local de reconhecimento de endonuclease.

De um modo mais importante, ambos os métodos de clivagem permitem a utilização de iniciadores 5' e 3' de modo a maximizar a

diversidade e de seguida clivar para remover sequências não pretendidas ou nocivas antes da clonagem, apresentação e/ou expressão.

Após a clivagem dos ADN amplificados utilizando um dos métodos desta invenção, o ADN é preparado para clonagem, apresentação e/ou expressão. Isto é realizado utilizando um adaptador de ADN sintético parcialmente em duplex, cuja sequência terminal é baseada no local de quebra específico no qual o ADN amplificado foi clivado.

O ADN sintético é concebido de modo que quando este é ligado ao ADN em cadeia simples clivado na grelha de leitura adequada de modo que o péptido, o polipéptido ou a proteína pretendidos possam ser apresentados e/ou expressos na superfície do pacote genético. De um modo preferido, a porção em cadeia dupla do adaptador compreende a sequência de vários codões que codificavam a sequência de aminoácidos característica da família de péptidos, polipéptidos ou proteínas até o local de quebra. Para cadeias pesadas humanas, são utilizados, de um modo preferido, os aminoácidos de estrutura 3-23 para proporcionar as sequências necessárias para a expressão do ADN clivado.

De um modo preferido, a porção em cadeia dupla do adaptador tem cerca de 12 a 100 bases de comprimento. De um modo mais preferido, são utilizadas cerca de 20 a 100 bases. A região padrão dupla do adaptador contém também, de um modo preferido, pelo menos, um local de reconhecimento de endonuclease útil para clonar o ADN numa vector de apresentação e/ou de expressão adequado (ou um vector receptor utilizado para arquivar a diversidade). Este local de endonuclease de restrição pode ser nativo das sequências génicas da linha germinal utilizadas para prolongar a sequência de ADN.

Este pode também ser construído utilizando sequências degeneradas em relação às sequências génicas da linha germinal nativas. Ou, este pode ser totalmente sintético.

A porção em cadeia simples do adaptador é complementar à região da quebra no ADN em cadeia simples. A sobreposição pode ser de cerca de 2 bases até cerca de 15 bases. Quanto mais longa for a sobreposição, mais eficiente será provavelmente a ligação. Uma extensão preferida para a sobreposição é 7 a 10. Isto permite alguns emparelhamentos incorrectos na região de modo que possa ser capturada diversidade nesta região.

A região em cadeia simples ou a sobreposição do adaptador parcialmente em duplex são vantajosas porque estas permitem que seja capturado ADN clivado no local seleccionado, mas não outros fragmentos. Esses fragmentos iriam contaminar a biblioteca com genes codificando sequências que não irão enrolar em anticorpos adequados e provavelmente irão ser aderentes não-especificamente.

Uma ilustração da utilização do adaptador parcialmente em duplex nos métodos desta invenção envolve a ligação desse adaptador a uma região FR3 humana que foi clivada, como descrito acima, em 5'-ACnGT-3' utilizando HpyCH4III, Bst4CI ou TaaI.

A tabela 4 F.2 mostra a cadeia inferior da porção em cadeia dupla do adaptador para ligação ao ADN da cadeia inferior clivado. Uma vez que o local HpyCH4III se situa tão à direita (como mostrado na Tabela 3), pode ser adicionadz uma sequência que inclui o local *Afl*II assim como o local *Xba*I. Esta porção de cadeia inferior do adaptador parcialmente em duplex, H43.XAExt, incorpora locais *Xba*I e *Afl*II. A cadeia superior da porção em cadeia dupla do adaptador não tem qualquer local (devido a emparelhamentos incorrectos

planeados nos segmentos opostos aos locais *Xba*I e *Afl*II de H43.XAExt), mas irá emparelhar muito compactamente com H43.XAExt. H43AExt contém apenas o local *Afl*II e destina-se a utilizar com as cadeias superiores H43.ABr1 e H43.ABr2 (que têm alterações intencionais para destruir o local *Afl*II).

Após a ligação, o ADN pretendido, capturado pode ser amplificado por PCR novamente, se pretendido, utilizando na forma de realização preferida um iniciador para a região constante a jusante do gene do anticorpo e um iniciador para parte da região padrão dupla do adaptador. Os iniciadores podem também conter locais de endonuclease de restrição para utilização na clonagem do ADN amplificado.

Após a ligação e possivelmente amplificação, do adaptador parcialmente em cadeia dupla ao ADN amplificado em cadeia simples, o ADN composto é clivado em locais de reconhecimento de endonuclease 5' e 3' seleccionados.

Os locais de clivagem úteis para a clonagem dependem do fago ou do fagemídeo ou de outros vectores em que a invólucro será inserido e dos locais disponíveis nos genes de anticorpo. A tabela 19 proporciona dados de endonuclease de restrição para 75 cadeias leves humanas. A Tabela 20 mostra dados correspondentes a 79 cadeias pesadas humanas. Em cada Tabela, as endonucleases são ordenadas pela frequência de corte crescente. Nestas Tabelas, Nch é o número de cadeias cortadas pela enzima e Ns é o número de locais (algumas cadeias têm mais de um local).

Desta análise, *Sfi*I, *Not*I, *Afl*II, *ApaL*I e *Asc*I são muito adequadas. *Sfi*I e *Not*I são, de um modo preferido, utilizadas em pCES1 para inserir o segmento de apresentação da cadeia pesada.

ApaLI e *Asc*I são, de um modo preferido, utilizadas em pCES1 para inserir o segmento de apresentação da cadeia leve.

Os locais *Bst*EII ocorrem em 97% dos genes JH da linha germinal. Em genes V rearranjados, apenas 54/79 (68%) dos genes de cadeia pesada contêm um local *Bst*EII e 7/61 destes contêm dois locais. Deste modo, 47/79 (59%) contêm um único local *Bst*EII. Uma alternativa à utilização de *Bst*EII é clivar através de URE na extremidade de JH e ligar a um oligonucleótido sintético que codifica parte de CH1.

Um exemplo de preparação de uma família de sequências de ADN utilizando os métodos desta invenção envolve capturar a diversidade da CDR 3 humana. Como descrito acima, são transcritos reversamente ARNm provenientes de vários doentes auto-imunitários na cadeia inferior de ADNc. Após a cadeia superior de ARN ser degradada, a cadeia inferior é imobilizada e é utilizado um oligonucleótido curto para quebrar o ADNc a montante de CDR3. É de seguida emparelhado um adaptador de ADN sintético parcialmente em duplex com o ADN e o ADN é amplificado utilizando um iniciador para o adaptador e um iniciador para a região constante (após FR4). O ADN é de seguida clivado utilizando *Bst*EII (em FR4) e uma endonuclease de restrição adequada ao adaptador parcialmente em cadeia dupla (*e. g.*, *Xba*I e *Afl*II (em FR3)). O ADN é então ligado numa estrutura base VH sintética tal como 3-23.

Um exemplo de preparação de um ADN em cadeia simples que foi clivado utilizando o método URE envolve a cadeia Kappa humana. O local de quebra na cadeia sentido desta cadeia é representado na Tabela 17. O oligonucleótido kapextURE é emparelhado com os oligonucleótidos (kaBR01UR, kaBR02UR, kaBR03UR e kaBR04UR) para formar um ADN parcialmente em duplex. Este ADN é então ligado às

cadeias kappa solúveis quebradas. O produto de ligação é de seguida amplificado utilizando iniciadores kapextUREPCR e CKForeAsc (que inserem um local AscI após o final de C kappa). Este produto é de seguida clivado com *ApaLI* e *Asc*I e ligado a um vector receptor cortado de um modo semelhante.

Um outro exemplo envolve a quebra de cadeias leves lambda, ilustradas na Tabela 18. Após a quebra, são emparelhados um extensor (on_LamEx133) e quatro oligonucleótidos de ponte (ON_LamB1-133, ON-LamB2-133, ON_LamB3-133 e ON-LamB4-133) para formar um ADN parcialmente em duplex. Esse ADN é ligado às cadeias sentido da cadeia lambda quebrada. Após a ligação, o ADN é amplificado com ON_Lam133PCR e um iniciador directo específico para o domínio constante de lambda, tal como CL2ForeAsc ou CL7ForeAsc (Tabela 130).

Nas cadeias pesadas humanas, podem-se clivar quase todos os genes em FR4 (a jusante, *i. e.*, em direcção à extremidade 3' da cadeia sentido, de CDR3) num local *Bst*EII que ocorre numa posição constante numa fracção muito grande de genes V da cadeia pesada humana. Necessita-se então de um local em FR3, se for pretendido capturar apenas a diversidade de CDR3, em FR2, se for pretendida a diversidade CDR2 e CDR3 ou em FR1, se for pretendida toda a diversidade CDR. Estes locais são, de um modo preferido, inseridos como parte do adaptador parcialmente em cadeia dupla.

A presente divulgação refere-se também aos vectores receptores (*e. g.*, para apresentação e/ou expressão) tendo locais que permitem a clonagem de cadeias quer leves quer pesadas. Esses vectores são bem conhecidos e amplamente utilizados na técnica. Um vector de apresentação de fagos preferido de acordo com invenção é o fago

MALIA3. Este apresenta o gene III. A sequência do fago MALIA3 é mostrada na Tabela 21A (anotada) e Tabela 21B (condensada).

O ADN codificando as regiões seleccionadas das cadeias leves ou pesadas pode ser transferido para os vectores utilizando endonucleases que cortam cadeias quer leves quer pesadas apenas muito raramente. Por exemplo, as cadeias leves podem ser capturadas com *ApaLI* e *Asc*I. Os genes de cadeia pesada são, de um modo preferido, clonados num vector receptor tendo locais *Sfi*I, *Nco*I, *Xba*I, *Afl*II, *Bst*EII, *Apa*I e *Not*I. As cadeias leves são, de um modo preferido, transferidas para biblioteca como fragmentos *ApaLI-Asc*I. As cadeias pesadas são, de um modo preferido, transferidas para biblioteca como fragmentos *Sfi*I-*Not*I.

De um modo muito preferido, a apresentação ocorre na superfície de um derivado do fago M13. O vector muito preferido contém todos os genes M13, um gene de resistência a antibiótico e o invólucro de apresentação. O vector preferido está proporcionado de locais de restrição que permitem a introdução e a excisão de membros da família diversa de genes, como invólucros. O vector preferido é estável contra o rearranjo nas condições de crescimento utilizadas para amplificar o fago.

Numa outra forma de realização desta invenção, a diversidade capturada pelos métodos da presente invenção pode ser apresentada e/ou expressa num vector fagemídico (*e. g.*, pCES1) que apresenta e/ou expressa o péptido, o polipéptido ou a proteína. Esses vectores podem ser também utilizados para armazenar a diversidade para apresentação e/ou expressão subsequente utilizando outros vectores ou fagos.

Numa outra forma de realização desta invenção, a diversidade capturada pelos métodos da presente invenção pode ser apresentada e/ou expressa num vector de levedura.

Numa outra forma de realização, a forma de apresentação pode ser através de um elemento de ligação curto para ancorar os domínios - compreendendo uma âncora possível a porção final de M13 III ("IIIstump") e sendo uma segunda âncora possível a proteína madura III completa.

O fragmento IIIstump contém suficiente M13 III para montagem no fago mas não os domínios envolvidos na mediação da infectividade. Dado que estão presentes proteínas III w. t. é improvável que o fago elimine os genes do anticorpo e o fago que elimina de facto estes segmentos recebe apenas uma muito pequena vantagem de crescimento. Para cada um dos domínios âncora, o ADN codifica a sequência de AA w. t., mas difere da sequência de ADN w. t. numa extensão muito elevada. Isto irá reduzir muito o potencial para recombinação homóloga entre a âncora e o gene w. t. que está também presente (ver o Exemplo 6).

De um modo muito preferido, A presente invenção utiliza um fago completo contendo um gene de resistência a antibiótico (tal como um gene de resistência à ampicilina) e a cassete de apresentação. Dado estão presentes os genes *iii* e possivelmente *viii* w. t., estão também presentes as proteínas w. t. A cassete de apresentação é transcrita a partir de um promotor regulável (*e. g.*, P_{Lacz}). A utilização de um promotor regulável permite que o controlo da proporção entre o gene de apresentação de fusão e a proteína do invólucro w. t. correspondente. Esta proporção determina o número médio de cópias da fusão de apresentação por partícula de fago (ou de fagemídeo).

Numa outra forma de realização dos métodos desta invenção, o fago ou fagemídeo podem apresentar e/ou expressar proteínas para além de Fab, substituindo as porções Fab indicadas acima, por outros genes de proteína.

Podem ser utilizados vários hospedeiros no aspecto da apresentação e/ou da expressão desta invenção. Esses hospedeiros são bem conhecidos na técnica. Na forma de realização preferida, em que estão a ser apresentados e/ou expressos Fab, o - hospedeiro preferido deve crescer a 30 °C e ser RecA- (para reduzir a recombinação genética não pretendida) e EndA- (para tornar mais fácil a recuperação do ADN RF). É também preferido que a estirpe hospedeira seja facilmente transformada por electroporação.

XL1-Blue MRF' satisfaz a maioria destas preferências, mas não cresce bem a 30 °C. XL1-Blue MRF' cresce de facto lentamente a 38 °C e, deste modo, é um hospedeiro aceitável. TG-1 é também um hospedeiro aceitável embora este seja RecA* e EndA*. XL1-Blue MRF' é mais preferida para o hospedeiro intermédio utilizado para acumular a diversidade antes da construção final da biblioteca.

Após apresentação e/ou expressão, as bibliotecas desta invenção podem ser rastreadas utilizando técnicas bem conhecidas e utilizadas convencionalmente. Os péptidos, os polipéptidos ou as proteínas seleccionados podem ser de seguida utilizados para tratar a doença. Geralmente, os péptidos, os polipéptidos ou as proteínas para utilização em terapia ou em composições farmacêuticas são produzidos por isolamento do ADN codificando o péptido, o polipéptido ou a proteína pretendidos a partir do membro da biblioteca seleccionado. Esse ADN é de seguida utilizado em métodos convencionais para produzir o péptido, os polipéptidos ou a

proteína que este codifica em células hospedeiras adequadas, de um modo preferido, células hospedeiras de mamífero, *e. g.*, células CHO. Após isolamento, o péptido, o polipéptido ou a proteína são utilizados isoladamente ou com composições farmaceuticamente aceitáveis em terapia para tratar a doença.

EXEMPLOS

Exemplo 1: Amplificação RACE de repertórios da cadeia pesada e leve de anticorpo de doentes auto-imunitários.

Foi isolado ARN total a partir de amostras de sangue individuais (50 mL) de 11 doentes utilizando um kit RNAzolTM (CINNA/Biotecx), como descrito pelo fabricante. Os doentes foram diagnosticados como se segue:

1. SLE e síndrome de fosfolípido

2. esclerose sistémica limitada

3. SLE e síndrome de Sjogren

4. Esclerose Sistémica limitada

5. Artrite Reumatóide com vasculite activa

6. Esclerose sistémica limitada e Síndrome de Sjogren

7. Artrite Reumatóide e vasculite (não activa)

8. SLE e síndrome de Sjogren

9. SLE

10. SLE e glomerulonefrite (activa)

11. Poliartrite/Fenómeno de Raynauds

Foi isolado ARN Poli-A+ a partir destas 11 amostras de ARN total, utilizando o kit Promega PolyATtract® mRNA Isolation (Promega).

Foram utilizadas 250 ng de cada amostra de ARN poli-A+ para amplificar as cadeias pesadas e leves de anticorpo com o kit GeneRAacerTM (n°. de cat. Invitrogen L1500-01). Na FIG. 3 é mostrado um resumo esquemático do processo RACE.

Foi ligado um adaptador de ARN à extremidade 5' de todos os ARNm utilizando o protocolo geral do kit GeneRAacer™. Depois, realizou-se uma reacção da transcritase reversa na presença de iniciador específico oligo (dT15) sob condições descritas pelo fabricante no kit GeneRAacer™.

Foi utilizado 1/5 do ADNc da reacção da transcritase reversa numa reacção PCR de 20 µL. Foi utilizado um iniciador directo baseado na cadeia CH1 de IgM [HuCmFOR] e um iniciador inverso baseado na sequência do adaptador sintético ligado [5'A], para a amplificação do repertório da cadeia pesada de IgM. (Ver a Tabela 22)

Foram utilizados um iniciador directo que contém a extremidade 3' de codificação do ADNc [HuCkFor e HuCLFor2+HuCLfor7] e um iniciador inverso baseado na sequência do adaptador sintético

ligado [5'A] para a amplificação das cadeias leves kappa e lambda, (Ver a Tabela 22). Foram obtidos produtos de amplificação específicos após 30 ciclos de PCR primária.

A FIG. 4 mostra os produtos de amplificação obtidos após a reacção de PCR primária de 4 amostras de doentes diferentes. Foram analisados 8 µL de produtos de PCR primária de 4 doentes diferentes [marcados 1,2, 3 e 4] num gel agarose. Para a cadeia pesada, é obtido um produto de, aproximadamente, 950 nt enquanto para as cadeias leves kappa e lambda o produto é, aproximadamente, 850 nt. M1-2 são marcadores de peso molecular.

Os produtos de PCR foram também analisados por sequenciação de ADN [10 clones dos repertórios da cadeia lambda, kappa ou pesada]. Todos os genes de anticorpo sequenciados recuperados continham a sequência codificante completa assim como a sequência líder 5' e a diversidade do gene V foi a diversidade esperada (em comparação com dados de literatura).

Foram misturadas 50 ng de todas as 11 amostras amplificadas individuais para as cadeias pesada, leve lambda ou leve kappa e foram utilizadas em reacções PCR secundárias.

Em todas as PCR secundárias foi utilizado, aproximadamente, 1 ng de ADN molde da mistura de PCR primária em reacções de PCR 50 µL múltiplas [25 ciclos].

Para a cadeia pesada, foi utilizado um iniciador directo biotinilado interno [HuCm-Interno] e foi utilizado um iniciador inverso da extremidade 5' interno localizado na sequência do adaptador sintético [5'NA]. Foi biotinilada a extremidade 5' da cadeia inferior da cadeia pesada.

Para as cadeias leves, foi utilizado iniciador interno biotinilado da extremidade 5' no adaptador sintético [5'NA] em combinação com um iniciador da extremidade 3' na região constante de Ckappa e Clambda, prolongado com uma sequência codificando para o local de restrição AscI [kappa: HuCkForAscI, Lambda: HuCL2-FOR-ASC + HuCL7-FOR-ASC]. [Foi biotinilado o ADN da extremidade 5' da cadeia superior]. Após a análise em gel, os produtos da PCR secundária foram reunidos e purificados com Promega Wizzard PCR cleanup. Foram isolados aproximadamente 25 µg de ADN da cadeia pesada, cadeia leve lambda e kappa biotinilado dos 11 doentes.

Exemplo 2: Captura de cadeias kappa com BsmAI.

Foi preparado um repertório de ARNm da cadeia kappa humana utilizando o método RACE do Exemplo 1 a partir de uma colecção de doentes tendo várias doenças auto-imunitárias.

Este Exemplo seguiu o protocolo do Exemplo 1. Foram imobilizados, aproximadamente, 2 microgramas (µg) de material RACE do gene da cadeia kappa humana (Igkappa) com biotina ligada à extremidade 5' da cadeia superior como no Exemplo 1 em 200 microlitros (µL) de esferas magnéticas Seradyn. A cadeia inferior foi removida por lavagem do ADN com 2 alíquotas de 200 µL de NaOH 0,1 M (pH 13) durante 3 minutos para a primeira alíquota seguidos de 30 segundos para a segunda alíquota. As esferas foram neutralizadas com 200 µL de Tris 10 mM (pH 7,5) NaCl 100 mM. Os oligonucleótidos curtos mostrados na Tabela 23 foram adicionados num excesso molar de 40 vezes em 100 µL de tampão NEB 2 (NaCl 50 mM, Tris-HCl 10 mM, $MgCl_2$ 10 mM, ditiotreitol 1 mM, pH 7,9) às esferas secas. A mistura foi incubada a 95 °C durante 5 minutos

depois arrefecida até 55 °C ao longo de 30 minutos. O excesso de oligonucleótido foi removido por lavagem com 2 lavagens com tampão NEB 3 (NaCl 100 mM, Tris-HCl 50 mM, $MgCl_2$ 10 mM, ditiotreitol 1 mM, pH 7,9). Foram adicionadas dez unidades de BsmAI (NEB) em tampão NEB 3 e incubadas durante 1 h a 55 °C. O ADN clivado a jusante foi recolhido e purificado através de uma coluna de purificação Qiagen PCR (FIG. 5 e 6).

A FIG. 5 mostra uma análise de ADN kappa em cadeia simples digerido. Foram emparelhadas, aproximadamente, 151,5 pmol do adaptador com 3,79 pmol do ADN kappa em cadeia simples imobilizado seguido por digestão com 15 U de *BsmAI*. O sobrenadante contendo o ADN pretendido foi removido e analisado com gel de poliacrilamida a 5% juntamente com as esferas remanescentes que continham ADN kappa completo não clivado. Foram purificadas 189 pmol de ADN em cadeia simples clivado para análise adicional. Cinco por cento do ADNcs completo original permaneceram nas esferas.

A FIG. 6 mostra uma análise do extensor - ligação kappa clivada. Foram ligadas 180 pmol de ponte/extensor pré-emparelhada a 1,8 pmol de ADN em cadeia simples digerido com *BsmAI*. O ADN ligado foi purificado com uma coluna de purificação Qiagen PCR e analisado num gel de poliacrilamida a 5%. Os resultados indicaram que a ligação do extensor ao ADN em cadeia simples foi eficiente a 95%.

Foi preparado um adaptador parcialmente em cadeia dupla utilizando o oligonucleótido mostrado na Tabela 23. O adaptador foi adicionado ao ADN em cadeia simples num excesso molar de 100 vezes juntamente com 1000 unidades de ADN ligase T4 e incubado durante a noite a 16 °C. O excesso de oligonucleótido foi removido com uma coluna de purificação Qiagen PCR. O material ligado foi amplificado

por PCR utilizando os iniciadores kapPCRt1 e kapfor mostrados na Tabela 23 durante 10 ciclos com o programa mostrado na Tabela 24.

O produto de PCR solúvel foi separado num gel e apresentou uma banda com, aproximadamente, 700 n, como esperado (FIG. 7 e 8). O ADN foi clivado com as enzimas *ApaLI* e *Asc*I, purificado em gel e ligado ao vector pCES1 clivado de um modo semelhante.

A FIG. 7 mostra uma análise do produto de PCR da amplificação extensor-kappa. O ADN em cadeia simples extensor-kappa ligado foi amplificado com iniciadores específicos para o extensor e para a região constante da cadeia leve. Foram utilizadas duas concentrações de molde diferentes, 10 ng *versus* 50 ng, como molde e foram utilizados 13 ciclos para produzir, aproximadamente, 1,5 µg de ADNcd como mostrado por análise em gel de agarose a 0,8%.

A FIG. 8 mostra uma análise do produto de PCR purificado a partir da amplificação extensor-kappa. Foram separados, aproximadamente, 5 µg de ADN em cadeia dupla extensor-kappa amplificado por PCR num gel de agarose a 0,8%, excisado e extraído com uma coluna de purificação de gel GFX. Foram preparados 3,5 µg de ADN em cadeia dupla por análise em gel.

O ensaio para capturar cadeias kappa com BsmA1 foi repetido e produziu resultados semelhantes. A FIG 9A mostra o ADN após este ter sido clivado e recolhido e purificado através de um coluna de purificação Qiagen PCR. A FIG. 9B mostra o adaptador parcialmente em cadeia dupla ligado ao ADN em cadeia simples. Este material ligado foi, de seguida, amplificado (FIG. 9C). O gel apresentou uma banda com, aproximadamente, 700 n.

A tabela 25 mostra a sequência de ADN de uma cadeia leve kappa capturada por este processo. A Tabela 26 mostra uma segunda sequência capturada por este processo. A sequência ponte mais próxima foi complementar à sequência 5'-agccacc-3', mas a sequência capturada apresenta 5'-Tgccacc-3', mostrando que é tolerado algum emparelhamento incorrecto na região que fica sobreposta.

Exemplo 3: Construção de Diversidade de CDR1 e CDR2 Sintéticas na Estrutura V-3-23 VH.

Foi criada diversidade sintética na Região Determinante da Complementaridade (CDR) 1 e 2 na estrutura 3-23 VH num processo de dois passos: primeiro, foi construído um vector contendo a estrutura 3-23 VH; e depois, foi montada e clonada uma CDR 1 e 2 sintéticas neste vector.

Para a construção de estrutura 3-23 VH, foram concebidos 8 oligonucleótidos e dois iniciadores de PCR (oligonucleótidos longos - TOPFR1A, BOTFR1B, BOTFR2, BOTFR3, F06, BOTFR4, ON-vgC1 e ON-vgC2 e iniciadores - SFPRMET e BOTPCRPRIM, mostrados na Tabela 27) que se sobrepõem com base na sequência Genebank da região da estrutura 3-23 VH. A concepção incorporou, pelo menos, um local de restrição útil em cada região estrutural, como mostrado na Tabela 27. Na Tabela 27, os segmentos que foram sintetizados são mostrados a negrito, as regiões sobrepostas estão sublinhadas e as regiões de iniciação de PCR em cada extremidade estão sublinhadas.

Foi combinada uma mistura destes 8 oligonucleótidos numa concentração final de 2,5 uM numa reacção de PCR de 20 uL. A mistura PCR continha dNTP 200 uM, $MgCl_2$ 2,5 mM, 0,02 U de ADN Polimerase *Pfu Turbo*™, 1 U de ADN Polimerase Taq Qiagen HotStart e

tampão de PCR Qiagen 1 x. O programa de PCR consistiu em 10 ciclos de 94 °C durante 30 s, 55 °C durante 30 s e 72 °C durante 30 s.

A sequência de ADN VH 3-23 montada foi, de seguida, amplificada, utilizando 2,5 µL de uma diluição 10 vezes da PCR inicial numa reacção de PCR de 100 uL. A reacção de PCR continha dNTP 200 uM, $MgCl_2$ 2,5 mM, 0,02 U de ADN Polimerase *Pfu Turbo*™, 1 U de ADN Polimerase Taq Qiagen HotStart, Tampão de PCR Qiagen 1 x e 2 iniciadores externos (SFPRMET e BOTPCRPRIM) numa concentração de 1 uM. O programa de PCR consistiu em 23 ciclos a 94 °C durante 30 s, 55 °C durante 30 s e 72 °C durante 60 s. A sequência de ADN 3-23 VH foi digerida e clonada em pCES1 (vector fagemídeo) utilizando os locais restrição de endonuclease *Sfi*I e *Bst*EII. Todas as enzimas de restrição aqui mencionadas foram fornecidas por New England BioLabs, Beverly, MA e utilizadas de acordo com as instruções do fabricante.

Foram introduzidas sequências de enchimento (mostradas na Tabela 28 e Tabela 29) em pCES1 para substituir as sequências CDR1/CDR2 (900 bases entre os locais RE BspEI e *Xba*I) e sequências CDR3 (358 bases entre *Afl*II e *Bst*EII) antes da clonagem da diversidade de CDR1/CDR2. Este novo vector foi denominado pCES5 e a sua sequência é dada na Tabela 29.

Ter os enchimentos no lugar das CDR evita o risco de uma sequência parental estar sobre-representada na biblioteca. As sequências de enchimento são fragmentos do gene da penicilase de *E. coli*. O enchimento CDR1-2 contém locais de restrição para *Bgl*II, *Bsu36*I, *Bcl*I, *Xcm*I, *Mlu*I, *Pvu*II, *Hpa*I e *Hinc*II, sendo os locais sublinhados únicos dentro do vector pCES5. O enchimento que substitui CDR3 contém o local de endonuclease de restrição *Rsr*II único.

Na FIG. 10 é mostrada uma representação esquemática da concepção para diversidade sintética de CDR1 e CDR2. A concepção foi baseada na presença de mutações em DP47/3-23 e genes da linha germinal relacionados. A diversidade foi concebida para ser introduzida nas posições dentro de CDR1 e CDR2 indicadas pelos números na FIG. 10. A diversidade em cada posição foi seleccionada para ser um dos três esquemas seguintes: 1 = ADEFGHIKLMNPQRSTVWY; 2 = YRWVGS; 3 = PS, em que as letras codificam misturas equimolares dos aminoácidos indicados.

Para a construção da diversidade CDR1 e CDR2, foram concebidos 4 oligonucleótidos sobrepostos (ON-vgC1, ON_Br12, ON_CD2Xba e ON-vgC2, mostrados na Tabela 27 e na Tabela 30) codificando CDR1/2, mais regiões flanqueantes. Foi combinada uma mistura destes 4 oligonucleótidos numa concentração final de 2,5 uM numa reacção de PCR de 40 uL. Dois dos 4 oligonucleótidos continham sequências variadas posicionadas nas CDR1 e CDR2. A mistura de PCR continha dNTP 200 uM, 2,5 U de ADN Polimerase Pwo (Roche) e tampão de PCR Pwo 1 x com $MgSO_4$ 2 mM. O programa de PCR consistiu em 10 ciclos em 94 °C durante 30 s, 60 °C durante 30 s e 72 °C durante 60 s. Esta sequência de ADN de CDR1/2 montada foi amplificada, utilizando 2,5 µL da mistura numa reacção de PCR de 100 uL. A reacção de PCR continha dNTP 200 uM, 2,5 U de ADN Polimerase Pwo, Tampão de PCR Pwo 1 x com $MgSO_4$ 2 mM e 2 iniciadores externos numa concentração de 1 uM. O programa de PCR consistiu em 10 ciclos a 94 °C durante 30 s, 60 °C durante 30 s e 72 °C durante 60 s. Essas sequências variegadas foram digeridas e clonadas na estrutura 3-23 VH no lugar do enchimento CDR1/2.

Os requerentes obtiveram aproximadamente 7×10^7 transformantes independentes. A diversidade de CDR3 quer de

populações dadoras quer de ADN sintético pode ser clonada no vector contendo diversidade CDR1 e CDR 2 sintética.

Na FIG. 11 é mostrada uma representação esquemática deste processo. Foi feita uma sequência codificando as regiões FR do segmento do gene V3-23 humano e regiões CDR com diversidade sintética por montagem de oligonucleótidos e clonagem através dos locais *Xba1* e *BspE1* num vector que complementa as regiões FR1 e FR3. Nesta biblioteca de segmentos VH sintéticos, a sequência VH-CDR3 complementar (direita superior) foi clonada através dos locais Xbal e BstEll. Os genes CH clonados resultante contêm uma combinação de diversidade sintética concebida e diversidade natural (ver a FIG. 11).

Exemplo 4: Clivagem e ligação de cadeias leves lambda com HinfI.

Nas FIG. 12A e 12B é mostrado uma esquema da quebra e ligação de cadeias leve do anticorpo. Foram imobilizadas, aproximadamente, 2 µg do ADN de Lambda humano biotinilado preparado como descrito no Exemplo 1 em 200 µL de esferas magnéticas Seradyn. A cadeia inferior foi removida por incubação do ADN com 200 µL de 0,1 M NaOH (pH = 13) durante 3 minutos, o sobrenadante foi removido e foi realizada uma lavagem adicional de 30 segundos com 200 µL de NaOH 0,1 M. O sobrenadante foi removido e as esferas foram neutralizadas com 200 µL de Tris 10 mM (pH = 7,5), NaCl 100 mM. Foram realizadas 2 lavagens adicionais com 200 µL de tampão NEB2 2, contendo Tris 10 mM (pH = 7,9), NaCl 50 mM, MgCl2 10 mM e ditiotreitol 1 mM. Após a imobilização, a quantidade de ADNcs foi estimada num gel de PAGE-UREIA a 5%.

Foram recuperadas cerca de 0,8 µg de ADNcs e incubadas em 100 µL de tampão NEB2 2 contendo um excesso molar de 80 vezes de uma mistura equimolar de ON_LamlaB7, ON_Lam2aB7, ON Lam31B7 e ON_Lam3rB7 [cada oligonucleótido num excesso molar de 20 vezes] (ver a Tabela 31).

A mistura foi incubada a 95 °C durante 5 minutos e depois arrefecida lentamente até 50 °C ao longo de um período de 30 minutos. O excesso do oligonucleótido foi removido por lavagem com 2 lavagens de 200 µL de tampão NEB 2. Foram adicionadas 4 U/µg de *Hinf* I e incubadas durante 1 hora a 50 °C. As esferas foram misturadas a cada 10 minutos.

Após a incubação a amostra foi purificada através de um coluna de purificação Qiagen PCR e, posteriormente, analisada num gel PAGE-ureia a 5% (ver a FIG. 13A, a clivagem foi eficiente a mais de 70%).

Na FIG. 12B é mostrado um esquema da ligação das cadeias leves clivadas. Foi preparada uma mistura de pares de ponte/extensor a partir dos oligonucleótidos de Pte/Ext listados na Tabela 31 (excesso molar total de 100 vezes) em 1000 U de ADN Ligase T4 (NEB) e foi incubada durante a noite a 16 °C. Após a ligação do ADN, o excesso de oligonucleótido foi removido com uma coluna de purificação Qiagen PCR e a ligação foi verificada num gel Ureia-PAGE (ver a FIG. 13B; a ligação foi eficiente a mais de 95%).

Foram realizados PCR múltiplos contendo 10 ng do material ligado numa reacção de PCR de 50 µL utilizando ON lamPlePCR 25 pMol e 25 pmol de uma mistura de equimolar do iniciador Hu-CL2AscI/HuCL7AscI (ver o Exemplo 1).

O PCR foi realizado a 60 °C durante 15 ciclos utilizando polimerase Pfu. Foi recuperado cerca de 1 µg de ADNcd por PCR (ver a FIG. 13C) e clivado com *ApaL1* e *Asc*I para clonar as cadeias leve lambda em pCES2.

Exemplo 5: Captura da população CDR3 da cadeia pesada humana.

Nas FIG. 14A e 14B é mostrado um esquema da clivagem e ligação das cadeias leves do anticorpo.

Foram imobilizadas, aproximadamente, 3 µg de material RACE do gene da cadeia pesada humana (IgM) com biotina ligada à extremidade 5' da cadeia inferior em 300 µL de esferas magnéticas Seradyn. A cadeia superior foi removida por lavagem do ADN com 2 alíquotas de 300 µL de NaOH 0,1 M (pH 13) durante 3 minutos para a primeira alíquota seguidos por 30 segundos para a segunda alíquota. As esferas foram neutralizadas com 300 µL de Tris 10 mM (pH 7,5) NaCl 100 mM. Foram adicionados os REdaptors (oligonucleótidos utilizados para tornar ADN localmente em cadeia simples em cadeia dupla) mostrados na Tabela 32 num excesso molar de 30 vezes em 200 µL de tampão NEB 4 (Acetato de Potássio 50 mM, Tris-Acetato 20 mM, Acetato de Magnésio 10 mM, ditiotreitol 1 mM pH 7,9) às esferas secas. Os REdaptors foram incubados com o ADN em cadeia simples a 80 °C durante 5 minutos depois foram arrefecidos até 55 °C ao longo de 30 minutos. O excesso de REdaptors foi removido por lavagem com 2 lavagens com tampão NEB 4. Foram adicionadas quinze unidades de HpyCH4III (NEB) a tampão NEB 4 e incubadas durante 1 hora a 55 °C. O ADN a jusante clivado que permanece nas esferas foi removido das esferas utilizando uma coluna de remoção Qiagen Nucleotide (ver a FIG. 15).

Os pares Ponte/Extensor mostrados na Tabela 33 foram adicionados num excesso molar de 25 juntamente com 1200 unidades do ADN ligase de T4 e incubados durante a noite a 16 °C. O excesso de Ponte/Extensor foi removido com um coluna de purificação Qiagen PCR. O material ligado foi amplificado por PCR utilizando os iniciadores H43.XAExtPCR2 e Hucumnest mostrados na Tabela 34 durante 10 ciclos com o programa mostrado na Tabela 35.

O produto de PCR solúvel foi separado num gel e apresentou uma banda com aproximadamente 500 n, como esperado (ver a FIG. 15B). O ADN foi clivado com as enzimas SfiI e *Not*I, purificado em gel e ligado ao vector PCES1 clivado da mesma forma.

Exemplo 6: Descrição do Vector de Apresentação em Fagos CJRA05, um membro da biblioteca incorporada no vector DY3F7.

A tabela 36 contém uma sequência de ADN anotada de um membro da biblioteca, CJRA05, ver a FIG. 16. A tabela 36 deve ser lida como se segue: em cada linha tudo que se segue a um ponto de exclamação "!" é um comentário. Todas as ocorrências de A, C, G e T antes de "!" são a sequência de ADN. As maiúsculas são utilizadas apenas para mostrar que algumas bases constituem características especiais, tais como locais de restrição, locais de ligação de ribossomas e semelhantes, que são legendados abaixo do ADN. O CJRA05 é um derivado do fago DY3F7, obtido clonando um fragmento *ApaL*I a *Not*I nestes locais em DY3F31. DY3F31 é semelhante a DY3F7 excepto por a cadeia leve e os genes de cadeia pesada terem sido substituídos pelo ADN "de enchimento" que não codifica para qualquer anticorpo. DY3F7 contém um anticorpo que se liga à estreptavidina, mas não provém da presente biblioteca.

Os genes de fago iniciam com o gene ii e continuam com genes x, v, vii, ix, viii, iii, vi, i e iv. O gene iii foi ligeiramente modificado por terem sido inseridos oito codões entre a sequência sinal e a proteína madura e por os aminoácidos finais da sequência sinal terem sido alterados. Isto permite que estejam presentes os locais de reconhecimento da enzima de restrição *Eag*I e *Xba*I. Após o gene iv está a origem de replicação (ori) do fago. Após ori está bla que confere resistência à ampicilina (ApR). Os genes fágicos e bla são transcritos no mesmo sentido.

Após bla, encontra-se o invólucro Fab (ilustrada na FIG. 17) compreendendo:

a) O promotor PlacZ,

b) Um primeiro Local de Ligação ao Ribossoma (RBS1),

c) A forma de sequência sinal M13 iii,

d) Um RERS *ApaL*I,

e) Uma cadeia leve (uma kappa L20::JK1 encurtada por um codão no limite V-J neste caso),

f) Um RERS *Asc*I,

g) Um segundo Local de Ligação ao Ribossoma (RBS2),

h) Uma sequência sinal, de um modo preferido, PelB, que contém,

i) Um RERS *Sfi*I,

j) Uma região 3-23 V sintética com diversidade em CDR1 e CDR2,

k) Uma CDR3 capturada,

l) Uma região J parcialmente sintética (FR4 após BstEII),

m) CH1,

n) Um RERS NotI,

o) Uma marcação His6,

p) Uma marcação cMyc,

q) Um codão âmbar,

r) Um ADN âncora que codifica a mesma sequência de aminoácidos que os codões 273 a 424 de M13 iii (como mostrado na Tabela 37).

s) Dois codões de terminação,

t) Um RERS *Avr*II, e

u) Um terminador trp.

A âncora (item r) codifica a mesma sequência de aminoácidos do que os codões 273 a 424 de M13 iii mas o ADN é, aproximadamente, tão diferente quanto possível da sequência de ADN de tipo selvagem. Na Tabela 36, III' stump vai da base 8997 à base 9455. Abaixo do ADN, como comentários, estão as diferenças em relação ao tipo

selvagem iii para os codões comparáveis com "!W.T" nos finais destas linhas. Refira-se que Met e Trp têm apenas um codão único e devem ser mantidos como tal. Estes tipos de AA são raros. Os codões Ser podem ser modificados na totalidade das três base, enquanto os codões Leu e Arg podem ser modificados em duas.

Na maioria de casos, pode ser introduzida uma modificação de uma base por codão. Isto tem três vantagens: 1) a recombinação com o gene do tipo selvagem contido em outro local no fago é menos provável, 2) podem ser introduzidos novos locais de restrição, facilitando a construção; e 3) podem ser concebidos iniciadores de sequenciação que se ligam em apenas uma das duas regiões.

O fragmento de M13 III mostrado em CJRA05 tem o comprimento preferido para o segmento âncora. Podem ser também utilizados segmentos âncora mais longos ou mais curtos alternativos definidos em relação à proteína III madura completa.

A sequência de M13 III consiste nos elementos seguintes: Sequência Sinal::Domínio 1 (D1)::Elemento de ligação 1 (L1)::Domínio 2 (D2)::Elemento de ligação 2 (L2)::Domínio 3 (D3)::Segmento Transmembranar (TM)::âncora Intracelular (IC) (ver a Tabela 38).

A âncora pIII (também conhecida como trpIII) consiste, de um modo preferido, em D2::L2::D3::TM::IC. Uma outra forma de realização da âncora pIII consiste em D2'::L2::D3::TM::IC (em que D2' compreende os 21 últimos resíduos de D2 com os 109 primeiros resíduos eliminados). Uma forma de realização adicional da âncora pIII consiste em D2'(C>S)::L2::D3::TM::IC (em que D2'(C>S) é D2' com o único C convertido em S) e d) D3::TM::IC.

A Tabela 38 mostra um fragmento génico compreendendo o local *Not*I, a marcação His6, a marcação cMyc, um codão âmbar, um local de quebra da enterocinase recombinante e a proteína M13 III madura completa. O ADN utilizado para codificar esta sequência é intencionalmente muito diferente do ADN do gene do tipo selvagem iii como mostrado pelas linhas designadas "W. T.". contendo as bases w. t. em que estas diferem deste gene. III está dividido em domínios designados "domínio 1", "elemento de ligação 1", "domínio 2", "elemento de ligação 2", "domínio 3", "segmento transmembranar" e "âncora intracelular".

Os segmentos âncora alternativos preferidos (definidos por referência à sequência da Tabela 38) incluem:

os codões 1-29 ligados aos codões 104-435, eliminando o domínio 1 e conservando o elemento de ligação 1 até à extremidade;

os codões 1-38 ligados aos codões 104-435, eliminando o domínio e conservando o local de clivagem rEK mais o elemento de ligação 1 até à extremidade de III;

os codões 1-29 ligados aos codões 236-435, eliminando o domínio 1, o elemento de ligação 1 e a maioria do domínio 2 e conservando o elemento de ligação 2 até à extremidade;

os codões 1-38 ligados aos codões 236-435, eliminando o domínio 1, elemento de ligação 1 e a maioria do domínio 2 e conservando o elemento de ligação 2 até à extremidade e o local de clivagem rEK;

os codões 1-29 ligados aos codões 236-435 e alterando o codão 240 para Ser (*e. g.*, agc), eliminando o domínio 1, o elemento de ligação 1 e a maioria do domínio 2 e conservando o elemento de ligação 2 até à extremidade; e

os codões 1-38 ligados aos codões 236-435 e alterando o codão 240 para Ser (*e. g.*, agc), eliminando o domínio 1, o elemento de ligação 1 e a maioria do domínio 2 e conservando o elemento de ligação 2 até à extremidade e o local de clivagem rEK.

As construções serão mais facilmente realizadas por métodos semelhantes aos de Wang e Wilkinson (*Biotechniques* 2001: **31**(4) 722-724) em que é utilizada PCR para copiar o vector excepto a parte a eliminar e são introduzidos ou conservados locais de restrição correspondentes em qualquer extremidade da parte a manter. A Tabela 39 mostra os oligonucleótidos a utilizar na eliminação de partes do segmento âncora III. O ADN mostrado na Tabela 38 tem um local *Nhe*I antes do local de clivagem da enterocinase recombinante DINDDRMA (rEKCS). Se for utilizada *Nhe*I no processo de eliminação com este ADN, o local rEKCS irá ser perdido. Este local poderá ser bastante útil na clivagem de Fab a partir do fago e poderá facilitar a captura de anticorpos de afinidade muito elevada. Pode-se mutagenizar esta sequência de modo que o local *Nhe*I siga o local rEKCS, uma sequência de aminoácidos Ala Ser está já presente. Alternativamente, pode-se utilizar *Sph*I para as eliminações. Isto iria envolver uma ligeira modificação na sequência de aminoácidos mas não teria consequência.

Exemplo 7: Selecção de ligantes de antigénio provenientes de uma biblioteca enriquecida de anticorpos humanos utilizando o vector fágico DY3F31.

Neste exemplo, a biblioteca de anticorpos humanos utilizada é descrita em de Haard *et al.*, (*Journal of Biological Chemistry*, **274**(26): 18218-30 (1999). Esta biblioteca, consistindo numa grande biblioteca de fagemídeos de Fab humanos não-imunitários, foi inicialmente enriquecida em antigénio, em estreptavidina ou em fenilo-oxazolona (phOx). Os métodos para isto são bem conhecidos na técnica. Foram seleccionadas para reclonagem duas bibliotecas de Fab pré-seleccionadas, a primeira seleccionada numa vez em phOx-BSA imobilizado (R1-ox) e a segunda seleccionada duas vezes em estreptavidina (R2-strep).

Estes repertórios enriquecidos dos anticorpos fágicos, em que apenas uma percentagem muito baixa tem a actividade de ligação ao antigénio utilizado na selecção, foram confirmados por rastreio de clones num ELISA para ligação do antigénio. Foram transferidos genes Fab seleccionados a partir do vector fagemídico desta biblioteca para o vector DY3F31 através de locais de restrição *ApaL1-Not1*.

Foi pré-tratado ADN do vector fágico DY3F31 com DNAse dependente de ATP para remover o ADN cromossómico e depois foi digerido com *ApaL1* e *Not1*. Foi realizada uma digestão adicional com *Asc*I intermediamente para prevenir a auto-ligação do vector. O fragmento Fab *ApaL1/Not*I das bibliotecas pré-seleccionadas foi subsequentemente ligado ao ADN do vector e transformado em células XL1-blue MRF' competentes.

As bibliotecas foram feitas utilizando proporções vector:inserção de 1:2 para a biblioteca phOx e 1:3 para a biblioteca STREP e utilizando 100 ng de ADN ligado por 50 µL de células competentes por electroporação (condições de electroporação: uma descarga de 1700 V, 1 hora de recuperação das células em meio SOC rico, plaqueamento em placas de agar contendo ampicilina).

Esta transformação resultou num tamanho de biblioteca de 1,6 x 10^6 para R1-ox em DY3F31 e 2,1 x 10^6 para R2-strep em DY3F31. Foram rastreadas dezasseis colónias de cada biblioteca para a inserção e todas apresentaram a inserção de tamanho correcto (± 1400 bp) (para ambas as bibliotecas).

O fago foi preparado a partir destes bibliotecas de Fab como se segue. Foi inoculada uma amostra representativa da biblioteca em meio com ampicilina e glicose e a OD 0,5, o meio foi permutado para ampicilina e IPTG 1 mM. Após crescimento durante a noite a 37 °C, o fago foi recolhido a partir do sobrenadante por precipitação com PEG-NaCl. O fago foi utilizado para a selecção do antigénio. Foi seleccionado R1-ox em phOx-revestido com BSA por adsorção passiva em imunotubos e R2-strep em esferas paramagnéticas revestidas com estreptavidina (Dynal, Noruega), em processos descritos em de Haard *et al.* e Marks *et al.*, *Journal of Molecular Biology*, **222**(3):581-97 (1991). Na Tabela 40 são proporcionados os títulos fágicos e os enriquecimentos.

Foram rastreados clones destas bibliotecas seleccionadas, respectivamente R2-ox duplicado e R3-strep, para ligação aos seus antigénios em ELISA. Foram seleccionados aleatoriamente e rastreados 44 clones de cada selecção como fago ou Fab solúvel para ligação em ELISA. Para as bibliotecas em DY3F31, os clones foram

inicialmente cultivados em 2TY-glicose a 2%-AMP 50 µg/mL até uma OD600 de aproximadamente 0,5 e de seguida foram cultivados durante a noite em 2TY- AMP 50 µg/mL +/- IPTG 1 mM. A indução com IPTG pode resultar na produção tanto de fago-Fab como de Fab solúvel. Consequentemente os (mesmos) clones foram também cultivados sem IPTG. A Tabela 41 mostra os resultados de um rastreio ELISA do sobrenadante resultante, quer para a detecção de partículas fágicas com ligação ao antigénio (Anti-M13 HRP = anticorpo anti-fago) ou para a detecção da Fab humanos, seja no fago ou como fragmentos solúveis, quer com utilização do anticorpo anti-myc 9E10 que detecta a marcação myc que cada Fab contém na extremidade do terminal C da cadeia pesada seguida por um soro de coelho-anti-Murganho marcado com HRP (coluna 9E10/RAM-HRP) ou com reagente anti-cadeia leve seguido por um antisoro de cabra-anti-coelho marcado com HRP (anti-CK/CL Gar-HRP).

Os resultados mostram que em ambas os casos os ligantes dos antigénios são identificados na biblioteca, com como Fab em fago ou com os reagentes anti-Fab (Tabela 41). A indução com IPTG produz um aumento no número de positivos. Pode ser também verificado que para os clones phOx, o ELISA com fago produz mais positivos do que o ELISA com Fab solúvel, muito provavelmente devido à ligação ávida do fago. Foram rastreados vinte e quatro dos clones positivos em ELISA utilizando PCR de Fab-inserção do vector, seguido por digestão com *BstN*I. Isto produziu 17 padrões diferentes para os Fab ligantes de phOx em 23 amostras que foram correctamente analisadas e 6 em 24 para os clones ligantes à estreptavidina. Deste modo, os dados proveniente da selecção e rastreio a partir desta biblioteca pré-enriquecida de Fab não-imunitários mostram que o vector DY3F31 é adequado para apresentação e selecção de fragmentos Fab e proporciona tanto Fab solúveis como Fab em fago para experiências de rastreio após selecção.

Exemplo 8: Selecção de bibliotecas anticorpo-Fago em esferas magnéticas com estreptavidina.

O exemplo seguinte descreve uma selecção em que inicialmente se esgota uma amostra da biblioteca de ligantes à estreptavidina e opcionalmente de ligantes a um não-alvo (*i. e.*, uma molécula que não o alvo que não se pretende a que o Fab seleccionado se ligue). É colocada a hipótese de ser especialmente útil ter uma molécula, denominada um "ligando competitivo", que se ligasse ao alvo e que um anticorpo que se ligasse ao mesmo local.

Para este processo foram bloqueadas Esferas Magnéticas de Estreptavidina (Dynal) uma vez com solução de bloqueamento (Marvel Milk 2%, PBS (pH 7,4), Tween-20 0,01% ("MPBST 2%")) durante 60 minutos à temperatura ambiente e depois foram lavadas cinco vezes com MPBST 2%. Foram bloqueados 450 µL de esferas para cada esgotamento e subsequente selecção estabelecida.

Por selecção, foram adicionados 6,25 µL do alvo de esgotamento biotinilado (armazenamento a 1 mg/mL em PBST) a 0,250 mL de esferas lavadas, bloqueadas (provenientes do passo 1). Foi permitida a ligação do alvo durante a noite, com rotação, a 4 °C. No dia seguinte, as esferas são lavadas 5 vezes com PBST.

Por selecção, foram adicionados 0,010 mL do antigénio alvo biotinilado (armazenamento a 1 mg/mL em PBST) a 0,100 mL de esferas bloqueadas e lavadas (do passo 1). Foi permitida a ligação do antigénio durante a noite, com rotação, a 4 °C. No dia seguinte, as esferas foram lavadas 5 vezes com PBST.

No ciclo 1, foram bloqueadas contra ligação não-específica 2 x 10^{12} até 10^{13} unidades formadoras de placas (pfu) por selecção por adição a 0,500 mL de MPBS 2% (=MPBST 2% sem Tween) durante 1 hora à TA (rotação). Em ciclos posteriores, foram bloqueadas 10^{11} pfu por selecção como realizado no ciclo 1.

Cada conjunto de fagos foi incubado com 50 µL de esferas alvo de esgotamento (removendo o sobrenadante da lavagem final imediatamente antes da utilização) num rotor Labquake durante 10 min à temperatura ambiente. Após incubação, o sobrenadante de fagos foi removido e incubado com mais 50 µL de esferas alvo de esgotamento. Isto foi repetido mais 3 vezes utilizando esferas alvo de esgotamento e duas vezes utilizando esferas com estreptavidina bloqueadas num total de 7 ciclos de esgotamento, pelo que cada conjunto de fago requereu 350 µL de esferas de esgotamento.

Foi recolhida uma pequena amostra de cada conjunto da biblioteca esgotada para titulação. Cada conjunto da biblioteca foi adicionado a 0,100 mL de esferas alvo (removendo o sobrenadante da lavagem final imediatamente antes da utilização) e foi permitida a incubação durante 2 horas à temperatura ambiente (rotação).

As esferas foram de seguida lavadas tão rapidamente quanto possível (*e. g.*, total de 3 minutos) com 5 x 0,500 mL de PBST e depois com PBS 2 x. O fago ainda ligado a esferas após a lavagem foi eluído uma vez com 0,250 mL de ligando competitivo (~1 µµM) em PBST durante 1 hora à temperatura ambiente num rotor Labquake. O eluído foi removido, misturado com 0,500 mL de solução de sais Mínimos A e armazenado. Para uma segunda selecção, foram utilizados 0,500 mL de TEA 100 mM para eluição durante 10 min à TA, depois foi neutralizado numa mistura de 0,250 mL de Tris 1 M, pH 7,4 + 0,500 mL de sais Min A.

Após a primeira selecção eluição, as esferas podem ser eluídas novamente com 0,300 mL de alvo não-biotinilado (1 mg/mL) durante 1 hora à TA num rotor Labquake. O fago eluído é adicionado a 0,450 mL de sais Mínimos A.

Foram mantidos separados três eluídos (competidor da 1ª selecção, alvo da 1ª selecção e eluição com TEA neutralizada da 2ª selecção) e uma pequena alíquota recolhida a partir de cada para titulação. Foram adicionados 0,500 mL de sais Mínimos A às alíquotas de esferas restantes após eluição do competidor e do alvo e após eluição com TEA. Foi realizada a recolha de uma pequena alíquota de cada para titulação.

Cada eluição e cada conjunto de esferas eluídas foi misturado com 2X YT e uma alíquota (*e. g.*, 1 mL com 1. E 10/mL) de células *E. coli* XL1-Blue MRF' (ou outra linha de células F') que tinha sido arrefecida em gelo após ter sido cultivada até ao meio da fase logarítmica, esgotada e concentrada (ver o processo abaixo - "preparação Mid-Log de células XL-1 blue MRF' para infecção").

Após, aproximadamente, 30 minutos à temperatura ambiente, as misturas fago/célula foram espalhadas em Placas Bio-Assay (243 x 243 x 18 mM, Nalge Nunc) contendo 2XYT, agar com IPTG 1 mM. As placas foram incubadas durante a noite a 30 °C. No dia seguinte, cada cultura de fagos amplificada foi recolhida a partir da sua placa respectiva. A placa foi submersa com 35 mL de TBS ou LB e as células foram raspadas a partir da placa. As células ressuspensas foram transferidas para um tubo de centrífuga. Foram utilizados 20 mL adicionais de TBS ou de LB para remover quaisquer células da placa e foram recolhidos com as células no tubo de centrífuga. As células foram removidas por centrifugação e foi recuperado o fago

no sobrenadante por precipitação com PEG. As preparações de fago amplificado foram tituladas ao longo do dia seguinte.

No primeiro ciclo, duas selecções produziram cinco eluídos amplificados. Estes eluídos amplificados foram isolados durante mais 2-3 ciclos adicionais da selecção utilizando ~1. E 12 fagos/ciclo. Para cada ciclo adicional, foram preparadas esferas de esgotamento e alvo na noite anterior ao ciclo ser iniciado.

Para os passos de eluição em ciclos subsequentes, foram realizados todas as eluições até ao passo de eluição de que proveio a eluição amplificada e as eluições anteriores foram tratadas como lavagens. Para, por exemplo, a esfera de fago amplificado por infecção, foram realizadas as eluições do ligando competitivo e do alvo e de seguida foram sacudidas como lavagens (ver abaixo). Depois as esferas foram utilizadas para infectar *E. coli*. Consequentemente, dois conjuntos produziram um total de 5 eluições finais no final da selecção.

1º conjunto de selecção

A. Eluição amplificada do ligando: eluir c/ ligando durante 1 hora, manter como eluição

B. Eluição amplificada do alvo: eluir c/ ligando durante 1 hora, sacudir como lavagem eluir o alvo durante 1 hora, manter como eluição

C. Eluição de esferas amp. infect.: eluir c/ ligando durante 1 hora, sacudir como lavagem eluir c/ alvo durante 1 hora,

sacudir como lavagem eluir c/ infecção celular, manter como eluição

2º conjunto de selecção

A. Eluição amplificada de TEA; eluir c/ TEA durante 10 min, manter como eluição

B. Eluição de esferas amp. infect.; eluir c/ TEA durante 10 min, sacudir como lavagem eluir c/ infecção celular, manter como eluição

Preparação mid-log de células XL1 blue MRF' para infecção (baseado em Barbas *et al. Phage Display manual procedure*)

Cultura de XL1 blue MRF' em NZCYM (12,5 mg/mL tet) a 37 °C e 250 rpm durante a noite. É iniciada uma cultura de 500 mL num frasco de 2 litros por diluição da células 1/50 em NZCYM/tet (são adicionados 10 mL de cultura durante a noite) e incubada a 37 °C a 250 rpm até que tivessesido atingida uma OD600 de 0,45 (1,5-2 horas). A agitação foi reduzida até 100 rpm durante 10 min Quando a OD600 atingiu entre 0,55-0,65, as células foram transferidas para tubos de centrífuga 2 x 250 mL, centrifugadas a 600 g durante 15 min a 4 °C. O sobrenadante foi vertido. O líquido residual foi removido com uma pipeta.

Os sedimentos foram ressuspensos suavemente (não pipetando para cima e para baixo) no volume original de 1 x de sais Mínimos A à temp. ambiente. As células ressuspensas foram transferidas de novo para o frasco de 2 litros, agitado a 100 rpm durante 45 min a

37 °C. Este processo foi realizado para esgotar as células e restaurar os *pili*. As células foram transferidas para tubos de centrífuga 2 x 250 mL e centrifugadas como anteriormente.

As células foram ressuspensas suavemente em sais Mínimos A gelados (5 mL por 500 mL de cultura original). As células foram colocadas em gelo para utilização em infecções logo que possível.

Os eluídos de fago foram levados até 7,5 mL com meio 2XYT e foram adicionados 2,5 mL de células. As esferas foram levadas até 3 mL com 2XYT e foi adicionado 1 mL de células. Incubadas a 37 °C durante 30 min. As células foram plaqueadas em placas de NUNC grandes com agar 2XYT, IPTG 1 mM e incubadas durante 18 h a 30 °C.

Exemplo 9: Incorporação de região sintética na região FR1/3.

Estão descritos abaixo exemplos para incorporar resíduos fixos em sequências do anticorpo para genes da cadeia leve kappa e lambda e para cadeias pesadas. As condições experimentais e os oligonucleótidos utilizados nos exemplos abaixo foram descritos em exemplos anteriores (*e. g.*, Exemplos 3 e 4).

O processo para incorporar resíduos FR1 fixos numa sequência lambda de anticorpo consiste em 3 passos (ver a FIG. 18): (1) emparelhamento do material de ADN em cadeia simples codificando genes VL com uma mistura de oligonucleótidos parcialmente complementares (indicados com Ext e Ponte), para emparelhar neste exemplo com a região codificando os resíduos 5-7 da FR1 dos genes lambda (indicada com X..X; a sobreposição pode não ser por vezes perfeita dentro dos genes lambda); (2) ligação deste complexo; (3) PCR do material ligado com o iniciador indicado ('PCRpr') e por

exemplo um iniciador baseado dentro do gene VL. Neste processo os primeiros poucos resíduos de todos os genes lambda irão ser codificados pelas sequências presentes nos oligonucleótidos (Ext., Ponte ou PCRpr). Após a PCR, os genes lambda podem ser clonados utilizando o local de restrição indicado para ApaLI.

O processo para incorporar resíduos FR1 fixos numa sequência kappa de anticorpo (FIG. 19) consiste em 3 passos: (1) emparelhamento do material de ADN em cadeia simples codificando genes VK com uma mistura de oligonucleótidos parcialmente complementares (indicados com Ext e Pon), para emparelhar neste exemplo à região codificando os resíduos 8-10 da FR1 dos genes kappa (indicada com X.. X; a sobreposição pode não ser por vezes perfeita dentro dos genes kappa); (2) ligação deste complexo; (3) PCR do material ligado com o iniciador indicado ('PCRpr') e por exemplo um iniciador baseado dentro do gene VK. Neste processo os primeiro poucos (8) resíduos de todos os genes kappa irão ser codificados pelas sequências presentes nos oligonucleótidos (Ext., Ponte ou PCRpr.). Após a PCR, os genes kappa podem ser clonados utilizando o local de restrição indicado para ApaLI.

O processo incorporar resíduos FR3 fixos num anticorpo numa sequência de cadeia pesada (FIG. 20) consiste em 3 passos: (1) emparelhamento do material de ADN em cadeia simples codificando parte dos genes VH (por exemplo codificando as regiões FR3, CDR3 e FR4) com uma mistura de oligonucleótidos parcialmente complementares (indicados com Ext e Ponte), para emparelhar neste exemplo com a região codificando os resíduos 92-94 (dentro da região FR3) dos genes VH (indicada com X..X; a sobreposição pode não ser por vezes perfeita dentro dos genes VH); (2) ligação deste complexo; (3) PCR do material ligado com o iniciador indicado ('PCRpr') e por exemplo um iniciador baseado dentro do gene VH (tal

como na região FR4). Neste processo alguns resíduos da totalidade dos genes VH serão codificados pelas sequências presentes nos oligonucleótidos aqui utilizados, em particular de PCRpr (para os resíduos 70-73) ou provenientes de oligonucleótidos Ext/Ponte (resíduos 74-91). Depois a PCR, os genes VH parciais podem ser clonados utilizando o local de restrição indicado para *Xba*I.

Irá ser entendido que o anterior é apenas ilustrativo dos princípios desta invenção e que podem ser realizadas várias modificações pelos especialista na técnica sem se afastarem do âmbito e do espírito da invenção.

Tabela 1: Sequências GLG FR3 humanas

```
! VH1
!   66  67   68  69   70  71  72  73  74  75  76  77  78  79  80
    agg gtc acc atg acc agg gac acg tcc atc agc aca gcc tac atg
!   81  82  82a 82b 82c 83  84  85  86  87  88   89  90  91   92
    gag ctg agc agg ctg aga tct gac gac acg gcc gtg tat tac tgt
!   93  94  95
    gcg aga ga  ! 1-02# 1
    aga gtc acc att acc agg gac aca tcc gcg agc aca gcc tac atg
    gag ctg agc agc ctg aga tct gaa gac acg gct gtg tat tac tgt
    gcg aga ga ! 1-03# 2
    aga gtc acc atg acc agg aac acc tcc ata agc aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
    gcg aga gg ! 1-08# 3
    aga gtc acc atg acc aca gac aca tcc acg agc aca gcc tac atg
    gag ctg agg agc ctg aga tct gac gac acg gcc gtg tat tac tgt
    gcg aga ga ! 1-18# 4
    aga gtc acc atg acc gag gac aca tct aca gac aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
    gca aca ga ! 1-24# 5
    aga gtc acc att acc agg gac agg tct atg agc aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac aca gcc atg tat tac tgt
    gca aga ta ! 1-45# 6
    aga gtc acc atg acc agg gac acg tcc acg agc aca gtc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
    gcg aga ga ! 1-46# 7
    aga gtc acc att acc agg gac atg tcc aca agc aca gcc tac atg
    gag ctg agc agc ctg aga tcc gag gac acg gcc gtg tat tac tgt
    gcg gca ga ! 1-58# 8
    aga gtc acg att acc gcg gac gaa tcc acg agc aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
    gcg aga ga ! 1-69# 9
    aga gtc acg att acc gcg gac aaa tcc acg agc aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc gtg tat tac tgt
    gcg aga ga ! 1-e# 10
    aga gtc acc ata acc gcg gac acg tct aca gac aca gcc tac atg
    gag ctg agc agc ctg aga tct gag gac acg gcc grg tat tac tgt
    gca aca ga ! 1-f# 11
```

! VH2

agg ctc acc atc acc aag gac acc tcc aaa aac cag gtg gtc ctt
aca atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt
gca cac aga c! 2-05# 12
agg ctc acc atc tcc aag gac acc tcc aaa agc cag gtg gtc ctt
acc atg acc aac atg gac cct gtg gac aca gcc aca tat tac tgt
gca cgg ata c! 2-26# 13
agg ctc acc atc tcc aag gac acc tcc aaa aac cag gtg gtc ctt
aca atg acc aac atg gac cct gtg gac aca gcc acg tat tac tgt
gca cgg ata c! 2-70# 14

! VH3

cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt
gcg aga ga ! 3-07# 15
cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg
caa atg aac agt ctg aga gct gag gac acg gcc ttg tat tac tgt
gca aaa gat a! 3-09# 16
cga ttc acc atc tcc agg gac aac gcc aag aac tca ctg tat ctg
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt
gcg aga ga ! 3-11# 17
cga ttc acc atc tcc aga gaa aat gcc aag aac tcc ttg tat ctt
caa atg aac agc ctg aga gcc ggg gac acg gct gtg tat tac tgt
gca aga ga ! 3-13# 18
aga ttc acc atc tca aga gat gat tca aaa aac acg ctg tat ctg
caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt
acc aca ga ! 3-15# 19
cga ttc acc atc tcc aga gac aac gcc aag aac tcc ctg tat ctg
caa atg aac agt ctg aga gcc gag gac acg gcc ttg tat cac tgt
gcg aga ga ! 3-20# 20
cga ttc acc atc tcc aga gac aac gcc aag aac tca ctg tat ctg
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt
gcg aga ga ! 3-21# 21
cgg ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg
caa atg aac agc ctg aga gcc gag gac acg gcc gta tat tac tgt
gcg aaa ga ! 3-233# 22
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt
gcg aaa ga ! 3-30# 23
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt

gag aga ga ! 3303# 24
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt
gcg aaa ga ! 3305#25
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctg
caa atg aac agc ctg aga gcc gag gac acg gct gtg tat tac tgt
gcg aga ga ! 3-33# 26
cga ttc acc atc tcc aga gac aac agc aaa aac tcc ctg tat ctg
caa atg aac agt ctg aga act gag gac acc gcc ttg tat tac tgt
gca aaa gat a ! 3-43# 27
cga ttc acc atc tcc aga gac aat gcc aag aac tca ctg tat ctg
caa atg aac agc ctg aga gac gag gac acg gct gtg tat tac tgt
gcg aga ga ! 3-48# 28
aga ttc acc atc tca aga gat ggt tcc aaa agc atc gcc tat ctg
caa atg aac agc ctg aaa acc gag gac aca gcc gtg tat tac tgt
act aga ga! 3-49# 29
cga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt
caa atg aac agc ctg aga gcc gag gac acg gcc gtg tat tac tgt
gcg aga ga ! 3-53# 30
aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt
caa atg ggc agc ctg aga gct gag gac atg gct gtg tat tac tgt
gcg aga ga ! 3-64# 31
aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg tat ctt
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt
gcg aga ga ! 3-66# 32
aga ttc acc atc tca aga gat gat tca aag aac tca ctg tat ctg
caa atg aac agc ctg aaa acc gag gac acg gcc gtg tat tac tgt
gct aga ga ! 3-72# 33
agg ttc acc atc tcc aga gat gat tca aag aac acg gcg tat ctg
caa atg aac agc ctg aaa acc gag gac acg gtg tat tac tgt
act aga ca ! 3-73# 34
cga ttc acc atc tcc aga gac aac gcc aag aac acg ctg tat ctg
caa atg aac agt ctg aga gcc gag gac acg gct gtg tat tac tgt
gca aga ga ! 3-74# 35
aga ttc acc atc tcc aga gac aat tcc aag aac acg ctg cat ctt
caa atg aac agc ctg aga gct gag gac acg gct gtg tat tac tgt
aag aaa ga ! 3-d# 36

! VH4

cga gtc acc ata tca gta gac aag tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg acc gcc gac acg gcc gtg tat tac tgt
gcg aga ga ! 4-04# 37
cga gtc acc atg tca gag ac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg acc gcc gtg gac acg gcc gtg tat tac tgt
gcg aga aa ! 4-28# 38
cga gtt acc ata tct gta gac acg tct aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gcg gac acg gcc gtg tat tac tgt
gcg aga ga ! 4301# 39
cga gtc acc ata tca gta gac agg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg acc gcc gcg gac acg gcc gtg tat tac tgt
gcc aga ga ! 4302# 40
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4304#41
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-31# 42
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-34# 43
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-39# 44
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-59# 45
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-61# 46
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt
gcc aga ga ! 4-b# 47
cga gtc acc ata tca gta gac acg tcc aag aac cag ttc tcc ctg
aag ctg agc tct gtg act gcc gca gac acg gcc gtg tat tac tgt

! VH5

cag gtc acc atc tca gcc gac aag tcc atc agc acc gcc tac ctg
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt
gcg aga ca! 5-51# 48

cac gtc acc atc tca gct gac aag tcc atc agc act gcc tac ctg
cag tgg agc agc ctg aag gcc tcg gac acc gcc atg tat tac tgt
gcg aga ! 5-a# 49

! VH6
cga ata acc atc aac cca gac aca tcc aag aac cag ttc tcc ctg
cag ctg aac tct gtg act ccc gag gac acg gct tat tac tgt
gca aga ga ! 6-1# 50
! VH7
cgg ttt gtc ttc tcc ttg gac acc tct gtc agc acg gca tat ctg cag atc tgc agc
cta aag gct gag gac act gcc gtg tat tac tgt
gcg aga ga ! 74.1# 51

Tabela 2: Enzimas que cortam 15 ou mais GLG humanos ou têm 5+- reconhecimento de bases em entrada típica de FR3T:

Reconhecimento de GLGid#:base#	Nome RE GLGid#:base#	#locais GLGid#:base#.....
BstEII Ggtnacc		2
1: 3	48: 3	
Existem 2 hits na base#	3	
MaeIII gtnac		36
1: 4	2: 4	3: 4
7: 4	8: 4	9: 4
37:58	38: 4	38: 58
40:58	91: 4	91: 58
43:58	44: 4	44: 58
46:58	47: 4	47: 58
Existem 24 hits na base#	**4**	
Tsp45I gtsac		33
1: 4	2: 4	3: 4
7: 4	8: 4	9: 4
37: 58	38: 4	38: 58
41: 58	42: 58	43: 4
45: 4	45: 58	46: 4
48: 4	49: 4	50: 58
Existem 21 hits na base#	4	
HphI tcacc		45

1: 5	2: 5	3: 5	4: 5	5: 5	6: 5
7: 5	8: 5	11: 5	12: 5	12: 11	13: 5
14: 5	15: 5	16: 5	17: 5	18: 5	19: 5
20: 5	21: 5	22: 5	23: 5	24: 5	25: 5
26: 5	27: 5	28: 5	29: 5	30: 5	31: 5
32: 5	33: 5	34: 5	35: 5	36: 5	37: 5

Reconhecimento de Nome RE #locais

GLGid#:base# GLGid#:base# GLGid#:base#.....

38: 5	40: 5	43: 5	44: 5	45: 5	46: 5
47: 5	48: 5	49: 5			

Existem 44 hits na base# 5

NlaIII CATG 26

1: 9	1: 42	2: 42	3: 9	3: 42	4: 9
4: 42	5: 9	5: 42	6: 42	6: 78	7: 9
7: 42	8: 21	8: 42	9: 42	10: 42	11: 42
12: 57	13: 48	13: 57	14: 57	31: 72	38: 9
48: 78	49: 78				

Existem 11 hits na base# 42

Existe 1 hit na base# 48 Pode provocar irregularidades.

BsaJI Ccnngg 37

1: 14	2: 14	5: 14	6: 14	7: 14	8: 14
8: 65	9: 14	10: 14	11: 14	12: 14	13: 14
14: 14	15: 65	17: 14	17: 65	18: 65	19: 65
20: 65	21: 65	22: 65	26: 65	29: 65	30: 65
33: 65	34: 65	35: 65	37: 65	38: 65	39: 65
40: 65	42: 65	43: 65	48: 65	49: 65	50: 65
51: 14					

Existem 23 hits na base# 65

Existem 14 hits na base# 14

AluI AGct 42

1: 47	2: 47	3: 47	4: 47	5: 47	6: 47
7: 47	8: 47	9: 47	10: 47	11: 47	16: 63
23: 63	24: 63	25: 63	31: 63	32: 63	36: 63
37: 47	37: 52	38: 47	38: 52	39: 47	39: 52
40: 47	40: 52	41: 47	41: 52	42: 47	42: 52
43: 47	43: 52	44: 47	44: 52	45: 47	45: 52
46: 47	46: 52	47: 47	47: 52	49: 15	50: 47

Existem 23 hits na base# 47

Existem 11 hits na base# 52 Apenas 5 bases de 47

BlpI GCtnagc **21**

1: 48	**2: 48**	**3: 48**	**5: 48**	**6: 48**	**7: 48**
8: 48	**9: 48**	**10: 48**	**11: 48**	**37: 48**	**38: 48**
39: 48	**40: 48**	**41: 48**	**42: 48**	**43: 48**	**44: 48**

Reconhecimento de Nome RE #locais

GLGid#:base# GLGid#:base# GLGid#:base#.....

45: 48	**46: 48**	**47: 48**			

Existem 21 hits na base# 48

MwoI GCNNNNNnngc 19

1: 48	2: 28	19: 36	22: 36	23: 36	24: 36
25: 36	26: 36	35: 36	37: 67	39: 67	40: 67
41: 67	42: 67	43: 67	44: 67	45: 67	46: 67
47: 67					

Existem 10 hits na base# 67

Existem 7 hits na base# 36

DdeI Ctnag 71

1: 49	1: 58	2: 49	2: 58	3: 49	3: 58
3: 65	4: 49	4: 58	5: 49	5: 58	5: 65
6: 49	6: 58	6: 65	7: 49	7: 58	7: 65
8: 49	8: 58	9: 49	9: 58	9: 65	10: 49
10: 58	10: 65	11: 49	11: 58	11: 65	15: 58
16: 58	16: 65	17: 58	18: 58	20: 58	21: 58
22: 58	23: 58	23: 65	24: 58	24: 65	25: 58
25: 65	26: 58	27: 58	27: 65	28: 58	30: 58
31: 58	31: 65	32: 58	32: 65	35: 58	36: 58
36: 65	37: 49	38: 49	39: 26	39: 49	40: 49
41: 49	42: 26	42: 49	43: 49	44: 49	45: 49
46: 49	47: 49	48: 12	49 12	51: 65	

Existem 29 hits na base# 58

Existem 22 hits na base# 49 Apenas nove bases de 58

Existem 16 hits na base# 65 Apenas sete bases de 58

BglII Agatct **11**

1: 61	2: 61	3: 61	4: 61	5: 61	6: 61
7: 61	9: 61	10: 61	11: 61	51: 47	

Existem 10 hits na base# 61

BstYI Rgatcy 12

1: 61	2: 61	3: 61	4: 61	5: 61	6: 61
7: 61	8: 61	9: 61	10: 61	11: 61	51: 47

Existem 11 hits na base# 61

Hpy188I TCNga 17

Reconhecimento de Nome RE #locais

GLGid#:base#	GLGid#:base#		GLGid#:base#.....		
1: 64	2: 64	3: 64	9: 64	5: 64	6: 64
7: 64	8: 64	9: 64	10: 64	11: 64	16: 57
20: 57	27: 57	35: 57	48: 67	49: 67	

Existem 11 hits na base# 64

Existem 4 hits na base# 57

Existem 2 hits na base# 67 Podem ser ireegulares.

Ms1I CAYNNnnRTG 44

1: 72	2: 72	3: 72	4: 72	5: 72	6: 72
7: 72	8: 72	9: 72	10: 72	11: 72	15: 72
17: 72	18: 72	19: 72	21: 72	23: 72	24: 72
25: 72	26: 72	28: 72	29: 72	30: 72	31: 72
32: 72	33: 72	34: 72	35: 72	36: 72	37: 72
38: 72	39: 72	40: 72	41: 72	42: 72	43: 72
44: 72	45: 72	46: 72	47: 72	48: 72	49: 72
50: 72	51: 72				

Existem 44 hits na base# 72

BsiEI CGRYcg 23

1: 74	3: 74	4: 74	5: 74	7: 74	8: 74
9: 74	10: 74	11: 74	17: 74	22: 74	30: 74
33: 74	34: 74	37: 74	38: 74	39: 74	40: 74
41: 74	42: 74	45: 74	46: 74	47: 74	

Existem 23 hits na base# 74

EaeI Yggccr 23

1: 74	3: 74	4: 74	5: 74	7: 74	8: 74
9: 74	10: 74	11: 74	17: 74	22: 74	30: 74
33: 74	34: 74	37: 74	38: 74	39: 74	40: 74
41: 74	42: 74	45: 74	46: 74	47: 74	

Existem 23 hits na base# 74

EagI Cggccg **23**

1: 74	**3: 74**	**4: 74**	**5: 74**	**7: 74**	**8: 74**
9: 74	**10: 74**	**11: 74**	**17: 74**	**22: 74**	**30: 74**
33: 74	**34: 74**	**37: 74**	**38: 74**	**39: 74**	**40: 74**

Reconhecimento de Nome RE #locais

GLGid#:base# GLGid#:base# GLGid#:base#.....

41: 74	**42: 74**	**45: 74**	**46: 74**	**47: 74**

Existem 23 hits na base# 74

HaeIII GGcc 27

1: 75	3: 75	4: 75	5: 75	7: 75	8: 75
9: 75	10: 75	11: 75	16: 75	17: 75	20: 75
22: 75	30: 75	33: 75	34: 75	37: 75	38: 75
39: 75	40: 75	41: 75	42: 75	45: 75	46: 75
47: 75	48: 63	49: 63			

Existem 25 hits na base# 75

Bst4CI ACNgt 65°C **63 Locais Existe um terceiro isosquizoerismo**

1: 86	**2: 86**	**3: 86**	**4: 86**	**5: 86**	**6: 86**
7: 34	**7: 86**	**8: 86**	**9: 86**	**10: 86**	**11: 86**
12: 86	**13: 86**	**14: 86**	**15: 36**	**15: 86**	**16: 53**
16: 86	**17: 36**	**17: 86**	**18: 86**	**19: 86**	**20: 53**
20: 86	**21: 36**	**21: 86**	**22: 0**	**22: 86**	**23: 86**
24: 86	**25: 86**	**26: 86**	**27: 53**	**27: 86**	**28: 36**
28: 86	**29: 86**	**30: 86**	**31: 86**	**32: 86**	**33: 36**
33: 86	**34: 86**	**35: 53**	**35: 86**	**36: 86**	**37: 86**
38: 86	**39: 86**	**40: 86**	**41: 86**	**42: 86**	**43: 86**
44: 86	**45: 86**	**46: 86**	**47: 86**	**48: 86**	**49: 86**
50: 86	**51: 0**	**51: 86**			

Existem 51 hits na base# 86 Todos os outros locais estão bem longe.

HpyCH4III ACNgt **63**

1: 86	**2: 86**	**3: 86**	**4: 86**	**5: 86**	**6: 86**
7: 34	**7: 86**	**8: 86**	**9: 86**	**10: 86**	**11: 86**
12: 86	**13: 86**	**14: 86**	**15: 36**	**15: 86**	**16: 53**
16: 86	**17: 36**	**17: 86**	**18: 86**	**19: 86**	**20: 53**
20: 86	**21: 36**	**21: 86**	**22: 0**	**22: 86**	**23: 86**
24: 86	**25: 86**	**26: 86**	**27: 53**	**27: 86**	**28: 36**

28: 86 29: 86 30: 86 31: 86 32: 86 33: 36
33: 86 34: 86 35: 53 35: 86 36: 86 37: 86
38: 86 39: 86 40: 86 41: 86 42: 86 43: 86
44: 86 45: 86 46: 86 47: 86 48: 86 49: 86
50: 86 51: 0 51: 86

Existem 51 hits na base# 86

Reconhecimento de Nome RE #locais
GLGid#:base# GLGid#:base# GLGid#:base#.....

HinfI Gantc 43
2: 2 3: 2 4: 2 5: 2 6: 2 7: 2
8: 2 9: 2 9: 22 10: 2 11: 2 15: 2
16: 2 17: 2 18: 2 19: 2 19: 22 20: 2
21: 2 23: 2 24: 2 25: 2 26: 2 27: 2
28: 2 29: 2 30: 2 31: 2 32: 2 33: 2
33: 22 34: 22 35: 2 36: 2 37: 2 38: 2
40: 2 43: 2 44: 2 45: 2 46: 2 47: 2
50: 60

Existem 38 hits na base# 2

MlyI GAGTCNNNNNn 18
2: 2 3: 2 4: 2 5: 2 6: 2 7: 2
8: 2 9: 2 10: 2 11: 2 37: 2 38: 2
40: 2 43: 2 44: 2 45: 2 46: 2 47: 2

Existem 18 hits na base# 2

PleI gagtc 18
2: 2 3: 2 4: 2 5: 2 6: 2 7: 2
8: 2 9: 2 10: 2 11: 2 37: 2 38: 2
40: 2 43: 2 44: 2 45: 2 46: 2 47: 2

Existem 18 hits na base# 2

AciI Ccgc 24
2: 26 9: 14 10: 14 11: 14 27: 74 37: 62
37: 65 38: 62 39: 65 40: 62 40: 65 41: 65
42: 65 43: 62 43: 65 44: 62 44: 65 45: 62
46: 62 47: 62 47: 65 48: 35 48: 74 49: 74

Existem 8 hits na base# 62

Existem 8 hits na base# 65

Existem 3 hits na base# 14

Existem 3 hits na base# 74

```
Existe 1 hit na base# 26
Existe 1 hit na base# 35
-"- Gcgg 11
    8: 91      9: 16      10: 16     11: 16     37: 67     39: 67
    40: 67     42: 67     43: 67     45: 67     46: 67
```
Existem 7 hits na base# 67

```
Reconhecimento de Nome RE                          #locais
GLGid#:base#      GLGid#:base#                  GLGid#:base#.....
```

```
Existem 3 hits na base# 16
Existe 1 hit na base# 91
BsiHKAI GWGCWc                        20
    2: 30      4: 30      6: 30      7: 30      9: 30      10: 30
    12: 89     13: 89     14: 89     37: 51     38: 51     39: 51
    40: 51     41: 51     42: 51     43: 51     44: 51     45: 51
    46: 51     47: 51
Existem 11 hits na base# 51

Bsp1286I GDGCHc                       20
    2: 30      4: 30      6: 30      7: 30      9: 30      10: 30
    12: 89     13: 89     14: 89     37: 51     38: 51     39: 51
    40: 51     41: 51     42: 51     43: 51     44: 51     45: 51
    46: 51     47: 51
Existem 11 hits na base# 51

HgiAI GWGCWc                          20
    2: 30      4: 30      6: 30      7: 30      9: 30      10: 30
    12: 89     13: 89     14: 89     37: 51     38: 51     39: 51
    40: 51     41: 51     42: 51     43: 51     44: 51     45: 51
    46: 51     47: 51
Existem 11 hits na base# 51

BsoFI GCngc                           26
    2: 53      3: 53      5: 53      6: 53      7: 53      8: 53
    8: 91      9: 53      10: 53     11: 53     31: 53     36: 36
    37: 64     39: 64     40: 64     41: 64     42: 64     43: 64
    44: 64     45: 64     46: 64     47: 64     48: 53     49: 53
    50: 45     51: 53
Existem 13 hits na base# 53
Existem 10 hits na base# 64
```

TseI Gcwgc 17

2: 53 3: 53 5: 53 6: 53 7: 53 8: 53

9: 53 10: 53 11: 53 31: 53 36: 36 45: 64

46: 64 48: 53 49: 53 50: 45 51: 53

Existem 13 hits na base# 53

MnlI gagg 34

3: 67 3: 95 4: 51 5: 16 5: 67 6: 67

Reconhecimento de Nome RE #locais

GLGid#:base# GLGid#:base# GLGid#:base#.....

7: 67 8: 67 9: 67 10: 67 11: 67 15: 67

16: 67 17: 67 19: 67 20: 67 21: 67 22: 67

23: 67 24: 67 25: 67 26: 67 27: 67 28: 67

29: 67 30: 67 31: 67 32: 67 33: 67 34: 67

35: 67 36: 67 50: 67 51: 67

Existem 31 hits na base# 67

HpyCH4V TGca 34

5: 90 6: 90 11: 90 12: 90 13: 90 14: 90

15: 44 16: 44 16: 90 17: 44 18: 90 19: 44

20: 44 21: 44 22: 44 23: 44 24: 44 25: 44

26: 44 27: 44 27: 90 28: 44 29: 44 33: 44

34: 44 35: 44 35: 90 36: 38 48: 44 49: 44

50: 44 50: 90 51: 44 51: 52

Existem 21 hits na base# 44

Existe 1 hit na base# 52

AccI GTmkac 13 reconhecimento de base 5

7: 37 11: 24 37: 16 38: 16 39: 16 40: 16

41: 16 42: 16 43: 16 44: 16 45: 16 46: 16

47: 16

Existem 11 hits na base# 16

SacII CCGCgg 8 reconhecimento de base 6

9: 14 10: 14 11: 14 37: 65 39: 65 40: 65

42: 65 43: 65

Existem 5 hits na base# 65

Existem 3 hits na base# 14

```
TfiI Gawtc                              24
    9: 22       15: 2       16: 2       17: 2       18: 2       19: 2
    19: 22      20: 2       21: 2       23: 2       24: 2       25: 2
    26: 2       27: 2       28: 2       29: 2       30: 2       31: 2
    32: 2       33: 2       33: 22      34: 22      35: 2       36: 2
Existem 20 hits na base# 2
BsmAI Nnnnnngagac                       19
    15: 11      16: 11      20: 11      21: 11      22: 11      23: 11
    24: 11      25: 11      26: 11      27: 11      28: 11      28: 56
Reconhecimento de Nome RE                           #locais
GLGid#:base#        GLGid#:base#                    GLGid#:base#.....
    30: 11      31: 11      32: 11      35: 11      36: 11      44: 87
    48: 87
Existem 16 hits na base# 11

BpmI ctccag                             19
    15: 12      16: 12      17: 12      18: 12      20: 12      21: 12
    22: 12      23: 12      24: 12      25: 12      26: 12      27: 12
    28: 12      30: 12      31: 12      32: 12      34: 12      35: 12
    36: 12
Existem 19 hits na base# 12

XmnI GAANNnnttc                         12
    37: 30      38: 30      39: 30      40: 30      41: 30      42: 30
    43: 30      44: 30      45: 30      46: 30      47: 30      50: 30
Existem 12 hits na base# 30

BsrI NCcagt                             12
    37: 32      38: 32      39: 32      40: 32      41: 32      42: 32
    43: 32      44: 32      45: 32      46: 32      47: 32      50: 32
Existem 12 hits na base# 32

BanII GRGCYc                            11
    37: 51      38: 51      39: 51      40: 51      41: 51      42: 51
    43: 51      44: 51      45: 51      46: 51      47: 51
Existem 11 hits na base# 51

Ecl136I GAGctc                          11
    37: 51      38: 51      39: 51      40: 51      41: 51      42: 51
    43: 51      44: 51      45: 51      46: 51      47: 51
Existem 11 hits na base# 51
```

```
SacI GAGCTc                          11
    37: 51     38: 51     39: 51     40: 51     41: 51     42: 51
    43: 51     44: 51     45: 51     46: 51     47: 51
Existem 11 hits na base# 51
```

Tabela 3: FR3 3-23 sintético de cadeias pesadas humanas mostrando posições de possíveis locais de clivagem

```
! Os locais contruídos no gene sintético são mostrados no ADN, em letra maiúscula,
! com o nome RE entre barras verticais (como em XbaI ).
! Os RERS frequentemente encontrados em GLG são mostrados abaixo da sequência sintética
! com o nome à direita (como em gtn ac=MaeIII(24), indicando que 24 dos 51 GLG contêm o local).
!
!                                                                    |---FR3---
!                                                                       89  90 (codão # em
!                                                                       R   F
3-23 sintético)
                                                                    |cgc|ttc|      6
! ADN permitido                                                     |cgn|tty|
!                                                                   |agr|
!                                                                       ga ntc =
HinfI (38)
!                                                                       ga gtc =
PleI (18)
!                                                                       ga wtc =
TfiI (20)
!                                                                              gtn ac =
MaeIII (24)
!                                                                              gts ac =
Tsp4SI (21)
!                                                                              tc acc =
HphI (44)
!
!          --------FR3-------------------------------------------------------------
!              91  92  93  94  95  96  97  98  99  100  101  102  103  104  105
!              T   I   S   R   D   N   S   K   N   T    L    Y    L    Q    M
            |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|     51
! permitido |acn|ath|tcn|cgn|gay|aay|tcn|aar|aay|acn|ttr|tay|ttr|car|atg|
!                   |agy|agr|       |agy|           |ctn|   |ctn|
!                   |    ga|gac = BsmAI (16)                        ag ct =
AluI(23)
```

```
!                  c|tcc ag = BpI(19)                                  g ctn agc =
BlpI (21)

!                      |          |              g aan nnn ttc = XmnI(12)
!                      | XbaI     |                          tg ca = HpyCH4V(21)
!
!          --------FR3-------------------------------------------------------------
!              106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!               N   S   L   R   A   E   D   T   A   V   Y   Y C   A   K
            |aac|agC|TTA|AGc|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|     96
! permitido |aay|tcn|ttr|cgn|gcn|gar|gay|acn|gcn|gtn|tay|tay|tgy|gcn|aar|
!               |agy|ctn|agr|           |     |
!                 |          |  cc nng g = BsaJi(23)      ac ngt = Bst4CI(51)
!                 |      aga tct = BglII(10)   |          ac ngt = HpyCH4III(51)
!                 |      Rga tcY = BstYI(11)   |           ac ngt = TaaI(51)
!                 |          |          c ayn nnn rtc = MslI(44)
!                 |          |             cg ryc g = BsiEI(23)
!                 |          |             yg gcc r = EaeI(23)
!                 |          |             cg gcc g = EagI(23)
!                 |          |             |g gcc = HaeIII(25)
!                 |          |      gag g = MnlI(31) |
!                 |  AflII   |             |  PstI   |
```

Tabela 4: REdaptors, Extensores e Pontes utilizados para Clivagem e Captura de Cadeias Pesadas Humanas em FR3.

A: HpyCH4V Sondas de genes HC humba reais

!HpyCH4V em FR3 de HC humano, bases 35-56; apenas aqueles com local TGca TGca;10,

reconhecimento RE:tgca de comprimento 4 é esperado em 10

1	6-1	agttctccctgcagctgaactc
2	3-11,3-07,3-21,3-72,3-48	cactgtatctgcaaatgaacag
3	3-09,3-43,3-20	ccctgtatctgcaaatgaacag
4	5-51	ccgcctacctgcagtggagcag
5	3-15,3-30,3-30.5,3-30.3,3-74,3-23,3-33	cgctgtatctgcaaatgaacag
6	7-4.1	cggcatatctgcagatctgcag
7	3-73	cggcgtatctgcaaatgaacag
8	5-a	ctgcctacctgcagtggagcag
9	3-49	tcgcctatctgcaaatgaacag

B: HpyCH4v REdaptors, Extensores e Pontes

B.1 REdaptor

! corte de cadeia inferior de HC:

! TmKeller para NaCl 100 mM, zero formamida

! REdapters para clivagem		T_m^W	T_m^K
(ON_HCFR36-1)	5'-agttctcccTGCAgctgaactc-3'	68.0	64.5
(ON_HCFR36-1A)	5'-TTCTCCCtgcaGCTGAACTC-3'	62.0	62.5
(ON_HCFR36-1B)	5'-ttctcccTGCAgctgaac-3'	56.0	59.9
(ON_HCFR33-15	5'-cgctgtatcTGCAaatgaacag-3'	64.0	60.8
(ON_HCFR33-15A	5'-ctgtatcTGCAaatgaacag-3'	56.0	56.3
(ON_HCFR33-15B)	5'-cactgtacTGCAaatgaacag-3'	62.0	58.9
(ON_HCFR35-51)	5'-ccgcctaccTGCAgtggagcag-3'	74.0	70.1

!

B.2 Segmento de gene 3-23 sintético no qual é para clonar o CDR3 capturado

```
!                    XbaI…
!D323*     cgCttcacTaag tcT aga gac aaC tcT aag aaT acT ctC taC
!          scab........ gene 3-23 concebido.................
```

```
!       HpyCH4V
!        .. ..            AflII...
!       Ttg caG atg aac agc TtA agG ...
!       ..............................
!
```

B.3 Extensor e Pontes

```
! Extensor (cadeia de fundo):
!
(ON_HCHpyEx01)    5'-cAAgTAgAgAgTATTcTTAgAgTTgTcTcTAgAcTTAgTgAAgcg-3'
! ON_HCHpyEx01 é o complemento reverso de
! 5'-cgCttcacTaag tcT aga gac aaC aag aaT acT ctC taC Ttg-3'
!
! Pontes (cadeia de topo, base 9 sobreposta):
(ON_HCHpyBr016-1) 5'-cgCttcacTaag tcT aga gac aaC tcT aag-
              aaT acT ctC taC Ttg CAgctgaac-3' { term C 3' está bloqueado}
!
! 3-15 et al. + 3-11
(ON_HCHpyBr023-15)    5'-chCttcacTaag tcT aga gac aac tcT aag-
              aaT acT ctC taC Ttg CAaatgaac-3' { term C 3' está bloqueado}
!
! 5-51
(ON_HCHpyBr045-51)    5'-cgCttcacTaag tcT aga gac aaC tcT aag-
              aaT acT ctC taC Ttg CAgtggagc-3' { term C 3' está bloqueado}
!
! Iniciador de PCR (cadeia de topo)
!
(ON_HCHpyPCR)         5'-chCttcacTaag tcT aga gac-3'
!
```

C: *BlpI* Sondas de GLG HC humano

```
    1          1-58, 1-03,1-08,1-69,1-24,1-45,1-46,1-f,1-e
acatggaGCTGAGCaggctgag
    2                                                  1-02
acatggaGCTGAGCaggctgag
    3                                                  1-18
acatggagctgaggagcctgag
    4                                                  5-51,5-a
acctgcagtggagcagcctgaa
    5                                        3-15,3-73, 3-49,3-72
atctgcaaatgaacagcctgaa
    6              3303,3-33,3-07,3-11,3-30,3-21,3-23,3305,3-48
```

atctgcaaatgaacagcctgag

7 3-20,3-74,3-09,3-43

atctgcaaatgaacagtctgag

8 74.1

atctgcagatctgcagcctaaa

9 3-66,3-13,3-53,3-d

atcttcaaatgaacagcctgag

10 3-64

atcttcaaatgggcagcctgag

11 4301,4-28,4302,4-04,4304,4-31,4-31,4-34,4-39,4-59,4-61,4-b

ccctgaaGCTGAGCtctgtgac

12 6-1

ccctgcagctgaactctgtgac

13 2-70,2-05

tccttacaatgaccaacatgga

14 2-26

tccttaccatgaccaacatgga

D: *BlpI* REdaptors, Extensores e Pontes

D.1 REdaptors

		T_m^W	T_m^K
(BlpF3HC1-58)	5'-ac atg gaG CTG AGC agc ct gag-3'	70	66.4
(BlpF3HC6-1)	5'-cc ctg aag ctg agc tct gtg ac-3'	70	66.4

! BlpF3HC6-1 combinações 4-30.1, not 6-1.

D.2 Segmento de gene 3-23 sintético no qual é para clonar o CDR3 capturado

!

BlpI

! XbaI...

... ...

!D323* cgCttcacTaag TCT AGA gac aaC tcT aag aaT acT ctC taC Ttg caG atg aac

!

! AflII...

! agC TTA AGG

D.3 Extensor e Pontes

! Pontes

(BlpF3Br1) 5'-chCttcacTcag tcT aga gaT aaC AGT aaA aaT acT TtG-

```
                    taC Ttg caG Ctg a|GC agc ctg-3'
(BlpF3Br2)     5'-cgCttcacTcag tcT aga gaT aaC AGT aaA aaT TtG-
                    taC Ttg caG Ctg a|gc tct gtg-3'
!                                    | a cadeia inferior é cortada aqui
! Extensor
(BlpF3Ext)     5'-TcAgcTgcAAgTAcAAAgTATTTTTAcTgTTATcTcTAgA cTgAgTgAAgcg-3'
! BlpF3Ext é o complemento reverso de :
! 5'-cgCttcacTcag tcT aga gaT aaC AGT aaA aaT TtG taC Ttg caG Ctg a-3'
!
(BlpF3PCR)     5'-cgCttcacTcag tct aga gaT aaC-3'
```

E: ***HpyCH4III*** local circundante de sequências GLG distintas, bases 77-98

```
    1      102#1,118#4,146#7,169#9,1e#10,311#17,353#30,404#37,4301
ccgtgtattactgtgcgagaga
    2      103#2,307#15#,321#21,3303#24,333#26,348#28,364#31,366#32
ctgtgtattactgtgcgagaga
    3                                                      108#3
ccgtgttattactgtgcgagagg
    4                                                124#5,1f#11
ccgtgtattactgtgcaacaga
    5                                                   145#6
ccatgtattactgtgcaagata
    6                                                   158#8
ccgtgtattactgtgcggcaga
    7                                                      205#12
ccacatattactgtgcacacag
    8                                                   226#13
ccacatattactgtgcacggat
    9                                                   270#14
ccacgtattactgtgcacggat
    10                                            309#16,343#27
ccttgtattactgtgcaaaaga
    11                                          313#18,374#35,61#50
ctgtgtattactgtgcaagaga
    12                                                  315#19
ccgtgtattactgtaccacaga
    13                                                  320#20
ccttgtatcactgtgcgagaga
    14                                                  323#22
ccgtatattactgtgcgaaaga
    15                                            330#23,3305#25
```

ctgtgtattactgtgcgaaaga

16 349#29

ccgtgtattactgtactagaga

17 372#33

ccgtgtattactgtgctagaga

18 373#34

ccgtgtattactgtactagaca

19 3d#36

ctgtgtattactgtaagaaaga

20 428#38

ccgtgtattactgtgcgagaaa

21 4302#40,4304#41

ccgtgtattactgtgccagaga

22 439#44

ctgtgtattactgtgcgagaca

23 551#48

ccatgtattactgtgcgagaca

24 5a#49

ccatgtattactgtgcgaga

F: *Hpy*CH4III REdaptors, Extensores e Pontes

F.1 REdaptors

! ON para clivagem de HC (inferior) em FR3 (bases 77-97)

! Para clivagem com HpyCH4III, Bst4CI, ou TaaI

! A clivagem é na cadeia inferior antes da base 88.

! 77 788 888 888 889 999 999 9

! 78 901 235 567 897 123 456 7 T_m^W T_m^K

(H43.77.97.1-02#1) 5'-cc gtg tat tAC TGT gcg aga g-3' 6462.6

(H43.77.97.1-03#2) 5'-ct gtg tat tAC TGT gcg aga g-3' 6260.6

(H43.77.97.108#3) 5'-cc gtg tat tAC TGT gcg aga g-3' 6462.6

(H43.77.97.323#22) 5'-cc gta tat tac tgt gcg aaa g-3' 6058.7

(H43.77.97.330#23) 5'-cc gtg tat tac tgt gcg aaa g-3' 6058.7

(H43.77.97.439#44) 5'-cc gtg tat tac tgt gcg aga c-3' 6260.6

(H43.77.97.551#48) 5'-cc atg tat tac tgt gcg aga c-3' 6260.6

(H43.77.97.5a#49) 5'-cc atg tat tAC TGT gcg aga c-3' 5858.3

F.2 Extensor e Pontes

! locais XbaI e AflII em pontes são tapados

(H43.XABr1) 5'-ggtgtagtga-

|TCT|AGt|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-

|aac|agC|TTt|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat tgt gcg aga-3'

(H43.XABr2) 5'-ggtgtagtga-

|TCT|AGt|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-

|aac|agC|TTt|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat tgt gcg aaa-3'

(H43.XAExt) 5'-ATAgTAgAcT gcAgTgTccT cAgcccTTAA gcTgTTcATc TgcAAgTAgA

gAgTATTcTT AgAgTTgTcT cTAgATcAct AcAcc-3'

!H43.XAExt é o complemento reverso de

! 5'-ggtgtagtga-

! |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat -3'

(H43.XAPCR) 5'-ggtgtagtga |TCT|AGA|gac|aac-3'

! locais XbaI e AflII em pontes são tapados

(H43.Abr1) 5'-ggtgtagtga-

|aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat tgt gcg aga-3'

(H43.Abr2) 5'-ggtgtagtga-

|aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat tgt gcg aaa-3'

(H43.XAExt) 5'-ATAgTAgAcTgcAgTgTccTcAgcccTTTAAgcTgTTTcAcTAcAcc-3'

! (H43.XAExt) é o complemento reverso de 5'-ggtgtagtga-

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat -3'

(H43.APCR) 5'-ggtgtagtga |aac|agC|TTA|AGg|gct|g-3'

Tabela 5: Análise da frequência de combinação de REdaptors em genes V reais

A: HpyCH4V em HC em bases 35-56

		Número de emparelhamento incorrectos..... Número													
Id	t	0	1	2	3	4	5	6	7	8	9	10	Corte	Id	Sonda
1	510	5	11	274	92	61	25	21	1	1	3	5	443	6-1	agttctcccTGCAgctgaactc
2	192	54	42	32	24	15	2	3	10	3	1	6	167	3-11	cactgtatcTGCAaatgaacag
3	58	19	7	17	6	5	1	0	1	0	2	0	54	3-09	ccctgtatcTGCAaatgaacag
4	267	42	33	9	8	82	83	42	2	8	11	1	100	5-51	ccgcctaccTGCAgtggagcag
5	250	111	59	41	24	7	5	1	0	0	2	0	242	3-15	cgctgtatcTGCAaatgaacag
6	7	0	2	0	1	0	0	0	0	0	4	0	3	7-4.1	cggcatatcTGCAgatctgcag
7	7	0	2	2	0	0	2	1	0	0	0	0	4	3-73	cggcgtatcTGCAaatgaacag
8	26	10	4	1	3	1	2	1	3	1	0	0	19	5-a	ctgcctaccTGCAgtggagcag
9	21	8	2	3	1	6	1	0	0	0	0	0	20	3-49	tcgcctatcTGCAaatgaacag
	1338	249	162	379	149	103	214	73	47	13	23	12	1052		
		249	411	790	939		1153		1200		1223				
						1042		1275		1288		1338			

Pontilhada

Sonda

Id	Sonda	sonda
6-1	agttctcccTGCAgctgaactc	agttctcccTGCAgctgaactc
3-11	cactgtatcTGCAaatgaacag	cac.g.at.....aa.....ag
3-09	ccctgtatcTGCAaatgaacag	ccc.g.at.....aa.....ag
5-51	ccgcctaccTGCAgtggagcag	ccgc..a.......tg..g.ag
3-15	cgctgtatcTGCAaatgaacag	c.c.g.at.....aa.....ag
7-4.1	cggcatatcTGCAgatctgcag	c.gca.at......a.ctg.ag
3-73	cggcgtatcTGCAaatgaacag	c.gcg.at.....aa.....ag
5-a	ctgcctaccTGCAgtggagcag	ctgc..a.......tg..g.ag
3-49	tcgcctatcTGCAaatgaacag	tcgc..at.....aa.....ag

Seqs apenas com o local RE esperado.....1004

(contam apenas casos com 4 ou menos emparelhamentos incorrectos)

Seqs apenas com um local inesperado..... 0

Seqs com esperado e inesperado.... 48

(contam apenas casos com 4 ou menos emparelhamentos incorrectos)

Seqs sem locais..... 0

B: Blpl em HC

Id	Ntot	0	1	2	3	4	5	6	7	8	N corte	Nome	
1	133	73	16	11	13	6	9	1	4	0	119	1-58	acatggaGCTGAGCagcctgag
2	14	11	1	0	0	0	0	1	0	1	12	1-02	acatggagctgagcaggctgag
3	34	17	8	2	6	1	0	0	0	0	0	1-18	acatggagctgaggagcctgag
4	120	50	32	16	10	9	1	1	1	0	2	5-51	acctgcagtggagcagcctgaa
5	55	13	11	10	17	3	1	0	0	0	0	3-15	atctgcaaatgaacagcctgaa
6	340	186	88	41	15	6	3	0	1	0	0	3303	atctgcaaatgaacagcctgag
7	82	25	16	25	12	1	3	0	0	0	0	3-20	atctgcaaatgaacagtctgag

8	3	0	2	0	1	0	0	0	0	0	0	74.1	atctgcagatctgcagcctaaa
9	23	18	2	2	1	0	0	0	0	0	0	3-66	atcttcaaatgaacagcctgag
10	2	1	0	1	0	0	0	0	0	0	0	3-64	atcttcaaatgggcagcctgag
11	486	249	78	81	38	21	10	4	4	1	467	4301	ccctgaagctgagctctgtgac
12	16	6	3	1	0	1	1	3	1	0	1	6-1	ccctgcagctgaactctgtgac
13	28	15	8	2	2	1	0	0	0	0	0	2-70	tccttacaatgaccaacatgga
14	2	0	2	0	0	0	0	0	0	0	0	2-26	tccttaccatgaccaacatgga

Nome	Sequência completa	Modo com pontos
1-58	acatggaGCTGAGCagcc tgag	acatggaGCTGAGCagcctgag
1-02	acatggagctgagcaggctga g	.....g.....
1-18	acatggagctgaggagcctga g	.....g.....
5-51	acctgcagtggagcagcctga a	..c..c..tg.....a
3-15	atctgcaaatgaacagcctgaa	.tc..c.aa...a.....a
3-30.3	atctgcaaatgaacagcctgag	.tc..c.aa...a.....
3-20	atctgcaaatgaacagtctgag	.tc..c.aa...a...t.....
7-4.1	atctgcagatctgcagcctaaa	.tc..c..a.ct.....a.a
3-66	atcttcaaatgaacagcctgag	.tc.tc.aa...a.....
3-64	atcttcaaatgggcagcctgag	.tc.tc.aa..g.....
4-30.1	ccctgaagctgagctctgtgac	c.c..a.....tctg...c
6-1	ccctgcagctgaactctgtgac	c.c..c.....a.tctg...c
2-70	tccttacaatgaccaacatgga	t.c.tacaa...c..a.a..ga
2-26	tccttaccatgaccaacatgga	t.c.tacca...c..a.a..ga

Seqs apenas com o local RE esperado..... 597 (contagem de sequências com 4 ou menos emparelhamentos incorrectos)

Seqs apenas com um local inesperado..... 2

Seqs com esperado e inesperado.... 2

Seqs sem locais..... 686

C: HpyCH4III, Bst4CI ou TaaI em HC

Ao classificar se o local RE de interesse está presente, contam-se apenas os ON que possuem 4 ou menos emparelhamentos incorrectos

Número de sequências..... 1617

Id	Nto t	0	1	2	3	4	5	6	7	8	NCUt	acngt	acngt
1	244	78	92	43	18	10	1	2	0	0	241	102#1,1 ccgtgtattACTGTgcgaga ga	ccgtgtattactgtgcgagaga
2	457	69	150	115	66	41	1	8	3	1	43	103#2,3 ctgtgtattactgtgcgagaga	.t.....
3	173	52	45	36	22	14	3	0	0	1	169	108#3 ccgtgtattactgtgcgagagg	.....g
4	16	0	3	2	2	1	6	0	1	1	8	124#5,1 ccgtgtattactgtgcaacaga	.....a.c...
5	4	0	0	1	0	1	1	0	1	0	2	145#6 ccatgtattactgtgcaagata	..a.....a...t.
6	15	1	0	1	0	6	4	1	1	1	8	158#8 ccgtgtattactgtgcggcaga	.....gc...
7	23	4	8	5	2	2	1	1	0	0	212	205#1 ccacatattactgtgcacacag	..aca.....acacag
8	9	1	1	1	0	3	2	1	0	0	6	226#1 ccacatattactgtgcacggat	..aca.....ac.gat

											3		
9	7	1	3	1	1	0	0	1	0	0	6	270#1	ccacgtattactgtgcacggat ..ac.....ac.gat
											4		
10	23	7	3	5	5	2	1	0	0	0	2	309#1	ccttgtattactgtgcaaaaga ..t.....a.a...
											2	6,	
11	35	5	10	7	6	3	3	0	1	0	3	313#1	ctgtgtattactgtgcaagaga .t.....a.....
											1	8,	
12	18	2	3	2	2	6	1	0	2	0	1	315#1	ccgtgtattactgtaccacaga .....a.c.c...
											5	9	
13	3	1	2	0	0	0	0	0	0	0	3	320#2	ccttgtatcactgtgcgagaga ..t.....c.....
											0		
14	117	29	23	28	2	8	4	2	1	0	1	323#2	ccgtatattactgtgcgaaaga ....a.....a...
					2						1	2	
											0		
15	75	21	25	13	9	1	4	2	0	0	6	330#2	ctgtgtattactgtgcgaaaga .t.....a...
											9	3,	
16	14	2	2	2	3	0	3	1	1	0	9	349#2	ccgtgtattactgtactagaga .....a.t.....
												9	
17	2	0	0	1	0	0	1	0	0	0	1	372#3	ccgtgtattactgtgctagaga .....t.....
												3	
18	1	0	0	1	0	0	0	0	0	0	1	373#3	ccgtgtattactgtactagaca .....a.t...c.
												4	
19	2	0	0	0	0	0	0	0	0	2	0	3d#36	ctgtgtattactgtaagaaaga .t.....aa..a...
20	34	4	9	9	4	5	3	0	0	0	3	428#3	ccgtgtattactgtgcgagaaa .....a.
											1	8	
21	17	5	4	2	2	3	1	0	0	0	1	4302#	ccgtgtattactgtgccagaga .....c.....
											6	40	
22	75	15	17	24	7	1	1	1	0	0	7	439#4	ctgtgtattactgtgcgagaca .t.....c.

3 4

23 40 14 15 4 5 1 0 1 0 0 3 551#4 ccatgtattactgtgcgagaca ..a.................c.

9 8

24 213 26 56 60 4 2 7 2 0 0 2 5a#49 ccatgtattactgtgcgagaA ..a.................AA

2 0 0 A

4

Grupo 337 471 363 218 130 58 23 11 6

Cumula 337 808 117 1389 1519 1577 1600 1611 1617

tivo

Seqs apenas com o local RE esperado..... 1

5

1

1

Seqs apenas com um local inesperado..... 0

Seqs com esperado e inesperado.... 8

Seqs sem locais......................... 0

Tabela 5D:

Análise repetida utilizando apenas 8 REdaptors best

Id	Ntot	0	1	2	3	4	5	6	7	8+		
1	301	78	101	54	32	16	9	10	1	0	281	102#1
ccgtgtattactgtgcgagaga												
2	493	69	155	125	73	37	14	11	3	6	459	103#2
ctgtgtattactgtgcgagaga												
3	189	52	45	38	23	18	5	4	1	3	176	108#3
ccgtgtattactgtgcgagagg												
4	127	29	23	28	24	10	6	5	2	0	114	323#22
ccgtatattactgtgcgaaaga												
5	78	21	25	14	11	1	4	2	0	0	72	330#23
ctgtgtattactgtgcgaaaga						6	79	15	17	25	8	11 1 2 0 0 76
439#44	ctgtgtattactgtgcgagaca											
7	43	14	15	5	5	3	0	1	0	0	42	551#48
ccatgtattactgtgcgagaca												
8	307	26	63	72	51	38	24	14	13	6	250	5a#49
ccatgtattactgtgcgaga												

1	102#1	ccgtgtattactgtgcgagaga	ccgtgtattactgtgcgagaga
2	103#2	ctgtgtattactgtgcgagaga	.t..................
3	108#3	ccgtgtattactgtgcgagagg	..................g
4	323#22	ccgtatattactgtgcgaaaga	....a.....a...
5	330#23	ctgtgtattactgtgcgaaaga	.t.....a...
6	439#44	ctgtgtattactgtgcgagaca	.t.....c.
7	551#48	ccatgtattactgtgcgagaca	..a.....c.
8	5a#49	ccatgtattactgtgcgagaAA	..a.....AA

Seqs apenas com o local RE esperado.....1463 / 1617

Seqs apenas com um local inesperado.....0

Seqs com esperado e inesperado.....7

Seqs sem locais.....0

Tabela 6: Sequências FR1 GLG HC Humanas

Exão VH- alinhamento de sequência nucleotídica

VH1

1-02 CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA AAG GTC TCC TGC AAG GCT TCT GGA TAC ACC TTC ACC

1-03 cag gtC cag ctT gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gct tct gga tac acc ttc acT

1-08 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct gga tac acc ttc acc

1-18 cag gtT cag ctg gtg cag tct ggA gct gag gtg aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gct tct ggt tac acc ttT acc

1-24 cag gtC cag ctg gta cag tct ggg gct gag aag aag cct ggg gcc tca gtg aag gtc tcc tgc aag gTt tcC gga tac acc Ctc acT

1-45 cag Atg cag ctg gtg cag tct ggg gct gag gtg aag aag Act ggg Tcc tca gtg aag gtT tcc tgc aag gct tcC gga tac acc ttc acc

1-46 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg gcc tca gtg aag gtT tcc tgc aag gcA tct gga tac acc ttc acc

1-58 ccA Atg cag ctg gtg cag tct ggg Cct gag gtg aag aag cct ggg Acc tca gtg aag gtc tcc tgc aag gct tct gga tTc acc ttT acT

1-69 cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tcG gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc

1-e cag gtg cag ctg gtg cag tct ggg gct gag gtg aag aag cct ggg Tcc tcG gtg aag gtc tcc tgc aag gct tct gga GGc acc ttc aGc

1-f Gag gtC cag ctg gtA cag tct ggg gct gag gtg aag aag cct ggg gcT Aca gtg aaA Atc tcc tgc aag gTt tct gga tac acc ttc acc

VH2

2-05 GAG ATC ACC TTG AAG GAG TCT GGT CCT ACG CTG GTG AAA CCC ACA CAG ACC CTC ACG CTG ACC TGC ACC TTC TCT GGG TTC TCA CTC AGC

VH2

2-26 CAG Gtc acc ttg aag gag tct ggt cct GTg ctg gtg aaa ccc aca Gag acc ctc acg ctg acc tgc acc Gtc tct ggg ttc tca ctc agc

2-70 cag Gtc acc ttg aag gag tct ggt cct Gcg ctg aaa ccc aca cag acc ctc acA ctg acc tgc acc ttc tct ggg ttc tca ctc agc

VH3

3-07 GAG GTG CAG CTG GTG GAG TCT GGG GGA GGC TTG GTC CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTT AGT

3-09 gaA gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggC Agg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt GAt

3-11 Cag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc Aag cct ggA ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttc agt

3-13 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-15 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA Aag cct ggg ggg tcc ctT aga ctc tcc tgt gca gcc tct gga ttc acT ttc agt

3-20 gag gtg cag ctg gtg tct ggg gga ggT Gtg gtA cGg cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt GAt

3-21 gag gtg cag ctg gtg gag tct ggg gga ggc Ctg gtc Aag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-23 gag gtg cag ctg Tgtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt agC

3-30 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-30.3 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

VH3

3-30.5 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga crc tcc tgt gca gcc tct gga ttc acc ttC agt

3-33 Cag gtg cag ctg gtg gag tct ggg gga ggc Gtg gtc cag cct ggg Agg tcc ctg aga ctc tcc tgt gca gcG tct gga ttc acc ttC agt

3-43 gaA gtg cag ctg gtg gag tct ggg gga gTc Gtg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttt GAt

3-48 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-49 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtA cag ccA ggg Cgg tcc ctg aga ctc tcc tgt Aca gcT tct gga ttc acc ttt Ggt

3-53 gag gtg cag ctg gtg gag Act ggA gga ggc ttg Atc cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct ggG ttc acc GtC agt

3-64 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-66 gag gtg cag ctg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc GtC agt

3-72 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggA ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-73 gag gtg cag ctg gtg gag tct ggg gga ggc ttg gtc cag cct ggg ggg tcc ctg aAa ctc tcc tgt gca gcc tct ggG ttc acc ttC agt

3-74 gag gtg cag ctg gtg gag tct ggg gga ttA gtT cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc ttC agt

3-d gag gtg cag ctg gtg gag tct Cgg gga gTc ttg gtA cag cct ggg ggg tcc ctg aga ctc tcc tgt gca gcc tct gga ttc acc GtC agt

VH4

4-04 CAG GTG CAG CTG CAG GAG TCG GGC CCA CTG GTG AAG CCT TCG GGG ACC CTG TCC CTC ACC TGC GCT GTC TCT GGT GGT GGC TCC ATC AGC

VH4

4-28 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAC acc ctg tcc ctc acc tgc gct gtc tct ggt TAc tcc atc agc

4-30.1 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcA CAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-30.2 cag Ctg cag ctg cag gag tcC ggc Tca gga ctg gtg aag cct tcA CAg acc ctg tcc ctc acc tgc gct gtc tct ggt gcc tcc atc agc

4-30.4 cag gtg cag gag tcg ggc cca gga ctg gtg aag cct tcA CAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-31 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcA CAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-34 cag gtg cag ctA cag Cag tGg ggc Gca gga ctg Ttg aag cct tcg gAg acc ctg tcc ctc acc tgc gct gct tAt ggt ggG tcc Ttc agT

4-39 cag Ctg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc atc agc

4-59 cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc atc agT

4-61 gac gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc ctc acc tgc Act gtc tct ggt ggc tcc Gtc agc

4-b cag gtg cag ctg cag gag tcg ggc cca gga ctg gtg aag cct tcg gAg acc ctg tcc ctc acc tgc gct gtc tct ggt TAc tcc atc agc

VH5

5-51 GAG GTG CAG CTG GTG CAG TCT GGA GCA GAG GTG AAA AAG CCC GGG GAG TCT CTG AAG ATC TCC TGT AAG GGT TCT GGA TAC AGC TTT ACC

5-a gaA gtg cag crg gtg cag tct gga gca gag gtg aaa aag ccc ggg gag tct ctg aGg atc tcc tgt aag ggt tct gga tac agc ttt acc

VH6

6-1 CAG GTA CAG CTG CAG CAG TCA GGT CCA GGA CTG GTG AAG CCC TCG CAG ACC CTC TCA CTC ACC TGT GCC ATC TCC GGG GAC AGT GTC TCT

VH7

7-4.1 CAG GTG CAG CTG GTG CAA TCT GGG TCT GAG TTC AAG AAG CCT GGG GCC TCA GTG AAG GTT TCC TGC AAG GCT TCT GGA TAC ACC TTC ACT

Tabela 7:locais RERS em FR1 GLG HC humanos em que existem, pelo menos, 20 cortes de GLG

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

1: 4	1: 13	2: 13	3: 4	3: 13	4: 13
6: 13	7: 4	7: 13	8: 13	9: 4	9: 13
10: 4	10: 13	15: 4	15: 65	16: 4	16: 65
17: 4	17: 65	18: 4	18: 65	19: 4	19: 65
20: 4	20: 65	21: 4	21: 65	22: 4	22: 65
23: 4	23: 65	24: 4	24: 65	25: 4	25: 65
26: 4	26: 65	27: 4	27: 65	28: 4	28: 65
29: 4	30: 4	30: 65	31: 4	31: 65	32: 4
32: 65	33: 4	33: 65	34: 4	39: 65	35: 4
35: 65	36: 4	36: 65	37: 4	38: 4	39: 4
41: 4	42: 4	43: 4	45: 4	46: 4	47: 4
48: 4	48: 13	49: 4	49: 13	51: 4	

Existem 39 hits na base# 4

Existem 21 hits na base# 65

-"- ctgcac 9

12: 63	13: 63	14: 63	39: 63	41: 63	42: 63
44: 63	45: 63	46: 63			

BbvI GCAGC 65

1: 6	3: 6	6: 6	7: 6	8: 6	9: 6
10: 6	15: 6	15: 67	16: 6	16: 67	17: 6
17: 67	18: 6	18: 67	19: 6	19: 67	20: 6
20: 67	21: 6	21: 67	22: 6	22: 67	23: 6
23: 67	24: 6	24: 67	25: 6	25: 67	26: 6
26: 67	27: 6	27: 67	28: 6	28: 67	29: 6
30: 6	30: 67	31: 6	31: 67	32: 6	32: 67
33: 6	33: 67	39: 6	39: 67	35: 6	35: 67
36: 6	36: 67	**37**: 6	**38**: 6	**39**: 6	**40**: 6
41: 6	**42**: 6	93: 6	**44**: 6	**45**: 6	**46**: 6
47: 6	98: 6	99: 6	**50**: 12	51: 6	

Existem 43 hits na base# 6 Locais a negrito muito próximos dos locais listados abaixo

Existem 21 hits na base# 67

-"- gctgc 13

37: 9	**38**: 9	**39**: 9	**40**: 3	**40**: 9	**41**: 9
42: 9	**44**: 3	**44**: 9	**45**: 9	**46**: 9	**47**: 9

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

50: 9

Existem 11 hits na base# 9

BsoFI GCnge 78

1: 6	3: 6	6: 6	7: 6	8: 6	9: 6
10: 6	15: 6	15: 67	16: 6	16: 67	17: 6
17: 67	18: 6	18: 67	19: 6	19: 67	20: 6
20: 67	21: 6	21: 67	22: 6	22: 67	23: 6
23: 67	24: 6	24: 67	25: 6	25: 67	26: 6
26: 67	27: 6	27: 67	28: 6	28: 67	29: 6
30: 6	30: 67	31: 6	31: 67	32: 6	32: 67
33: 6	33: 67	34: 6	39: 67	35: 6	35: 67
36: 6	36: 67	37: 6	37: 9	38: 6	38: 9
39: 6	39: 9	40: 3	40: 6	40: 9	41: 6
41: 9	42: 6	42: 9	43: 6	44: 3	44: 6
44: 9	45: 6	45: 9	46: 6	46: 9	47: 6
47: 9	48: 6	49: 6	50: 9	50: 12	51: 6

Existem 43 hits na base# 6 Muitas vezes, estes ocorrem em conjunto.

Existem 11 hits na base# 9

Existem 2 hits na base# 3

Existem 21 hits na base# 67

TseI Gcwgc 78

1: 6	3: 6	6: 6	7: 6	8: 6	9: 6
10: 6	15: 6	15: 67	16: 6	16: 67	17: 6
17: 67	18: 6	18: 67	19: 6	19: 67	20: 6
20: 67	21: 6	21: 67	22: 6	22: 67	23: 6
23: 67	24: 6	24: 67	25: 6	25: 67	26: 6
26: 67	27: 6	27: 67	28: 6	28: 67	29: 6
30: 6	30: 67	31: 6	31: 67	32: 6	32: 67
33: 6	33: 67	34: 6	34: 67	35: 6	35: 67
36: 6	36: 67	37: 6	37: 9	38: 6	38: 9
39: 6	39: 9	40: 3	40: 6	40: 9	41: 6
41: 9	42: 6	42: 9	43: 6	44: 3	44: 6
44: 9	45: 6	45: 9	46: 6	46: 9	47: 6
47: 9	48: 6	49: 6	50: 9	50: 12	51: 6

Existem 43 hits na base# 6 Muitas vezes em conjunto.

Existem 11 hits na base# 9

Existem 2 hits na base# 3

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

Existe 1 hit na base# 12

Existem 21 hits na base# 67

MspAlI CMGckg 48

1: 7	3: 7	4: 7	5: 7	6: 7	7: 7
8: 7	9: 7	10: 7	11: 7	15: 7	16: 7
17: 7	18: 7	19: 7	20: 7	21: 7	22: 7
23: 7	29: 7	25: 7	26: 7	27: 7	28: 7
29: 7	30: 7	31: 7	32: 7	33: 7	34: 7
35: 7	36: 7	37: 7	38: 7	39: 7	40: 1
40: 7	41: 7	42: 7	44: 1	44: 7	45: 7
46: 7	47: 7	48: 7	49: 7	50: 7	51: 7

Existem 46 hits na base# 7

PvuII CAGctg 48

1: 7	3: 7	4: 7	5: 7	6: 7	7: 7
8: 7	9: 7	10: 7	11: 7	15: 7	16: 7
17: 7	18: 7	19: 7	20: 7	21: 7	22: 7
23 : 7	24: 7	25 : 7	26: 7	27: 7	28: 7
29: 7	30: 7	31: 7	32: 7	33: 7	34: 7
35: 7	36: 7	37: 7	38: 7	39: 7	40: 1
40: 7	41: 7	42: 7	44: 1	44: 7	45: 7
46: 7	47: 7	48: 7	49: 7	50: 7	51: 7

Existem 46 hits na base# 7

Existem 2 hits na base# 1

AluI AGct 54

1: 8	2: 8	3: 8	4: 8	4: 24	5: 8
6: 8	7: 8	8: 8	9: 8	10: 8	11: 8
15: 8	16: 8	17: 8	18: 8	19: 8	20: 8
21: 8	22: 8	23: 8	24: 8	25: 8	26: 8
27: 8	28: 8	29: 8	29: 69	30: 8	31: 8
32: 8	33: 8	34: 8	35: 8	36: 8	37: 8
38: 8	39: 8	40: 2	40: 8	41: 8	42: 8
43: 8	44: 2	44: 8	45: 8	46: 8	47: 8
48: 8	48: 82	49: 8	49: 82	50: 8	51: 8

Existem 48 hits na base# 8

Existem 2 hits na base# 2

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

DdeI Ctnag 48

1: 26	1: 48	2: 26	2: 48	3: 26	3: 48
4: 26	4: 48	5: 26	5: 48	6: 26	6: 48
7: 26	7: 48	8: 26	8: 48	9: 26	10: 26
11: 26	12: 85	13: 85	14: 85	15: 52	16: 52
17: 52	18: 52	19: 52	20: 52	21: 52	22: 52
23: 52	24: 52	25: 52	26: 52	27: 52	28: 52
29: 52	30: 52	31: 52	32: 52	33: 52	35: 30
35: 52	36: 52	40: 24	49: 52	51: 26	51: 48

Existem 22 hits na base# 52 52 e 48 nunca em conjunto.

Existem 9 hits na base# 48

Existem 12 hits na base# 26 26 e 24 nunca em conjunto.

HphI tcacc 42

1: 86	3: 86	6: 86	7: 86	8: 80	11: 86
12: 5	13: 5	14: 5	15: 80	16: 80	17: 80
18: 80	20: 80	21: 80	22: 80	23: 80	24: 80
25: 80	26: 80	27: 80	28: 80	29: 80	30: 80
31: 80	32: 80	33: 80	34: 80	35: 80	36: 80
37: 59	38: 59	39: 59	40: 59	41: 59	42: 59
43: 59	44: 59	45: 59	46: 59	47: 59	50: 59

Existem 22 hits na base# 80 80 e 86 nunca em conjunto

Existem 5 hits na base# 86

Existem 12 hits na base# 59

BssKI Nccngg 50

1: 39	2: 39	3: 39	4: 39	5: 39	7: 39
8: 39	9: 39	10: 39	11: 39	15: 39	16: 39
17: 39	18: 39	19: 39	20: 39	21: 29	21: 39
22: 39	23: 39	24: 39	25: 39	26: 39	27: 39
28: 39	29: 39	30: 39	31: 39	32: 39	33: 39
34: 39	35: 19	35: 39	36: 39	37: 24	38: 24
39: 24	41: 24	42: 24	44: 24	45: 24	46: 24
47: 24	48: 39	48: 40	49: 39	49: 40	50: 24
50: 73	51: 39				

Existem 35 hits na base# 39 39 e 40 em conjunto duas vezes.

Existem 2 hits na base# 40

BsaJI Ccnngg 47

1: 40	2:	40 3:	4: 40	5: 40	7: 40

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

8: 40	9: 40	9: 47	10: 40	10: 47	11: 40
15: 40	18: 40	19: 40	20: 40	21: 40	22: 40
23: 40	24: 40	25: 40	26: 40	27: 40	28: 40
29: 40	30: 40	31: 40	32: 40	34: 40	35: 20
35: 40	36: 40	37: 24	38: 24	39: 24	41: 24
42: 24	44: 24	45: 24	46: 24	47: 24	48: 40
48: 41	49: 40	49: 41	50: 74	51: 40	

Existem 32 hits na base# 40 40 e 41 em conjunto duas vezes

Existem 2 hits na base# 41

Existem 9 hits na base# 24

Existem 2 hits na base# 47

BstNI CCwgg 44

PspGI ccwgg

ScrFI(SM.HpaII) CCwgg

1: 40	2: 40	3: 40	4: 40	5: 40	7: 40
8: 40	9: 40	10: 40	11: 40	15: 40	16: 40
17: 40	18: 40	19: 40	20: 40	21: 30	21: 40
22: 40	23: 40	24: 40	25: 40	26: 40	27: 40
28: 40	29: 40	30: 40	31: 40	32: 40	33: 40
34: 40	35: 40	36: 40	37: 25	38: 25	39: 25
41: 25	42: 25	44: 25	45: 25	46: 25	47: 25
50: 25	51: 40				

Existem 33 hits na base# 40

ScrFI CCngg 50

1: 40	2: 40	3: 40	4: 40	5: 40	7: 40
8: 40	9: 40	10: 40	11: 40	15: 40	16: 40
17: 40	18: 40	19: 40	20: 40	21: 30	21: 40
22: 40	23: 40	24: 40	25: 40	26: 40	27: 40
28: 40	29: 40	30: 40	31: 40	32: 40	33: 40
34: 40	35: 20	35: 40	36: 40	37: 25	38: 25
39: 25	41: 25	42: 25	44: 25	45: 25	46: 25
47: 25	48: 40	48: 41	49: 40	49: 41	50: 25
50: 74	51: 40				

Existem 35 hits na base# 40

Existem 2 hits na base# 41

EcoO109I RGgnccy 34

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

1: 43	2: 43	3: 43	4: 43	5: 43	6: 43
7: 43	8: 43	9: 43	10: 43	15: 46	16: 46
17: 46	18: 46	19: 46	20: 46	21: 46	22: 46
23: 46	24: 46	25: 46	26: 46	27: 46	28: 46
30: 46	31: 46	32: 46	33: 46	34: 46	35: 46
36: 46	37: 46	43: 79	51: 43		

Existem 22 hits na base# 46 46 e 43 nunca em conjunto

Existem 11 hits na base# 43

N1aIV GGNncc 71

1: 43	2: 43	3: 43	4: 43	5: 43	6: 43
7: 43	8: 43	9: 43	9: 79	10: 43	10: 79
15: 46	15: 47	16: 47	17: 46	17: 47	18: 46
18: 47	19: 46	19: 47	20: 46	20: 47	21: 46
21: 47	22: 46	22: 47	23: 47	24: 47	25: 47
26: 47	27: 46	27: 47	28: 46	28: 47	29: 47
30: 46	30: 47	31: 46	31: 47	32: 46	32: 47
33: 46	33: 47	34: 46	34: 47	35: 46	35: 47
36: 46	36: 47	37: 21	37: 46	37: 47	37: 79
38: 21	39: 21	39: 79	40: 79	41: 21	41: 79
42: 21	42: 79	43: 79	44: 21	44: 79	45: 21
45: 79	46: 21	46: 79	47: 21	51: 43	

Existem 23 hits na base# 47 46 & 47 muitas vezes em conjunto

Existem 17 hits na base# 46 Existem 11 hits na base# 43

Sau96I Ggncc 70

1: 44	2: 3	2: 44	3: 44	4: 44	5: 3	5: 44	6: 44
7: 44	8: 22	8: 44	9: 44	10: 44	11: 3	12: 22	13: 22
14: 22	15: 33	15: 47	16: 47	17: 47	18: 47	19: 47	20: 47
21: 47	22: 47	23: 33	23: 47	24: 33	24: 47	25: 33	25: 47
26: 33	26: 47	27: 47	28: 47	29: 47	30: 47	31: 33	31: 47
32: 33	32: 47	33: 33	33: 47	34: 33	34: 47	35: 47	36: 47
37: 21	37: 22	37: 47	38: 21	38: 22	39: 21	39: 22	41: 21
41: 22	42: 21	42: 22	43: 80	44: 21	44: 22	45: 21	45: 22
46: 21	46: 22	47: 21	47: 22	50: 22	51: 44		

Existem 23 hits na base# 47 Estes não ocorrem em conjunto.

Existem 11 hits na base# 44

Existem 14 hits na base# 22 Estes ocorrem em conjunto.

Existem 9 hits na base# 21

BsmAI GTCTCNnnnn 22

1: 58	3: 58	4: 58	5: 58	8: 58	9: 58

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

10: 58 13: 70 36: 18 37: 70 38: 70 39: 70

40: 70 41: 70 42: 70 44: 70 45: 70 46: 70

47: 70 48: 48 49: 48 50: 85

Existem 11 hits na base# 70

-"- Nnnnnngagac 27

13: 40 15: 48 16: 48 17: 48 18: 48 20: 48

21: 48 22: 48 23: 48 24: 48 25: 48 26: 48

27: 48 28: 48 29: 48 30: 10 30: 48 31: 48

32: 48 33: 48 35: 48 36: 48 43: 40 44: 40

45: 40 46: 40 47: 40

Existem 20 hits na base# 48

AvaII Ggwcc 44

Sau96I($M.HaeIII) Ggwcc 44

2: 3 5: 3 6: 44 8: 44 9: 44 10: 44

11: 3 12: 22 13: 22 14: 22 15:33 15: 47

16: 47 17: 47 18: 47 19: 47 20: 47 21: 47

22: 47 23: 33 23: 47 24: 33 24: 47 25: 33

25: 47 26: 33 26: 47 27: 47 28: 47 29: 47

30: 47 31: 33 31: 47 32: 33 32: 47 33: 33

33: 47 34: 33 34: 47 35: 47 36: 47 37: 47

43: 80 50: 22

Existem 23 hits na base# 47 44 & 47 nunca em conjunto

Existem 4 hits na base# 44

PpuMI RGgwccy 27

6: 43 8: 43 9: 43 10: 43 15: 46 16: 46

17: 46 18: 46 19: 46 20: 46 21: 46 22: 46

23: 46 24: 46 25: 46 26: 46 27: 46 28: 46

30: 46 31: 46 32: 46 33: 46 34: 46 35: 46

36: 46 37: 46 43: 79

Existem 22 hits na base# 46 43 e 46 nunca ocorrem em conjunto.

Existem 4 hits na base# 43

BsmFI GGGAC 3

8: 43 37: 46 50: 77

-"- gtccc 33

15: 48 16: 48 17: 48 1: 0 1: 0 20: 48

21: 48 22: 48 23: 48 24: 48 25: 48 26: 48

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

27: 48 28: 48 29: 48 30: 48 31: 48 32: 48

33: 48 34: 48 35: 48 36: 48 37: 54 38: 54

39: 54 40: 54 41: 54 42: 54 43: 54 44: 54

45: 54 46: 54 47: 54

Existem 20 hits na base# 48

Existem 11 hits na base# 54

HinfI Gantc 80

8: 77 12: 16 13: 16 14: 16 15: 16 15: 56

15: 77 16: 16 16: 56 16: 77 17: 16 17: 56

17: 77 18: 16 18: 56 18: 77 19: 16 19: 56

19: 77 20: 16 20: 56 20: 77 21: 16 21: 56

21: 77 22: 16 22: 56 22: 77 23: 16 23: 56

23: 77 24: 16 24: 56 24: 77 25: 16 25: 56

25: 77 26: 16 26: 56 26: 77 27: 16 27: 26

27: 56 27: 77 28: 16 28: 56 28: 77 29: 16

29: 56 29: 77 30: 56 31: 16 31: 56 31: 77

32: 16 32: 56 32: 77 33: 16 33: 56 33: 77

34: 16 35: 16 35: 56 35: 77 36: 16 36: 26

36: 56 36: 77 37: 16 38: 16 39: 16 40: 16

41: 16 42: 16 44: 16 45: 16 46: 16 47: 16

48: 46 49: 46

Existem 34 hits na base# 16

TfiI Gawtc 21

8: 77 15: 77 16: 77 17: 77 18: 77 19: 77

20: 77 21: 77 22: 77 23: 77 24: 77 25: 77

26: 77 27: 77 28: 77 29: 77 31: 77 32: 77

33: 77 35: 77 36: 77

Existem 21 hits na base# 77

MlyI GAGTC 38

12: 16 13: 16 14: 16 15: 16 16: 16 17: 16

18: 16 19: 16 20: 16 21: 16 22: 16 23: 16

24: 16 25: 16 26: 16 27: 16 27: 26 28: 16

29: 16 31: 16 32: 16 33: 16 34: 16 35: 16

36: 16 36: 26 37: 16 38: 16 39: 16 40: 16

41: 16 42: 16 44: 16 45: 16 46: 16 47: 16

48: 46 49: 46

BsgI GTGCAG 71 (16 cortes/14 bases à direita)

Existem 34 hits na base# 16

-"- GACTC 21

15: 56 16: 56 17: 56 18: 56 19: 56 20: 56
21: 56 22: 56 23: 56 24: 56 25: 56 26: 56
27: 56 28: 56 29: 56 30: 56 31: 56 32: 56
33: 56 35: 56 36: 56

Existem 21 hits na base# 56

PleI gagtc 38

12: 16 13: 16 14: 16 15: 16 16: 16 17: 16
18: 16 19: 16 20: 16 21: 16 22: 16 23: 16
24: 16 25: 16 26: 16 27: 16 27: 26 28: 16
29: 16 31: 16 32: 16 33: 16 34: 16 35: 16
36: 16 36: 26 37: 16 38: 16 39: 16 40: 16
41: 16 42: 16 44: 16 45: 16 46: 16 47: 16
48: 46 49: 46

Existem 34 hits na base# 16

-"- gactc 21

15: 56 16: 56 17: 56 18: 56 19: 56 20: 56
21: 56 22: 56 23: 56 24: 56 25: 56 26: 56
27: 56 28: 56 29: 56 30: 56 31: 56 32: 56
33: 56 35: 56 36: 56

Existem 21 hits na base# 56

AlwNI CAGNNNctg 26

15: 68 16: 68 17: 68 18: 68 19: 68 20: 68
21: 68 22: 68 23: 68 24: 68 25: 68 26: 68
27: 68 28: 68 29: 68 30: 68 31: 68 32: 68
33: 68 34: 68 35: 68 36: 68 39: 46 40: 46
41: 46 42: 46

Existem 22 hits na basel 68

Tabela 8: GLG FR1 Kappa

!	1	2	3	4	5	6	7	8	9	10	11	12	
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
!	13	14	15	16	17	18	19	20	21	22	23		
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	012
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	CTC	ACC	ATC	ACT	TGC	!	02
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	018
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	08
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	CTC	ACC	ATC	ACT	TGC	!	A20
`	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	A30
	AAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	GCC	ATG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L14
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L1
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L15
	GCC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L4
	GCC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L18
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	TCC	GTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L5
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCT	TCT	GTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L19
	GAC	ATC	CAG	TTG	ACC	CAG	TCT	CCA	TCC	TTC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L8
	GCC	ATC	CGG	ATG	ACC	CAG	TCT	CCA	TTC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L23
	GCC	ATC	CGG	ATG	ACC	CAG	TCT	CCA	TCC	TCA	TTC	TCT	
	GCA	TCT	ACA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGT	!	L9
	GTC	ATC	TGG	ATG	ACC	CAG	TCT	CCA	TCC	TTA	CTC	TCT	
	GCA	TCT	ACA	GGA	GAC	AGA	GTC	ACC	ATC	AGT	TGT	!	L24
	GCC	ATC	CAG	ATG	ACC	CAG	TCT	CCA	TCC	TCC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L11
	GAC	ATC	CAG	ATG	ACC	CAG	TCT	CCT	TCC	ACC	CTG	TCT	
	GCA	TCT	GTA	GGA	GAC	AGA	GTC	ACC	ATC	ACT	TGC	!	L12

GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC
GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC ! O11
GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC
GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC ! 01
GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC
GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC ! A17
GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC
GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC ! A1
GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC
GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC ! A18
GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC
GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC ! A2
GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC
GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC ! A19
GAT ATT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC
GTC ACC CCT GGA GAG CCG gcc tcc atc tcc tgc ! A3
GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC TCA CCT
GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC ! A23
GAA ATT GTG TTG ACG CAG TCT CCA GGC ACC CTG TCT
TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! A27
GAA ATT GTG TTG ACG CAG TCT CCA GCC ACC CTG TCT
TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! A11
GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT
GTG TCT CCA GGG GAA GAG GCC ACC CTC TCC TGC ! L2
GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT
GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L16
GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT
TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L6
GAA ATT GTG TGG ACA CAG TCT CCA GCC ACC CTG TCT
TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L20
GAA ATT GTA ATG ACA CAG TCT CCA GCC ACC CTG TCT
TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC ! L25
GAT ATC GTG ATG ACC CAG TCT CCA GAC TCC CTG GCT
GTG TCT CTG GGC GAG AGG GCC ACC ATC AAC TGC ! B3
GAA ACG ACA CTC ACG CAG TCT CCA GCA TTC ATG TCA
GCG ACT CCA GGA GAC AAA GTC AAC ATC TCC TGC ! B2
GAA ATT GTG CTG ACT CAG TCT CCA GAC TTT CAG TCT
GTG ACT CCA AAG GAG AAA GTC ACC ATC ACC TGC ! A26
GAA ATT GTG CTG ACT CAG TCT CCA GAC TTT CAG TCT
GTG ACT CCA AAG GAG AAA GTC ACC ATC ACC TGC ! A10
GAT GTT GTG ATG ACA CAG TCT CCA GCT TTC CTC TCT

GTG ACT CCA GGG GAG AAA GTC ACC ATC ACC TGC ! A14

Tabela 9 locais RERS encontrados em GLG FR1 Kappa Humanos

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
O12 1-69	3	3 23	12 49	15	18 47	26	36
O2 101.169	103	103 123	112 149	115	118 147	126	136
O18 201-269	203	203 223	212 249	215	218 247	226	236
O8 301-369	303	303 323	312 349	315	318 347	326	336
A20 401-469	403	403 423	412 449	415	418 447	426	436
A30 501-569	503	503 523	512 549	515	518 547	526	536
L14 601-669	603	603	612 649	615	618 647	-	636
L1 701-769	703	703 723	712 749	715	718 747	726	736
L15 801-869	803	803 823	812 849	815	818 847	826	836
L4 901-969	-	903 923	912 949	906 915	918 947	926	936
L18 1001-1069	-	1003	101 1049	1006 1015	1018 1047	1026	1036
L5 1101-1169	1103	-	1112 1149	1115	1118 147	-	1136
L19 1201-1269	1203	1203	1212 1249	1215	1218 247	-	1236
L8 1301-1369	-	1303 1323	1312 1349	1306 1315	1318 347	-	1336
L23 1401-1469	1403	1408 1403	1412 1449	1415	1418 447	-	1436
L9 1501-1569	1503	1503 1508 1523	1512 1549	1515	1518 1547	1526	1536
L24 1601-1669	1603	1608 1623	1612 1649	1615	1618 1647	-	1636
L11 1701-1769	1703	1703 1723	1712 1749	1715	1718 1747	1726	1736
L12 1801-1869	1803	1803	1812 1849	1815	1818 1847	-	1836
VKII							
O11 1901-1969	-	-	-	-	-	1956	-
O1 2001-2069	-	-	-	-		2056	-
A17 2101-2169	-	-	2112	-	2118	2156	-

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
A1 2201-2269	-	-	2212	-	2218	2256	-
A18 2301-2369	-	-	-	-	-	2356	-
A2 2401-2469	-	-	-	-	-	2456	-
A19 2501-2569	-	-	2512	-	2518	2556	-
A3 2601-2669	-	-	2612	-	2618	2656	-
A23 2701-2769		- -	-	-	-	2729 2756	-
VKIII							-
A27 2801-2869	-	-	2812	-	2818 2839	2860	-
A11 2901-2969	-	-	2912	-	2918 2939	2960	-
L2 3001-3069	-	-	3012	-	3018 3039	3060	-
L16 3101-3169	-	-	3112	-	3118 3139	3160	-
L6 3201-3269	-	-	3212	-	3218 3239	3260	-
L20 3301-3369	-	-	3312	-	3318 3339	3360	-
L25 3401-3469	-	-	3412	-	3418 3439	3460	
VKIV							
B3 3501-3569	3503	-	3512	3515	3539 3518	3551<	-
VKV							
82 3601-3669	-	-	3649	-	3647 3618		-
VKVI							
A26 3701-3769	-	-	3712	-	3718		-
A10 3801-3869	-	-	3812	-	3818		-
A14 3901-3969	-	-	3912	-	3918	3930>	-
VKI							
O12 1-69	37	41	53	53	55	56	-
O2 101-169	137	141	153	153	155	156	-
O18 201-269	237	241	253	253	255	256	-

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
O8 301-369	337	341	353	353	355	356	-
A20 401-469	437	441	453	453	455	456	-
A30 501-569	537	541	553	553	555	556	-
L14 601-669	637	641	653	653	655	656	-
L1 701-769	737	741	753	753	755	756	-
L15 801-869	837	841	853	853	855	856	-
L4 901-969	937	941	953	953	955	956	-
L18 1001-1069	1037	1041	1053	1053	1055	1056	-
L5 1101-1169	1137	1141	1153	1153	1155	1156	-
L19 1201-1269	1237	1241	1253	1253	1255	1256	-
L8 1301-1369	1337	1341	1353	1353	1355	1356	-
L23 1401-1469	1437	1441	1453	1453	1455	1456	1406
L9 1501-1569	1537	1541	1553	1553	1555	1556	1506
	SfaNI	SfcI	HinfI	MlyI --> --> <--	mesmos locais MaeIII TspEI	HphI xx38 xx56 xx62	HpaII MspI xx06 xx52
L24 1601-1669	1637	1641	1653	1653	1655	1656	
L11 1701-1769	1737	1741	1753	1753	1755	1756	
L12 1801-1869	1837	1841	1853	1853	1855	1856	
VKII							
O11 1901-1969	-	-	1918	1918	1937	1938	1952
O1 2001-2069	-	-	2018	2018	2037	2038	2052
A17 2101-2169	-	-	2112	2112	2137	2138	2152
A1 2201-2269	-	-	2212	2212	2237	2238	2252
A18 2301-2369	-	-	2318	2318	2337	2338	2352
A2 2401-2469	-	-	2418	2418	2437	2438	2452
A19 2501-2569	-	-	2512	2512	2537	2538	2552
A3 2601-2669	-	-	2612	2612	2637	2638	2652

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
A23 2701-2769	-	-	2718	2718	2737	2731* 2738*	-
VKIII							
A27 2801-2869	-	-	-	-			-
A11 2901-2969	-	-	-	-			-
L2 3001-3069	-	-	-	-			-
L16 3101-3169	-	-	-	-			-
L6 3201-3269	-	-	-	-			-
L20 3301-3369	-	-	-	-			-
L25 3401-3469	-	-	-	-			-
VKIV							
B3 3501-3569	-	-	3525	3525			-
VKV							
B2 3601-3669	-	-	3639	3639			-
VKVI							
A26 3701-3769	-	-	3712 3739	3712 3739	3737 3755	3756 3762	-
A10 3801-3869	-	-	3839 3812	3812 3839	3837 3855	3856 3862	-
A14 3901-3969	-	-	3939	3939	3937 3955	3956 3962	-
	BsaJI xx29 xx42 xx43	BssKI (NstNI) xx22 xx30 xx43	BpmI xx20 xx41 xx44 --> --> <--	BsrFI Cac8I NaeI NgoMIV	HaeIII	Tsp509I	
VKI							
O12 1-69	-	-	-	-	-	-	
O2 101-169	-	-	-	-	-	-	
O18 201-269	-	-	-	-	-	-	
O8 301-369	-	-	-	-	-	-	
A20 401-469	-	-	-	-	-	-	
A30 501-569	-	-	-	-	-	-	

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
L14 601-669	-	-	-	-	-	-	
L1 701-769	-	-	-	-	-	-	
L15 801-869	-	-	-	-	-	-	
L4 901-969	-	-	-	-	-	-	
L18 1001-1069	-	-	-	-	-	-	
L5 1101-1169	-	-	-	-	-	-	
L19 1201-1269	-	-	-	-	-	-	
L8 1301-1369	-	-	-	-	-	-	
L23 1401-1469	-	-	-	-	-	-	
L9 1501-1569	-	-	-	-	-	-	
L24 1601-1669	-	-	-	-	-	-	
L11 1701-1769	-	-	-	-	-	-	
L12 1801-1869	-	-	-	-	-	-	
VKII							
O11 1901-1969	1942	1943	1944	1951	1954	-	
O1 2001-2069	2042	2043	2044	2051	2054	-	
A17 2101-2169	2142	-	-	2151	2154	-	
A1 2201-2269	2242	-	-	2251	2254	-	
A18 2301-2369	2342	2343	-	2351	2354	-	
A2 2401-2469	2442	2443	-	2451	2454	-	
A19 2501-2569	2542	2543	2544	2551	2554	-	
A3 2601-2669	2642	2643	2644	2651	2654	-	
A23 2701-2769	2742	-	-	2751	2754	-	
VKIII							
A27 2801-2869	2843	2822 2843	2820 2841	-	-	2803	
A11 2901-2969	2943	2943	2920 2941	-	-	2903	
L2 3001-3069	3043	3043	3041	-	-	-	

	MslI	FokI --> <-- ->	PflFI	BsrI	BsmAI	MnlI	HpyCH 4V
VKI							
L16 3101-3169	3143	3143	3120 3141	-	-	-	
L6 3201-3269	3243	3243	3220 3241	-	-	3203	
L20 3301-3369	3343	3343	3320 3341	-	-	3303	
L25 3401-3469	3443	3443	3420 3441	-	-	3403	
VKIV							
B3 3501-3569	3529	3530	3520	-	3554		
VKV							
B2 3601-3669		3643	3620 3641	-	-		
VKVI							
A26 3701-3769		-	3720	-	-	3703	
A10 3801-3869		-	3820	-	-	3803	
A14 3901-3969	3943	3943	3920 3941	-	-	-	

122

Tabela 10 Sequências FR1 GLG Lambda

! VL1

CAG TCT GTG CTG ACT CAG CCA CCC TCG GIG TCT GAA

GCC CCC AGG CAG AGG CTC ACC AIC TCC TGT ! 1a

cag tct gtg ctg acG cag ccG ccc tcA gtg tct gGG

gcc ccA Ggg cag agg gtc acc atc tcc tgC ! 1e

cag tct gtg ctg act cag cca ccc tcA gCg tct gGG

Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1c

cag tct gtg ctg act cag cca ccc tcA gCg tct gGG

Acc ccc Ggg cag agg gtc acc atc tcT tgt ! 1g

cag tct gtg Itg acG cag ccG ccc tcA gtg tct gCG

gcc ccA GaA cag aAg gtc acc atc tcc tgC ! 1b

! VL2

CAG TCT GCC CTG ACT CAG CCT CCC TCC GCG TCC GGG

TCT CCT GGA CAG TCA GTC ACC ATC TCC TGC ! 2c
cag tct gcc ctg act cag cct cGc tcA gTg tcc ggg
tct cct gga cag tca gtc acc act tcc tgc ! 2e
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg
tct cct gga cag tcG Atc acc atc tcc tgc ! 2a2
cag tct gcc ctg act cag cct ccc tcc gTg tcc ggg
tct cct gga cag tca gtc acc atc tcc tgc ! 2d
cag tct gcc ctg act cag cct Gcc tcc gTg tcT ggg
tct cct gga cag tcG Atc acc atc tcc tgc ! 2b2

! VL3

TCC TAT GAG CTG ACT CAG CCA CCC TCA GTG TCC GTG
TCC CCA GGA CAG ACA GCC AGC ATC ACC TGC! 3r
tcc tat gag ctg act cag cca cTc tca gtg tcA gtg
Gcc cTG gga cag acG gcc agG atT acc tgT ! 3j
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg
tcc cca gga caA acG gcc agG atc acc tgc ! 3p
tcc tat gag ctg acA cag cca ccc tcG gtg tcA gtg
tcc cTa gga cag aTg gcc agG atc acc tgc ! 3a
tT tCt gag ctg act cag GAC ccT GcT gtg tcT gtg
Gcc TTG gga cag aca gTc agG atc acA tgc ! 31
tcc tat gTg ctg act cag cca ccc tca gtg tcA gtg
Gcc cca gga Aag acG gcc agG atT acc tgT ! 3h
tcc tat gag ctg acA cag cTa ccc tcG gtg tcA gtg
tcc cca gga cag aca gcc agG atc acc tgc ! 3e
tcc tat gag ctg aTG cag cca ccc tcG gtg tcA gtg
tcc cca gga cag acG gcc agG atc acc tgc ! 3m
tcc tat gag ctg acA cag cca Tcc tca gtg tcA gtg
tcT ccG gga cag aca gcc agG atc acc tgc ! V2-19

VL4

CTG CCT GTG CTG ACT CAG CCC CCG TCT GCA TCT GCC
TTG CTG GGA GCC TCG ATC AAG CTC ACC TGC ! 4c
cAg cct gtg ctg act caA TcA TcC tct gcC tct gcT
tCC ctg gga Tcc tcg Gtc aag ctc acc tgc ! 4a
cAg cTt gtg ctg act caA TcG ccC tct gcC tct gcc
tCC ctg gga gcc tcg Gtc aag ctc acc tgc ! 4b

! VL5

CAG CCT GTG CTG ACT CAG CCA CCT TCC TCC TCC GCA
TCT CCT GGA GAA TCC GCC AGA CTC ACC TGC ! 5e
cag Gct gtg ctg act cag ccG Gct tcc CTc tcT gca
tct cct gga gCa tcA gcc agT ctc acc tgc ! 5c
cag cct gtg ctg act cag cca Tct tcc CAT tcT gca

```
   Tct Tct gga gCa tcA gTc aga ctc acc tgc ! 5b
! VL6
   AAT TTT ATG CTG ACT CAG CCC CAC TCT GTG TCG GAG
   TCT CCG GGG AAG ACG GTA ACC ATC TCC TGC ! 6a
! VL7
   CAG ACT GTG GTG ACT CAG GAG CCC TCA CTG ACT GTG
   TCC CCA GGA GGG ACA GTC ACT CTC ACC TGT ! 7a
   cag Gct gtg gtg act cag gag ccc tca ctg act gtg
   tcc cca gga ggg aca gtc act ctc acc tgt ! 7b
! VL8
   CAG ACT GTG GTG ACC CAG GAG CCA TCG TTC TCA GTG
   TCC CCT GGA GGG ACA GTC ACA CTC ACT TGT ! 8a
! VL9
   CAG CCT GTG CTG ACT CAG CCA CCT TCT GCA TCA GCC
   TCC CTG GA GCC TCG GTC ACA CTC ACC TGC ! 9a
! VL10
   CAG GCA GGG CTG ACT CAG CCA CCC TCG GTG TCC AAG
   GGC TTG AGA CAG ACC GCC ACA TCT ACC TGC ! 10a
```

Tabela 11 RERS encontrados em GLG FR1 lambda humanos

! Existem 31 GLG lambda

MlyI NnnnnnGACTC 25

1: 6	3: 6	4: 6	6: 6	7: 6	8: 6
9: 6	10: 6	1: 6	12: 6	15: 6	16: 6
20: 6	21: 6	22: 6	23: 6	23: 50	24: 6
25: 6	25: 50	26: 6	27: 6	28: 6	30: 6
31: 6					

Existem 23 hits na base# 6

-"- GAGTCNNNNNn 1

26: 34

MwoI GCNNNNNnngc 20

1: 9	2: 9	3: 9	4: 9	11: 9	11: 56
12: 9	13: 9	14: 9	16: 9	17: 9	18: 9
19: 9	20: 9	23: 9	24: 9	25: 9	26: 9
30: 9	31: 9				

Existem 19 hits na base# 9

HinfI Gantc 27

1: 12	3: 12	4: 12	6: 12	7: 12	8: 12
9: 12	10: 12	11: 12	12: 12	15: 12	16: 12
20: 12	21: 12	22: 12	23: 12	23: 46	23: 56
24: 12	25: 12	25: 56	26: 12	26: 34	27: 12
28: 12	30: 12	31: 12			

Existem 23 hits na base# 12

PleI gactc 25

1: 12	3: 12	4: 12	6: 12	7: 12	8: 12
9: 12	10: 12	11: 12	12: 12	15: 12	16: 12
20: 12	21: 12	22: 12	23: 12	23: 56	24: 12
25: 12	25: 56	26: 12	27: 12	28: 12	30: 12
31: 12					

Existem 23 hits na base# 12

-"- gagtc 1

26: 34

DdeI Ctnag 32

1: 14	2: 24	3: 14	3: 24	4: 14	4: 24
5: 24	6: 14	7: 14	7: 24	8: 14	9: 14
10: 14	11: 14	11: 24	12: 14	12: 24	15: 5

15: 14 16: 14 16: 24 19: 24 20: 14 23: 14
24: 14 25: 14 26: 14 27: 14 28: 14 29: 30
30: 14 31: 14
Existem 21 hits na base# 14

BsaJI Ccnngg 38
1: 23 1: 40 2: 39 2: 40 3: 39 3: 40
4: 39 9: 40 5: 39 11: 39 12: 38 12: 39
13: 23 13: 39 14: 23 14: 39 15: 38 16: 39
17: 23 17: 39 18: 23 18: 39 21: 38 21: 39
21: 47 22: 38 22: 39 22: 47 26: 40 27: 39
28: 39 29: 14 29: 39 30: 38 30: 39 30: 47
31: 23 31: 32
Existem 17 hits na base# 39
Existem 5 hits na base# 38
Existem 5 hits na base# 40 Fazem clivagem irregular.
Mn1I cctc 35
1: 23 2: 23 3: 23 4: 23 5: 23 6: 19
6: 23 7: 19 8: 23 9: 19 9: 23 10: 23
11: 23 13: 23 14: 23 16: 23 17: 23 18: 23
19: 23 20: 47 21: 23 21: 29 21: 47 22: 23
22: 29 22: 35 22: 47 23: 26 23: 29 24: 27
27: 23 28: 23 30: 35 30: 47 31: 23
Existem 21 hits na base# 23
Existem 3 hits na base# 19
Existem 3 hits na base# 29
Existe 1 hit na base# 26
Existe 1 hit na base# 27 Estes podem fazer clivagem irregular.
-"- gagg 7
1: 48 2: 48 3: 48 4: 48 27: 44 28: 44
29: 44

BssKI Nccngg 39
1: 40 2: 39 3: 39 3: 40 4: 39 9: 40
5: 39 6: 31 6: 39 7: 31 7: 39 8: 39
9: 31 9: 39 10: 39 11: 39 12: 38 12: 52
13: 39 13: 52 14: 52 16: 39 16: 52 17: 39
17: 52 18: 39 18: 52 19: 39 19: 52 21: 38
22: 38 23: 39 24: 39 26: 39 27: 39 28: 39
29: 14 29: 39 30: 38

Existem 21 hits na base# 39

Existem 4 hits na base# 38

Existem 3 hits na base# 31

Existem 3 hits na base# 40 Irregular

BstNI CCwgg 30

1: 41	2: 40	5: 40	6: 40	7: 40	8: 40
9: 40	10: 40	11: 40	12: 39	12: 53	13: 40
13: 53	14: 53	16: 40	16: 53	17: 40	17: 53
18: 40	18: 53	19: 53	21: 39	22: 39	23: 40
24: 40	27: 40	28: 40	29: 15	29: 40	30: 39

Existem 17 hits na base# 40

Existem 7 hits na base# 53

Existem 4 hits na base# 39

Existe 1 hit na base# 41 irregular

PspGI ccwgg 30

1: 41	2: 40	5: 40	6: 40	7: 40	8: 40
9: 40	10: 40	11: 40	12: 39	12: 53	13: 40
13: 53	14: 53	16: 40	16: 53	17: 40	17: 53
18: 40	18: 53	19: 53	21: 39	22: 39	23: 40
24: 40	27: 40	28: 40	29: 15	29: 40	30: 39

Existem 17 hits na base# 40

Existem 7 hits na base# 53

Existem 4 hits na base# 39

Existe 1 hit na base# 41

ScrFI CCngg 39

1: 41	2: 40	3: 40	3: 41	4: 40	4: 41
5: 40	6: 32	6: 40	7: 32	7: 40	8: 40
9: 32	9: 40	10: 40	11: 40	12: 39	12: 53
13: 40	13: 53	14: 53	16: 40	16: 53	17: 40
17: 53	18: 40	18: 53	19: 40	19: 53	21: 39
22: 39	23: 40	29: 40	26: 40	27: 40	28: 40
29: 15	29: 40	30: 39			

Existem 21 hits na base# 40

Existem 4 hits na base# 39

Existem 3 hits na base# 41

MaeIII gtnac 16

1: 52	2: 52	3: 52	4: 52	5: 52	6: 52
7: 52	9: 52	26: 52	27: 10	27: 52	28: 10
28: 52	29: 10	29: 52	30: 52		

Existem 13 hits na base# 52

Tsp45I gtsac 15

1: 52	2: 52	3: 52	4: 52	5: 52	6: 52
7: 52	9: 52	27: 10	27: 52	28: 10	28: 52
29: 10	29: 52	30: 52			

Existem 12 hits na base# 52

HphI tcacc 26

1: 53	2: 53	3: 53	4: 53	5: 53	6: 53
7: 53	8: 53	9: 53	10: 53	11: 59	13: 59
14: 59	17: 59	18: 59	19: 59	20: 59	21: 59
22: 59	23: 59	24: 59	25: 59	27: 59	28: 59
30: 59	31: 59				

Existem 16 hits na base# 59

Existem 10 hits na base# 53

BspMI ACCTGCNNNNn 14

11: 61	13: 61	14: 61	17: 61	18: 61	19: 61
20: 61	21: 61	22: 61	23: 61	24: 61	25: 61
30: 61	31: 61				

Existem 14 hits na base# 61 Vai para CDR1

Tabela 12: Combinações com adaptadores URE FR3 em 79 HC humanos.

A. Lista de genes de cadeias pesadas amostrados

AF008566	AF103367	HSA235674	HSU94417	S83240
AF035043	AF103368	HSA235673	HSU94418	SABVH369
AF103026	AF103369	HSA240559	HSU96389	SADEIGVH
af103033	AF103370	HSCB201	HSU96391	SAH2IGVH
AF103061	af103371	HSIGGVHC	HSU96392	SDA3IGVH
Af103072	AF103372	HSU44791	HSU96395	SIGVHTTD
af103078	AF158381	HSU44793	HSZ93849	SUK4IGVH
AF103099	E05213	HSU82771	HSZ93850	
AF103102	E05886	HSU82949	HSZ93851	
AF103103	E05887	HSU82950	HSZ93853	
AF103174	HSA235661	HSU82952	HSZ93855	

AF103186	HSA235664	HSU82961	HSZ93857
af103187	HSA235660	HSU86522	HSZ93860
AF103195	HSA235659	HSU86523	HSZ93863
af103277	HSA235678	HSU92452	MCOMFRAA
af103286	HSA235677	HSU94412	MCOMFRVA
AF103309	HSA235676	HSU94415	582745
af103343	HSA235675	HSU94416	S82764

Tabela 12B. Teste de todos os GLG diferentes das bases 89.1 a 93.2 do domínio variável pesado

Id	Nb	0	1	2	3	4			SEQ ID Nº
1	38	15	11	10	0	2	Seq1	gtgtattactgtgc	25
2	19	7	6	4	2	0	Seq2	gtAtattactgtgc	26
3	1	0	0	1	0	0	Seq3	gtgtattactgtAA	27
4	7	1	5	1	0	0	Seq4	gtgtattactgtAc	28
5	0	0	0	0	0	0	Seq5	Ttgtattactgtgc	29
6	0	0	0	0	0	0	Seq6	TtgtatCactgtgc	30
7	3	1	0	1	1	0	Seq7	ACAtattactgtgc	31
8	2	0	2	0	0	0	Seq8	ACgtattactgtgc	32
9	9	2	2	4	1	0	Seq9	ATqtattactqtqc	33
Grupo		26	26	21	4	2			
Cumulativo		26	52	73	77	79			

Tabela 12C Seqs de reconhecimento de URE mais importante em FR3 Pesado

1	VHSzy1	GTGtattactgtgc	(ON_SHC103)	(SEQ ID Nº 25)
2	VHSzy2	GTAtattactgtgc	(ON_SHC323)	(SEQ ID Nº 26)
3	VHSzy4	GTGtattactgtac	(ON_SHC349)	(SEQ ID Nº 28)
4	VHSzy9	ATGtattactgtgc	(ON_SHC5a)	(SEQ ID Nº 33)

Tabela 12D, teste de genes V de 79 HC de humanos com quarto sondas

Número de sequências										79
Número de bases										29143
Número de combinações										
Id	Best	0	1	2	3	4	5			
1	39	15	11	10	1	2	0	Seq1 gtgtattactgtgc	(SEQ ID Nº 25)	
2	22	7	6	5	3	0	1	Seq2 gtAtattactgtgc	(SEQ ID Nº 26)	
3	7	1	5	1	0	0	0	Seq4 gtgtattactgtAc	(SEQ ID Nº 28)	
4	11	2	4	4	1	0	0	Seq9 ATgtattactgtgc	(SEQ ID Nº 33)	

Grupo	25	26	20	5	2
Cumulativo	25	51	71	76	78

Uma sequência tem cinco combinações com as sequências 2, 4, e 9; é classificada como a melhor para 2.

Id é o número do adaptador.

Best é o número da sequência para a qual o adaptador identificado era o melhor disponível.

O resto da Tabela mostra como as sequências combinam bem com os adaptadores. Por exemplo, existem 10 sequências que combinam com VHSzy1(Id=1 com 2 emparelhamentos incorrectos e são piores para todos os outros adaptadores. Nesta amostra, 90% vem dentro de 2 bases de um dos quarto adaptadores.

Tabela 13

A seguinte lista de enzimas foi retirada de http://rebase.neb.com/cgi-bin/asymmlist.

Removeram-se as enzimas que a) cortam dentro do reconhecimento, b) cortam em ambos os lados do reconhecimento ou c) têm menos do que 2 bases entre o local de reconhecimento e de corte mais próximo.

Enzimas REBASE 04/13/2001

Enzimas de restrição Tipo II com sequências de reconhecimento assimétricas:

Enzimas	Sequência de Reconhecimento	Isosquizómeros	Fornecedores
AarI	CACCTGCNNNN^NNNN_	-	y
AceIII	CAGCTCNNNNNNN^NNNN_	-	-
Bbr7I	GAAGACNNNNNNN^NNNN_	-	-
BbvI	GCAGCNNNNNNNN^NNNN_		y
BbvII	GAAGACNN^NNNN_		
Bce83I	CTTGAGNNNNNNNNNNNNN_NN^_	-	-
BceAI	ACGGCNNNNNNNNNNNN^NN_	-	y
BcefI	ACGGCNNNNNNNNNNNN^N_	-	-
BciVI	GTATCCNNNNN_N^	BfuI	y
BfiI	ACTGGGNNNN_N^	BmrI	y
BinI	GGATCNNNN^N_		
BscAI	GCATCNNNN^NN_	-	-
BseRI	GAGGAGNNNNNNNN_NN^	-	y
BsmFI	GGGACNNNNNNNNNN^NNNN_	BspLU11III	y
BspMI	ACCTGCNNNN^NNNN_	Acc36I	y
EciI	GGCGGANNNNNNNNN_NN^	-	y
Eco57I	CTGAAGNNNNNNNNNNNNNN_NN^	BspKT5I	y
FauI	CCCGCNNNN^NN_	BstFZ438I	y
FokI	GGATGNNNNNNNNN^NNNN_	BstPZ418I	y

GsuI	NN^ CTGGAGNNNNNNNNNNNNNN_ NN^	-	y
HgaI	GACGCNNNNN^NNNNN_	-	y
HphI	GGTGANNNNNNN N^	AsuHPI	y
MboII	GAAGANNNNNNN_N^	-	y

Enzimas	Sequência de Reconhecimento	Isosquizómeros	Fornecedores
MlyI	GAGTCNNNNN^	SchI	y
MmeI	TCCRACNNNNNNNNNNNNNNNNN_ NN^	-	-
MnlI	CCTCNNNNNN_N^	-	y
PleI	GAGTCNNNN^N_	PpsI	y
RleAI	CCCACANNNNNNNNN_NNN^	-	-
SfaNI	GCATCNNNNN^NNNN_	BspST5I	y
SspD5I	GGTGANNNNNNNN^	-	-
Sth132I	CCCGNNNN^NNNN_	-	-
StsI	GGATGNNNNNNNNNN^NNNN_	-	-
TaqII	GACCGANNNNNNNNN_NN^, CACCCANNNNNNNNN_NN^	-	-
TthlllII	CAARCANNNNNNNNN_NN^	-	-
UbaPI	CGAACG		

A notação é ^ significa corte da cadeia superior e … significa corte da cadeia inferior. Se as cadeias superior e inferior estão cortadas no mesmo local, então aparece apenas ^.

Tabela 14

(FOKIact) 5′-**cAcATccgTg** TTgTT **cAcggATgTg**-3′ a rever pelos ADN emies

(VHEx881) 5′-AATAgTAgAc TgcAgTgTcc TcAgcccTTA AgcTgTTcAT cTgcAAgTAgAgAgTATTcT TAgAgTTgTc TcTAgAcTTA gTgAAgcg-3′

! note-se que VHEx881 é o complemento reverso do ON abaixo

! **[RC]** 5′-cgCttcacTaag-

! Scab.....

! 3-23 Sintética como na Tabela 206

! |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|XbaI...

! |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|t-3′ Af1II...

(VHBA881) 5′-cgCttcacTaag|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg||aac|agC|TTA|AGg|gct|gag |gac|aCT|GCA|Gtc|tac|tat|tgt gcg ag-3′

(VHBB881) 5′-cgCttcacTaag|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg||aac|agC|TTA|AGg|gct|gag |gac|aCT|GCA|Gtc|tac|tat|tgt Acg ag-3′

(VH881PCR) 5′-cgCttcacTaag|TCT|AGA|gac|aac -3′

Tabela 15: Utilização de *Fok*I como "Enzima de Restrição Universal"

*Fok*I - para ADNcd, I representa locais de clivagem

locais de clivagem 5'-cacGGATGtq--nnnnnnn|nnnnnnn-3'(SEQ ID N° 15)

3'-gtgCCTACac--nnnnnnnnnnn|nnn-5' (SEQ ID Nº 16)

RECONHE

CImento de *Fok*I

Caso I

5'-...gtg|tatt-actgtgc..Substrato....-3' (SEQ ID N° 17)

3'-cac-ataa|tgacacg¬

gtGTAGGcac

5'- caCATCCgtg/ (SEQ ID Nº 18)

Caso II

5'-...gtgtatt|agac-tgc..Substrato....-3' (SEQ ID N° 19)

⌐cacataa-tctg|acg-5'

/gtgCCTACac

\cacGGATGtg-3' (SEQ ID Nº 20)

Caso III (Caso I rodado 180 graus)

/gtgCCTACac-5'

\cacGGATGtg¬

gtgtctt|acag-tcc-3' Adaptador (SEQ ID Nº 21)

3'-...cacagaa-tgtc|agg..substrato....-5'(SEQ ID N° 22)

Caso IV (Caso II rodado 180 graus)

3'- gtGTAGGcac\(SEQ ID Nº 23)

⌐caCATCCgtg/

5'-gag|tctc-actgagc

caCATCCgtg/

5'-gag|tctc-actgagc

Substrato 3'-...ctc-agag|tgactcg...-5'(SEQ ID N° 24)

Adaptadores *Fok*I melhorados

*Fok*I - para ADNcd, I representa locais de clivagem

Caso I

Haste 11, ansa 5, haste 11, reconhecimento 17

5'-...catgtg|tatt-actgtgc..Substrato....-3'

3'-gtacac-ataa|tgacacg¬ ⌐T¬

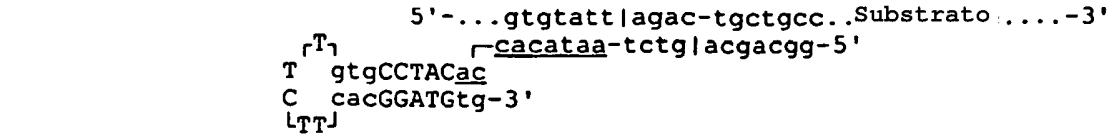
gtGTAGGcacG T

5'- caCATCCgtgc C

└TT┘

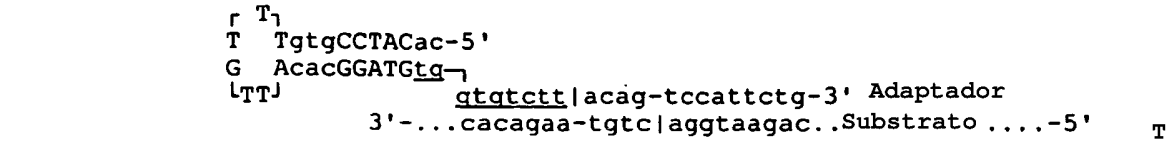
Caso II

Haste 10, ansa 5, haste 10, reconhecimento 18



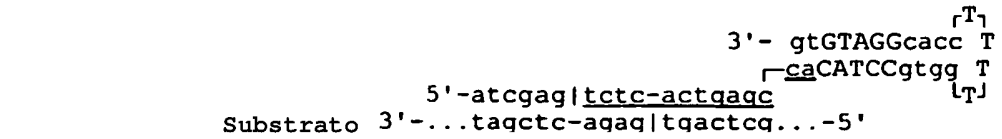
Caso III (Caso I rodado 180 graus)

Haste 11, ansa 5, haste 11, reconhecimento 20



Caso IV (Caso II rodado 180 graus)

Haste 11, ansa 4, haste 11, reconhecimento 17



BseRI

|locais de clivagem

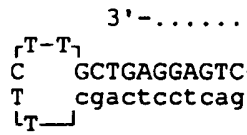
5′-cacGAGGAGnnnnnnnnnn|nnnnn-3′

3′-gtgctcctcnnnnnnnn|nnnnnnn-5′

RECONHE

CIMento de *Bse*RI

Haste 11, ansa 5, haste 11, reconhecimento 19



Um adaptador para BseRI clivar o substrato acima.

Tabela 16 Bases de cadeias pesadas humanas 88.1 a 94.2

Número de sequências 840

		Número de Emparelhadores Incorrectos.....								**Sonda**	
Id	**Ntot**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Nome**	**Sequência.....Modo com Pontos.....**
1	364	152	97	76	26	7	4	2	0	VHS881-1.1	gctgtgtattactgtgcgag gctgtgtattactgtgcgag
2	265	150	60	33	13	5	4	0	0	VHS881-1.2	gccgtgtattactgtgcgag ..c.....
3	96	14	34	16	10	5	7	9	1	VHS881-2.1	gccgtatattactgtgcgag ..c..a.....
4	20	0	3	4	9	2	2	0	0	VHS881-4.1	gccgtgtattactgtacgag ..c.....a....
5	95	25	36	18	11	2	2	0	1	VHS881-9.1	gccatgtattactgtgcgag ..ca.....
	840	341	230	147	69	21	19	11	2		
		341	571	718	787	808	827	838	840		

88 89 90 91 92 93 94 95 Número de codão como na Tabela 195

Reconhecimento..... Haste.....ansa. Haste.....

(VHS881-1.1) 5′-gctgtgtat|tact-gtgcgag **cAcATccgTg** TTgTT **cAcggATgTg-3′**

(VHS881-1.2) 5′-gctgtgtat|tact-gtgcgag **cAcATccgTg** TTgTT **cAcggATgTg-3′**

(VHS881-2.1) 5′-gctgtgtat|tact-gtgcgag **cAcATccgTg** TTgTT **cAcggATgTg-3′**

(VHS881-4.1) 5′-gctgtgtat|tact-gtgcgag **cAcATccgTg** TTgTT **cAcggATgTg-3′**

(VHS881-9.1) 5′-gctgtgtat|tact-gtgcgag **cAcATccgTg** TTgTT **cAcggATgTg-3′**

| local de clivagem de substrato

(FOKIact) 5′-cAcATccgTg TTgTT cAcggATgtg-3′

(VHEx881) 5′-AATAgTAgAc TgcAgTgTcc TcAgcccTTA AgcTgTTcAT cTgcAAgTAg-

AgAgTCTTcT TAgAgTTgTc TcTAgAcTTA gTgAAgcg-3′

!note-se que CHEx881 é o complemento reverso do ON abaixo

! [RC] 5′-cgCttcacTaag-

```
!          Scab.......
!          3-23 Sintética como na Tabela 206
!          |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-
!          XbaI...
!          |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|t-3′
!          AflII...
(VHBA881)  5′-cgCttcacTaag-
           |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-
           |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt gcg ag-3′
(VHBB881)  5′-cgCttcacTaag-
           |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|-
           |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgt Acg ag-3′
(VH881PCR) 5′-cgCttcacTaag|TCT|AGA|gac|aac -3′
```

Tabela 17: Kappa, bases 12-30

!

! ID	Ntot	0	1	2	3	4	5	6	Nome Sequência.....Modo com pontos .....
! 1	84	40	21	20	1	2	0	0	SK12O12gacccagtctccatcctcc gaccagtctccatcctcc
! 2	32	19	3	6	2	1	0	1	SK12A17gactcagtctccactctcc ...t.....ct....
! 3	26	17	8	1	0	0	0	0	SK12A27gacgcagtctccaggcacc ...g.....gg.a..
! 4	40	21	18	1	0	0	0	0	SK12A11gacgcagtctccagccacc ...g.....g..a..
!	182	97	50	28	3	3	0	1	
!		97	147	175	178	181	181	182	

!

adaptadores URE:

```
!            Haste..... ansa. Haste..... Reconhecimento.....
(SzKB1230-O12)    5′-cAcATccgTg TTgTT cAcggATgTg ggAggATggAgAcTgggTc-3′
!         [RC] 5′-gacccagtctccatcctcc cAcATccgTg AAcAA cAcggATgTg-3′
!            Reconhecimento..... Haste..... ansa. Haste.....
!                           FokI.         FokI.
!
!            Haste.....ansa. Haste..... Reconhecimento.....
(SzKB1230-A17)    5′-cAcATccgTg TTgTT cAcggATgTg ggAgAgTggAgAcTgAgTc-3′
!         [RC] 5′-gactcagtctccactctcc cAcATccgTg AAcAA cAcggATgTg-3′
!            Reconhecimento..... Haste..... ansa. Haste.....
!                           FokI.         FokI.
!
!            Haste.....ansa. Haste..... Reconhecimento.....
(SzKB1230-A27)    5′-cAcATccgTg TTgTT cAcggATgTg ggTgccTggAgAcTgcgTc-3′
!         [RC] 5′-gacgcagtctccaggcacc cAcATccgtg AAcAA cAcggATgtg-3′
!            Reconhecimento..... Haste..... ansa. Haste.....
!                           FokI.         FokI.
!
```

```
!             Haste.......ansa. Haste...... Reconhecimento........
(SzKB1230-A11)    5′-cAcATccgTg TTgTT cAcggATgTg ggTggcTggAgAcTgcgTc-3′
!          [RC] 5′-gacgcagtctccagccacc cAcATccgtg AAcAA cAcggATgTg-3′
!             Reconhecimento....... Haste....... ansa. Haste......
!                           FokI.         FokI.
```

O que acontece na cadeia superior:

```
(SzKB1230-012*)             5′-gac cca gtc|tcc a-tc ctc c-3′
!                            |Local de clivagem no substrato
!
(SzKB1230-A17*)             5′-gac tct gtc|tcc a-ct ctc c-3′
!
(SzKB1230-A27*)             5′-gac gca gtc|tcc a-gg cac c-3′
!
(SzKB1230-A11*)             5′-gac gca gtc|tcc a-gc cac c-3′
```



```
(kapextURE)       5′-ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg-3′ !cadeia de sentido
                        Scab.............ApaLI.
(kapextUREPCR)    5′-ccTctactctTgTcAcAgtg-3′
                        Scab.............

(kaBR01UR)    5′-ggAggATggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA cAgg-3′
!      [RC]   5′-ccTctactcTgTcAcAgtgcAcAA gAc ATc cAg tcc a-tc ctc c-3′ ON acima é R.C. deste
(kaBR02UR)    5′-ggAgAgTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3′
!      [RC]   5′-ccTctactctTgTcAcAgTgcAcAA gAc ATC cAg tcc a-ct ctc c-3′ ON acima é R.C. deste
(kaBR03UR)    5′-ggTgccTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3′
!      [RC]   5′-ccTctactcTgTcAcAgTgcAcAA gAc ATC cAg tcc a-gg cac c-3′ 3′ ON acima é R.C. deste
(kaBR04UR)    5′-ggTggcTggA cTggATgTcT TgTgcAcTgT gAcAAgAgTA gAgg-3′
!      [RC]   5′-ccTctactctTgTcAcAgTgcAcAA gAc ATc cAg tcc a-gc cac c-3′ ON acima é R.C. deste
          Scab.............ApaLI.
```

Tabela 18 adaptadores de URE Lambda bases 13.3 a 19.3

```
!
!Número de sequências..... 128
!
!    Número de Emparelhadores Incorrectos.....
!Id  Ntot   0  1  2  3  4  5  6  7  8  Nome      Sequência.....Modo com pontos.....
! 1  58     45 7  1  0  0  0  2  2  1  VL133-2a2 gtctcctggacagtcgatc gtctcctggacagtcgatc
! 2  16     10 1  0  1  0  1  1  0  2  VL133-31  ggccttgggacagacagtc.g.cttg......a.ag..
! 3  17     6  0  0  0  4  1  1  5  0  VL133-2c  gtctcctggacagtcagtc  ............ag..
! 4  37     3  0  10 4  4  3  7  4  2  VL133-1c  ggccccagggcagagggtc.g.c..a..g...ag.g..
!    128    64 8  11 5  8  5  11 11 5
!          64  72 83 88 96 101 112 123 128
!
!          Haste.......ansa. Haste...... Reconhecimento........
(VL133-2a2)    5'-cAcATccgTg TTGTT cAcggATgTg gATcgAcTgTccAggAgAc-3'
!    [RC] 5'-gtctcctggacacagtcgatc cAcATccgTg AAcAA cAcggATgTg-3'
!          Reconhecimento....... Haste....... ansa. Haste......
!
!          Haste.......ansa. Haste...... Reconhecimento........
(VL133-31)         5'-cAcATccgTg RRgTT cAcggATgTg gAcTgTcTgTcccAAggcc-3'
!    [RC] 5'-gtctcctggacagtcgatc cAcATccgtg AAcAA cAcggATgTg-3'
!          Reconhecimento....... Haste....... ansa. Haste......
!
!          Haste.......ansa. Haste...... Reconhecimento........
(VL133-2c)         5'-cAcATccgTg TTgTT cAcggATgTg gAcTgAcTgTccAggAgac-3'
    [RC] 5'-gtctcctggacagtcagtc cAcATccgTg AAcAA cAcggATgTg-3'
!          Reconhecimento....... Haste....... ansa. Haste......
!
!          Haste.......ansa. Haste...... Reconhecimento........
(VL133-1c)  5'cAcATccgTg TTgTT cAcggATgTg gAcccTcTgcccTggggcc-3
```

o que acontece na cadeia de topo:

```
          !                 |local de clivagem na cadeia superior
(VL133-2a2*)   5'-g tct cct g|ga cag tcg atc
!
(VL133-3l*)    5'-g gcc ttg g|ga cag aca gtc
!
(VL133-2c*)    5'-g tct cct g|ga cag tca gtc
!
(VL133-1c*)    5'-g gcc caa g|gg cag agg gtc
!
!Os Extensores e Pontes seguintes codificam todos a Sequência AA de 2a2 para
codões 1-15
!                   1
(ON_LamEx133)  5'-ccTcTgAcTgAgT gcA cAg-
!
!      2   3   4   5   6   7   8   9   10  11  12
       AGt gcT TtA acC caA ccG gcT AGT gtT AgC ggT-
!
!      13 14 15
       tcC ccG g! 2a2
!                   1
(ON_LamB1-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg-
!
!      2 3 4 5 6 7 8 9 10 11 12
       AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-
!
!      13 14 15
       tcC ccG g ga cag tcg at-3'! 2a2 N.B. a seq real é o
!                                       complemento reverso
!                                       da mostrada.
!
(ON_LamB2-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg-
!
!      2 3 4 5 6 7 8 9 10 11 12
       AGt gcT TtA acC caA ccG gcT AGT gtT aGC ggT-
!
!      13 14 15
       tcC ccG g ga cag tca gt -3'!2c N.B a seq real é o
!                                       complemento reverso
!                                       da mostrada.
```

(ON_LamB3-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg-

!

! 2 3 4 5 6 7 8 9 10 11 12

AGt gcT TtA acC caA ccG gcT AGT gtT aGC ggT-

!

! 13 14 15

tcC ccG g ga cag tca gt -3'!2c ***N.B*** a seq real é o

! complemento reverso

! da mostrada.

!(ON_LamB4-133) [RC] 5'-ccTcTgAcTgAgT gcA cAg -

!

! 2 3 4 5 6 7 8 9 10 11 12

AGt gcT TtA acC caA ccG gcT AGT gtT AGC ggT-s

!

! 13 14 15

tcC ccG g gg cag agg gt-3' ! 1c ***N.B*** a seq real é o

! complemento reverso

! da mostrada.

!

(ON_Lam133PCR) 5'-ccTcTgAcTgAgT gcA cAg AGt gc-3'

Tabela 19: Clivagem de 75 cadeias leves humanas.

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
AfeI	AGCgct	0	0	
AflII	Cttaag	0	0	FR3HC
AgeI	Accggt	0	0	
AscI	GGcgcgcc	0	0	após LC
BglII	Agatct	0	0	
BsiWI	Cgtacg	0	0	
BspDI	ATcgat	0	0	
BssHII	Gcgcgc	0	0	
BstBI	TTcgaa	0	0	
DraIII	CACNNNgtg	0	0	
EagI	Cggccg	0	0	
FseI	GGCCGGcc	0	0	
FspI	TGCgca	0	0	
HpaI	GTTaac	0	0	
MfeI	Caattg	0	0	HC FR1
MluI	Acgcgt	0	0	

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
NcoI	Ccatgg	0	0	Sinal de cadeia pesada
NheI	Gctagc	0	0	HC/elemento de ligação em âncora
NotI	GCggccgc	0	0	no elemento de ligação após HC
NruI	TCGcga	0	0	
PacI	TTAATtaa	0	0	
PmeI	GTTTaaac	0	0	
PmlI	CACgtg	0	0	
PvuI	CGATcg	0	0	
SacII	CCGCgg	0	0	
SalI	Gtcgac	0	0	
SfiI	GGCCNNNNNnggcc	0	0	Sinal de Cadeia Pesada
SgfI	GCGATcgc	0	0	
SnaBI	TACgta	0	0	
StuI	AGGcct	0	0	
XbaI	Tctaga	0	0	HC FR3
AatII	GACGTc	1	1	
AclI	AAcgtt	1	1	
AseI	ATtaat	1	1	
BsmI	GAATGCN	1	1	
BspEI	Tccgga	1	1	HC FR1
BstXI	CCANNNNNTGG	1	1	HC FR2
DrdI	GACNNNNnngtc	1	1	
HindIII	Aagctt	1	1	
PciI	Acatgt	1	1	
SapI	gaagagc	1	1	
ScaI	AGTact	1	1	
SexAI	Accwggt	1	1	
SpeI	Actagt	1	1	
TliI	Ctcgag	1	1	
XhoI	Ctcgag	1	1	
BcgI	cgannnnnntgc	2	2	
BlpI	GCtnagc	2	2	
BssSI	Ctcgtg	2	2	
BstAPI	GCANNNNntgc	2	2	
EspI	GCtnagc	2	2	
KasI	Ggcgcc	2	2	
PflMI	CCANNNNntgg	2	2	
XmnI	GAANNnnttc	2	2	
ApaLI	Gtgcac	3	3	seq de sinal LC
NaeI	GCCggc	3	3	

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
NgoMI	Gccggc	3	3	
PvuII	CAGctg	3	3	
RsrII	CGgwccg	3	3	
BsrBI	GAGcgg	4	4	
BsrDI	GCAATGNNn	4	4	
BstZ17I	GTAtac	4	4	
EcoRI	Gaattc	4	4	
SphI	GCATGc	4	4	
SspI	AATatt	4	4	
AccI	GTmkac	5	5	
BclI	Tgatca	5	5	
BsmBI	Nnnnnngagacg	5	5	
BsrGI	Tgtaca	5	5	
OraI	TTTaaa	6	6	
NdeI	CAtatg	6	6	HC FR4
SwaI	ATTTaaat	6	6	
BamHI	Ggatcc	7	7	
SacI	GAGCTc	7	7	
BciVI	GTATCCNNNNNN	8	8	
BsaBI	GATNNnnatc	8	8	
NsiI	ATGCAt	8	8	
Bsp120I	Gggccc	9	9	CH1
ApaI	GGGCCc	9	9	CH1
PspOOM1	Gggccc	9	9	
BspHI	Tcatga	9	11	
EcoRV	GATatc	9	9	
AhdI	GACNNNnngtc	11	11	
BbsI	GAAGAC	11	14	
PsiI	TTAtaa	12	12	
BsaI	GGTCTCNnnnn	13	15	
XmaI	Cccggg	13	14	
AvaI	Cycgrg	14	16	
BglI	GCCNNNNnggc	14	17	
AlwNI	CAGNNNctg	16	16	
BspMI	ACCTGC	17	19	
XcmI	CCANNNNNnnnntgg	17	26	
BstEII	Ggtnacc	19	22	HC FR4
Sse8387I	CCTGCAgg	20	20	
AvrII	Cctagg	22	22	
HincII	GTYrac	22	22	

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
BsgI	GTGCAG	27	29	
MscI	TGGcca	30	34	
BseRI	NNnnnnnnnnctcctc	32	35	
Bsu36I	CCtnagg	35	37	
PstI	CTGCAg	35	40	
EciI	nnnnnnnnntccgcc	38	40	
PpuMI	RGgwccy	41	50	
StyI	Ccwwgg	44	73	
EcoO109I	RGgnccy	46	70	
Acc65I	Ggtacc	50	51	
KpnI	GGTACc	50	51	
BpmI	ctccag	53	82	
AvaII	Ggwcc	71	124	

* a clivagem ocorre na cadeia de topo depois da última base de letra maiúscula. Para RE que cortam sequências palindrómicas, a cadeia inferior corta-se no local simétrico.

Tabela 20: Clivagem de 79 cadeias pesadas

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
AfeI	AGCgct	0	0	
AflII	**Cttaag**	**0**	**0**	**HC FR3**
AscI	**GGcgcgcc**	**0**	**0**	**após LC**
BsiWI	Cgtacg	0	0	
BspDI	ATcgat	0	0	
BssHII	Gcgcgc	0	0	
FseI	GGCCGGcc	0	0	
HpaI	GTTaac	0	0	
Nhel	**Gctagc**	**0**	**0**	**Elemento de ligação de HC**
NotI	**Gcggccgc**	**0**	**0**	**No elemento de ligação, HC/âncora**
NruI	TCGcga	0	0	
NsiI	ATGCAt	0	0	
PacI	TTAATtaa	0	0	
PciI	Acatgt	0	0	
PmeI	GTTTaaac	0	0	
PvuI	CGATcg	0	0	
RsrII	CGgwccg	0	0	
SapI	gaagagc	0	0	
SfiI	**GGCCIJNNNnggcc**	**0**	**0**	**seq de sinal HC**
SgfI	GCGATcgc	0	0	

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
SwaI	ATTTaaat	0	0	
AclI	AAcgtt	1	1	
AgeI	Accggt	1	1	
AseI	ATtaat	1	1	
AvrII	Cctagg	1	1	
BsmI	GAATGCN	1	1	
BsrBI	GAGcgg	1	1	
BsrDI	GCAATGNNn	1	1	
DraI	TTTaaa	1	1	
FspI	TGCgca	1	1	
HindIII	Aagctt	1	1	
MfeI	**Caattg**	**1**	**1**	**HC FR1**
NaeI	GCCggc	1	1	
NgoMI	Gccggc	1	1	
SpeI	Actagt	1	1	
Acc65I	Ggtacc	2	2	
BstBI	TTcgaa	2	2	
KpnI	GGTACc	2	2	
MluI	Acgcgt	2	2	
NcoI	**Ccatgg**	**2**	**2**	**Em seq de sinal HC**
NdeI	**CAtatg**	**2**	**2**	**HC FR4**
PmlI	CACgtg	2	2	
XcmI	CCANNNNNnnnntgg	2	2	
BcgI	cgannnnnntgc	3	3	
BclI	Tgatca	3	3	
BglI	GCCNNNNnggc	3	3	
BsaBI	GATNNnnatc	3	3	
BsrGI	Tgtaca	3	3	
SnaBI	TACgta	3	3	
Sse8387I	CCTGCAgg	3	3	
ApaLI	**Gtgcac**	**4**	**4**	**Sinal LC /FR1**
BspHI	Tcatga	4	4	
BssSI	Ctcgtg	4	4	
PsiI	TTAtaa	4	5	
SphI	GCATGc	4	4	
AhdI	GACNNNnngtc	5	5	
BspEI	**Tccgga**	**5**	**5**	**HC FR1**
MscI	TGGcca	5	5	
SacI	GAGCTc	5	5	
ScaI	AGTact	5	5	

Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
SexAI	Accwggt	5	6	
SspI	AATatt	5	5	
TliI	Ctcgag	5	5	
XhoI	Ctcgag	5	5	
BbsI	GAAGAC	7	8	
BstAPI	GCANNNNntgc	7	8	
BstZ17I	GTAtac	7	7	
EcoRV	GATatc	7	7	
EcoRI	Gaattc	8	8	
BlpI	GCtnagc	9	9	
Bsu36I	CCtnagg	9	9	
DraIII	CACNNNgtg	9	9	
EspI	GCtnagc	9	9	
StuI	AGGcct	9	13	
XbaI	**Tctaga**	**9**	**9**	**HC FR3**
Bsp120I	**Gggccc**	**10**	**11**	**CH1**
ApaI	**GGGCCc**	**10**	**11**	**CH1**
PspOOMI	Gggccc	10	11	
BciVI	GTATCCNNNNNN	11	11	
SalI	Gtcgac	11	12	
DrdI	GACNNNNnngtc	12	12	
KasI	Ggcgcc	12	12	
XmaI	Cccggg	12	14	
BglII	Agatct	14	14	
HincII	GTYrac	16	18	
BamHI	Ggatcc	17	17	
PflMI	CCANNNNntgg	17	18	
BsmBI	Nnnnnngagacg	18	21	
BstXI	**CCANNNNNntgg**	**18**	**19**	**HC FR2**
XmnI	GAANNnnttc	18	18	
SacII	CCGCgg	19	19	
PstI	CTGCAg	20	24	
PvuII	CAGctg	20	22	
AvaI	Cycgrg	21	24	
EagI	Cggccg	21	22	
AatII	GACGTc	22	22	
BspMI	ACCTGC	27	33	
AccI	GTmkac	30	43	
StyI	Ccwwgg	36	49	
AlwNI	CAGNNNctg	38	44	

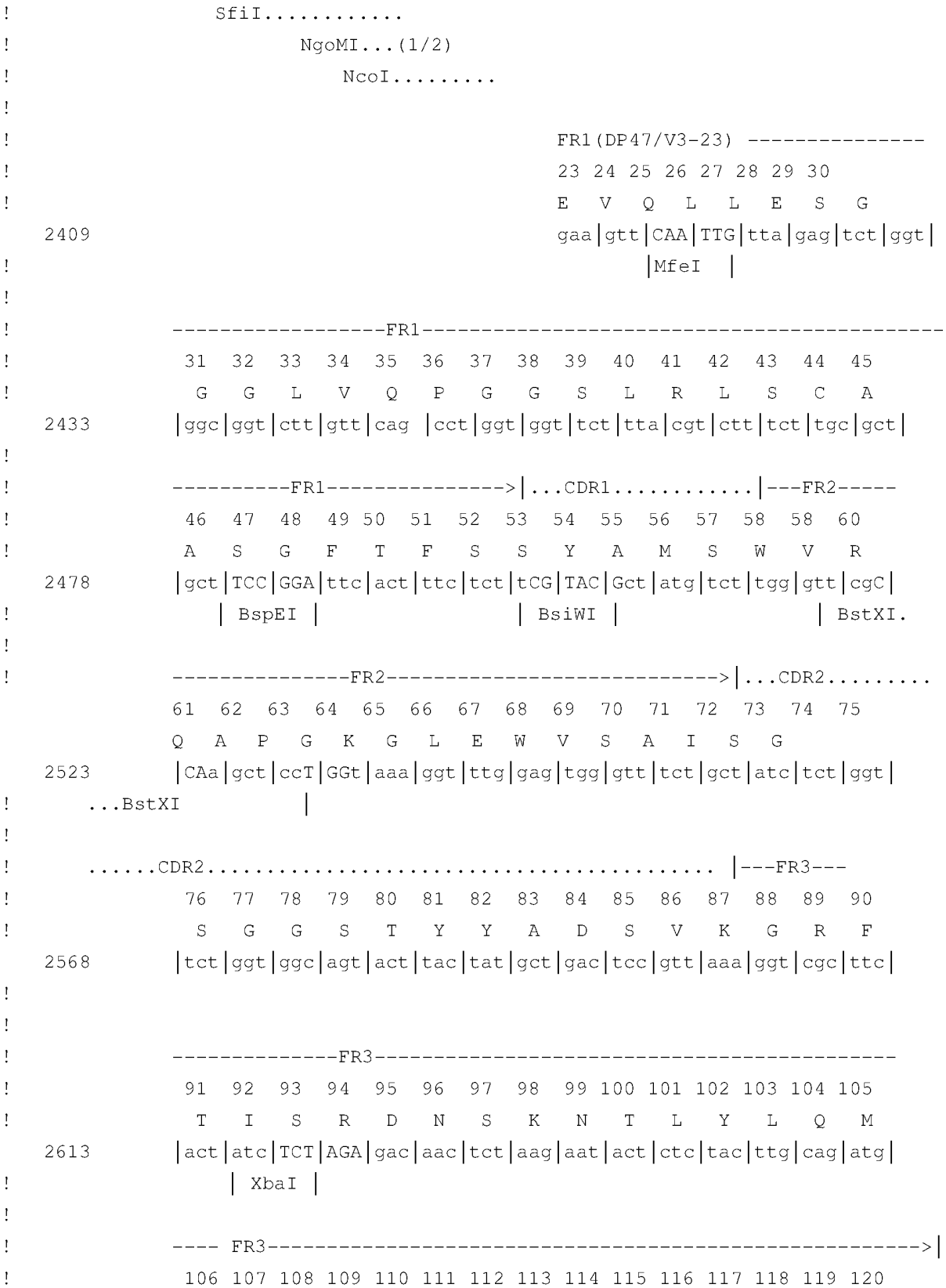
Enzima	Reconhecimento*	Nch	Ns	Posição planeada do local
BsaI	GGTCTCNnnnn	38	44	
PpuMI	RGgwccy	43	46	
BsgI	GTGCAG	44	54	
BseRI	NNnnnnnnnnctcctc	48	60	
EciI	nnnnnnnnntccgcc	52	57	
BstEII	**Ggtnacc**	**54**	**61**	**HC Fr4, 47/79 têm um**
EcoO109I	RGgnccy	54	86	
BpmI	ctccag	60	121	
AvaII	Ggwcc	71	140	

Tabela 21: MALIA3, anotada

```
! 9532 bases de MALIA3
! ------------------------------------------------------------------------------
    1 aat act gct act att agt aga att gat gcc acc ttt tca gct cgc gcc
!   gene ii continua
   49 cca aat gaa aat ata gct aaa cag gtt att gac cat ttg cga aat gta
   97 tct aat ggt caa act aaa tct act cgt tcg cag aat tgg gaa tca act
  145 gtt aca tgg aat gaa act tcc aga cac cgt act tta gtt gca tat tta
  193 aaa cat gtt gag cta cag cac cag att cag caa tta agc tct aag cca
  241 tcc gca aaa atg acc tct tat caa aag gag caa tta aag gta ctc tct
  289 aac cct gac ctg ttg gag ttt gct tcc ggt ctg gtt cgc ttt gaa gct
  337 cga att aaa acg cga tat ttg aag tct ttc ggg ctt cct ctt aat ctt
  385 ttt gat gca atc cgc ttt gct tct gac tat aat agt cag ggt aaa gac
  433 ctg att ttt gat tta tgg tca ttc tcg ttt tct gaa ctg ttt aaa gca
  481 ttt gag ggg gat tca ATG aat att tat gac gat tcc gca gta ttg gac
!         RBS?.....         início do gene x, ii continua
  529 gct atc cag tct aaa cat ttt act att acc ccc tct ggc aaa act tct
  577 ttt gca aaa gcc tct cgc tat ttt ggt ttt tat cgt cgt gtg gta aac
  625 gag ggt tat gat agt gtt gct ctt act atg cct cgt aat tcc ttt tgg
  673 cgt tat gta tct gca tta gtt gaa tgt ggt att cct aaa tct caa ctg
  721 atg aat ctt tct acc tgt aat aat gtt gtt ccg tta gtt cgt ttt att
  769 aac gta gat ttt tct tcc caa cgt cct gac tgg tat aat gag cca gtt
  817 ctt aaa atc gca TAA
!                   Fim de X & II
  832 ggtaattca ca
!         M1                E5                Q10               T15
  843 ATG att aaa gtt gaa att aaa cca tct caa gcc caa ttt act act cgt
!     Início do gene V
!
```

```
!         S17               S20               P25               E30
   891 tct ggt gtt tct cgt cag ggc aag cct tat tca ctg aat gag cag ctt
!
!               V35                  E40              V45
   939 tgt tac gtt gat ttg ggt aat gaa tat ccg gtt ctt gtc aag att act
!
!         D50                      A55              L60
   987 ctt gat gaa ggt cag cca gcc tat gct cct ggt cTG TAC Acc gtt cat
!                                                     BsrGI...
!      L65                     V70               S75               R80
   1035   ctg tcc tct tta aaa gtt ggt cag ttc ggt tcc ctt atg att gac cgt
!
!                           P85        K87 Fim de V
   1083 ctg cgc ctc gtt ccg gct aag TAA C
!
!  1108 ATG gag cag gtc gcg gat ttc gac aca att tat cag gcg atg
!      Início do gene VII
!
   1150 ata caa atc tcc gtt gta ctt tgt ttc gcg ctt ggt ata atc
!
!                              VII e IX sobrepostos.
!                              ..... S2 V3 L4 V5                  S10
   1192 gct ggg ggt caa agA TGA gt gtt tta gtg tat tct ttc gcc tct ttc gtt
!                                    Fim de VIII
!                                  | Início de IX
!         L13        W15                      G20        T25        E29
   1242 tta ggt tgg tgc ctt cgt agt ggc att acg tat ttt acc cgt tta atg gaa
!
   1393 act tcc tc
!
!      ....paragem de IX. IX e VIII sobrepostos por quatro bases
   1301 ATG aaa aag tct tta gtc ctc aaa gcc tct gta gcc gtt gct acc ctc
!       Início da Sequência de sinal de viii.
!
!  1349 gtt ccg atg ctg tct ttc gct gct gag ggt gac gat ccc gca aaa gcg
!                                  VIII maduro --->
   1397 gcc ttt aac tcc ctg caa gcc tca gcg acc gaa tat atc ggt tat gcg
   1445 tgg gcg atg gtt gtt gtc att
   1466 gtc ggc gca act atc ggt atc aag ctg ttt aag
   1499 aaa ttc acc tcg aaa gca ! 1515
!         .............. -35    ..
```

```
!
   1517       agc tga taaaccgat acaattaaag gctccttttg
!                       ..... -10  ...
!
   1552 gagccttttt ttttGGAGAt ttt ! S.D. sublinhado
!
!          <------ Sequência de sinal III ------------------------------------>
!              M   K   K   L   L   F   A   I   P   L   V
   1575 caac GTG aaa aaa tta tta ttc gca att cct tta gtt ! 1611
!
!        V   P   F   Y   S   H   S   A   Q
   1612 gtt cct ttc tat tct cac aGT gcA Cag TCT
!                           ApaLI...
!
   1642       GTC GTG ACG CAG CCG CCC TCA GTG TCT GGG GCC CCA GGG CAG
              AGG GTC ACC ATC TCC TGC ACT GGG AGC AGC TCC AAC ATC GGG GCA
!                 BstEII...
   1729       GGT TAT GAT GTA CAC TGG TAC CAG CAG CTT CCA GGA ACA GCC CCC AAA
   1777       CTC CTC ATC TAT GGT AAC AGC AAT CGG CCC TCA GGG GTC CCT GAC CGA
   1825       TTC TCT GGC TCC AAG TCT GGC ACC TCA GCC TCC CTG GCC ATC ACT
   1870       GGG CTC CAG GCT GAG GAT GAG GCT GAT TAT
   1900       TAC TGC CAG TCC TAT GAC AGC AGC CTG AGT
   1930       GGC CTT TAT GTC TTC GGA ACT GGG ACC AAG GTC ACC GTC
!                                                    BstEII...
   1969       CTA GGT CAG CCC AAG GCC AAC CCC ACT GTC ACT
   2002       CTG TTC CCG CCC TCC TCT GAG GAG CTC CAA GCC AAC AAG GCC AGA CTA
   2050       GTG TGT CTG ATC AGT GAC TTC TAC CCG GGA GCT GTG ACA GTG GCC TGG
   2098       AAG GCA GAT AGC AGC CCC GTC AAG GCG GGA GTG GAG ACC ACC ACA CCC
   2146       TCC AAA CAA AGC AAC AAC AAG TAC GCG GCC AGC AGC TAT CTG AGC CTG
   2194       ACG CCT GAG CAG TGG AAG TCC CAC AGA AGC TAC AGC TGC CAG GTC ACG
   2242       CAT GAA GGG AGC ACC GTG GAG AAG ACA GTG GCC CCT ACA GAA TGT TCA
   2290       TAA TAA ACCG CCTCCACCGG GCGCGCCAAT TCTATTTCAA GGAGACAGTC ATA
!                               AscI.....
!
!             sinal PelB ---------------------------------------------->
!              M   K   Y   L   L   P   T   A   A   A   G   L   L   L   L
   2343       ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC
!
!             16  17  18  19   20   21  22
!              A   A   Q   P   A    M   A
   2388       gcG GCC cag ccG GCC   atg gcc
```


SfiI.....
NgoMI...(1/2)
NcoI.....
FR1(DP47/V3-23) ----------
23 24 25 26 27 28 29 30
E V Q L L E S G
2409 gaa|gtt|CAA|TTG|tta|gag|tct|ggt|
|MfeI |
----------FR1----------
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
G G L V Q P G G S L R L S C A
2433 |ggc|ggt|ctt|gtt|cag |cct|ggt|ggt|tct|tta|cgt|ctt|tct|tgc|gct|
----------FR1---------->|...CDR1.....|---FR2-----
46 47 48 49 50 51 52 53 54 55 56 57 58 58 60
A S G F T F S S Y A M S W V R
2478 |gct|TCC|GGA|ttc|act|ttc|tct|tCG|TAC|Gct|atg|tct|tgg|gtt|cgC|
| BspEI | | BsiWI | | BstXI.
----------FR2----------->|...CDR2.....
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
Q A P G K G L E W V S A I S G
2523 |CAa|gct|ccT|GGt|aaa|ggt|ttg|gag|tgg|gtt|tct|gct|atc|tct|ggt|
...BstXI |
.....CDR2..... |---FR3---
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
S G G S T Y Y A D S V K G R F
2568 |tct|ggt|ggc|agt|act|tac|tat|gct|gac|tcc|gtt|aaa|ggt|cgc|ttc|
----------FR3----------
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
T I S R D N S K N T L Y L Q M
2613 |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|
| XbaI |
---- FR3----------
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120

```
!                N   S   L   R   A   E   D   T   A   V   Y   Y   C   A   K
   2658        |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|
                  | AflII |               | PstI |
!
!
!       ......CDR3..................... |---FR4-----------------------------
!            121 122 123 124 125 126 127 128 128 130 131 132 133 134 135
              D   Y   E   G   T   G   Y   A   F   D   I   W   G   Q   G
   2703     |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ATA|TGg|ggt|caa|ggt|
!                                                 |NdeI | (1/4)
!
!              ----------------FR3----------->|
!            136 137 138 139 140 141 142
!             T   M   V   T   V   S   S
   2748     |act|atG|GTC|ACC|gtc|tct|agt
!                 | BstEII |
! Da frente de BstEII, pV323 é o mesmo que PCEs1, excepto como referido.
!os locais BstEII podem ocorrer em cadeias leves; não como sendo únicas no
vector final
!.
!
!                               143 144 145 146 147 148 149 150 151 152
!                               A   S   T   K   G   P   S   V   F   P
   2769                        gcc tcc acc aaG GGC CCa tcg GTC TTC ccc
!                                           Bsp120I.      BbsI...(2/2)
!                                           ApaI....
!
!          153 154 155 156 157 158 159 160 161 162 163 164 165 166 167
!           L   A   P   S   S   K   S   T   S   G   G   T   A   A   L
   2799    ctg gca ccC TCC TCc aag agc acc tct ggg ggc aca gcg gcc ctg
!                 BseRI...(2/2)
!
!          168 169 170 171 172 173 174 175 176 177 178 179 180 181 182
!           G   C   L   V   K   D   Y   F   P   E   P   V   T   V   S
   2844    ggc tgc ctg GTC AAG GAC TAC TTC CCc gaA CCG GTg acg gtg tcg
!                                       AgeI....
!
!          183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
!           W   N   S   G   A   L   T   S   G   V   H   T   F   P   A
   2889    tgg aac tca GGC GCC ctg acc agc ggc gtc cac acc ttc ccg gct
!                    KasI...(1/4)
```

```
!
!          198 199 200 201 202 203 204 205 206 207 208 209 210 211 212
!           V   L   Q   S   S   G   L   Y   S   L   S   S   V   V   T
    2934   gtc cta cag tCt agc GGA ctc tact cc ctc agc agc gta gtg acc
!                     (Bsu36I...)(destruídos)
!
!          213 214 215 216 217 218 219 220 221 222 223 224 225 226 227
!           V   P   S   S   S   L   G   T   Q   T   T   I   C   N   V
    2979   gtg ccC tCt tct agc tTG Ggc acc cag acc tac atc tgc aac gtg
!                  (BstXI...........)N.B. destruição dos locais BstXi & BpmI.
!
!          228 229 230 231 232 233 234 235 236 237 238 239 240 241 242
!           N   H   K   P   S   N   T   K   V   D   K   K   V   E   P
    3024   aat cac aag ccc agc aac acc aag gtg gac aag aaa gtt gag ccc
!
!          243 244 245
!           K   S   C   A   A   A   H   H   H   H   H   H   S   A
    3069   aaa tct tgt GCG GCC GCt cat cac cac cat cat cac tct gct
!                        NotI.......
!
!           E   Q   K   L   I   S   E   E   D   L   N   G   A   A
    3111   gaa caa aaa ctc atc tca gaa gag gat ctg aat ggt gcc gca
!
!
!           D   I   N   D   D   R   M       A   S   G   A
    3153   GAT ATC aac gat gat cgt atg     gct AGC  ggc gcc
!          local de clivagem rEK...........NheI...    KasI...
!          EcoRV..
!
! Domínio 1 ------------------------------------------------------------------
!           A   E   T   V   E   S   C   L   A
    3183   gct gaa act gtt gaa agt tgt tta gca
!
!
!           K   P   H   T   E   I   S   F
    3210   aaa ccc cat aca gaa aat tca ttt
!
!           T   N   V   W   K   D   D   K   T
    3234   aCT AAC GTC TGG AAA GAC GAC AAA Act
!
!           L   D   R   Y   A   N   Y   E   G   C   L   W   N   A   T   G   V
```

```
   3261   tta gat cgt tac gct aac tat gag ggt tgt ctg tgG AAT GCt aca ggg gtt
!                                                      BsnI_____
!
!          V   V   C   T   G   D   E   T   Q   C   Y   G   T   W   V   P   I
   3312   gta gtt tgt act ggt GASC GAA ACT CAG TGT TAC GGT ACA TGG GTT cct att
!
!          G   L   A   I   P   E   N
   3363   ggg ctt gct act cct gaa aat
!
! elemento de ligação L1 -----------------------------------
!          E   G   G   G   S   E   G   G   G   S
   3384   gag ggt ggt ggc tct gag ggt ggc ggt tct
!
!          E   G   G   G   S   E   G   G   G   T
   3414   gag ggt ggt ggc tct gag ggt ggc ggt act
!
! Domínio 2 ------------------------------------------------------------------
   3444   aaa cct cct gag tac ggt gat aca cct att ccg ggc tat act tat atc aac
   3495   cct ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct
   3546   aat ctc tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag aat
!                      BseRI____
   3597   aat agg ttc cga aat agg cag ggg gca tta act gtt tat acg ggc act
   3645   gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act cct
   3693   gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttC AGA
!                                                                 AlwNI
   3741   GAC TGc gct ttc cat tct ggc ttt aat gaa gat cca ttc gtt tgt gaa
!             AlwNI
   3789   tat caa ggc caa tcg tct gac ctg cct ca acct cct gtc aat gct
!
   3834   ggc ggc ggc tct
! Início de L2 ---------------------------------------------------------------
   3846   ggt ggt ggt tct
   3858   ggt ggt ggt tct
   3870   gag ggt ggt ggc tct gag ggt ggc ggt tct
   3900   gag ggt ggt ggc tct gag ggt ggc ggt tct
   3930   ggt ggt ggt tct ggt   !   Fim de L2
!
! Domínio 3 ------------------------------------------------------------------
!          S   G   D   F   D   Y   E   K   M   A   N   A   N   K   G   A
   3945   tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat aag ggg gct
!
```

```
!          M   T   E   N   A   D   E   N   A   L   Q   S   D   A   K   G
    3993   atg acc gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc
!
!          K   L   D   S   V   A   T   D   Y   G   A   A   I   D   G   F
    4041   aaa ctt gat tct gtc gct act gat tac ggt gct gct atc gat ggt ttc
!
!          I   G   D   V   S   G   L   A   N   G   N   G   A   T   G   D
    4089   att ggt gac gtt tcc ggc ctt gct aac ggt aat ggt gct act ggt gat
!
!          F   A   G   S   N   S   Q   M   A   Q   V   G   D   G   D   N
    4137   ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt gat aat
!
!          S   P   L   M   N   N   F   R   Q   Y   L   P   S   L   P   Q
    4185   tca cct tta atg aat aat ttc agt caa tat tta cct tcc ctc cct caa
!
!          S   V   E   C   R   P   F   V   F   S   A   G   K   P   Y   E
    4233   tcg gtt gaa tgt cgc cct ttt gtc ttt agc gct ggt aaa cca tat gaa
!          F   S   I   D   C   D   K   I   N   L   F   R
    4281   ttt tct att gat tgt gac aaa ata aac tta ttc cgt
!                                                  Fim do Domínio 3
!
!          G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V F140
    4317   ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat gta ttt
!          início do segmento transmembranar
!
!          S   T   F   A   N   I   L
    4365   tct acg ttt gct aac ata ctg
!
!          R   N   K   E   S
    4386   cgt aat aag gag tct TAA ! paragem de iii
!      Âncora intracelular
!
!             M1  P2  V   L   L5   G   I   P   L  L10  L  R   F   L  G15
    4404   tc ATG cca gtt ctt ttg ggt att ccg tta tta ttg cgt ttc ctc ggt
!             Início de VI
!
    4451   ttc ctt ctg gta act ttg ttc ggc tat ctg ctt act ttt ctt aaa aag
    4499   ggc ttc ggt aag ata gct att gct att tct ttg ttt ctt gct ctt att
    4547   att ggg ctt aac tca att ctt gtg ggt tat ctc tct gat att agc gct
    4595   caa tta ccc tct gac ttt gtt cag ggt gtt cag tta att ctc ccg tct
    4643   aat gcg ctt ccc tgt ttt tat gtt att ctc tct gta aag gct gct att
```

```
   4691   ttc att ttt gac gtt aaa caa aaa atc gtt tct tat ttg gat tgg gat
!
!                    M1  A2  V3     F5             L10        G13
   4739   aaa TAA t ATG gct gtt tat ttt gta act gcc aaa tta ggc tct gga
!            Fim de VI      Início do gene I
!
!          14  15  16  17  18  19  20  21  22  23  24  25  26  27  28
!           K   T   L   V   S   V   G   K   I   Q   D   K   I   V   A
   4785   aag acg ctc gtt agc gtt ggt aag att cag gat aaa att gta gct
!
!          29  30  31  32  33  34  35  36  37  38  39  40  41  42  43
!          G   C   K   I   A   T   N   L   D   L   R   L   Q   N   L
   4830   ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa aac ctc
!
!          44  45  46  47  48  49  50  51  52  53  54  55  56  57  58
!           P   Q   V   G   R   F   A   K   T   P   R   V   L   R   I
   4875   ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga ata
!
!          59  60  61  62  63  64  65  66  67  68  69  70  71  72  73
!           P   D   K   P   S   I   S   D   L   L   A   I   G   R   G
   4920   ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt
!
!          74  75  76  77  78  79  80  81  82  83  84  85  86  87  88
!           N   D   S   Y   D   E   N   K   N   G   L   L   V   L   D
   4965   aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat
!
            89  90  91  92  93  94  95  96  97  98  99 100 101 102 103
!           E   C   G   T   W   F   N   T   R   S   W   N   D   K   E
   5010   gag tgc gtg act tgg ttt aat acc cgt tct tgg aat gat aag gaa
!           104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
!            R   Q   P   I   I   D   W   F   L   H   A   R   K   L   G
   5055    aga cag ccg att att gat tgg ttt cta cat gct cgt aaa tta gga
!
!           119 120 121 122 123 124 125 126 126 128 129 130 131 132 133
!            W   D   I   I   F   L   V   Q   D   L   S   I   V   D   K
   5100   tgg gat att att ttt ctt gtt cag gac tta tct att gtt gat aaa
!
!          134 135 136 137 138 139 140 141 142 143 144 145 146 147 148
!           Q   A   R   S   A   L   A   E   H   V   V   Y   C   R   R
   5145   cag gcg gct tct gca tta gct gaa cat gtt gtt tat tgt cgt cgt
!
```

```
!          149 150 151 152 153 154 155 156 157 158 159 160 161 162 163
!           L   D   R   I   T   L   P   F   V   G   T   L   Y   S   L
    5190   ctg gac aga att act tta cct ttt gtc ggt act tta tat tct ctt
!
!          164 165 166 167 168 169 170 171 172 173 174 175 176 177 178
!           I   T   G   S   K   M   P   L   P   K   L   H   V   G   V
    5235   att act ggc tcg aaa atg cct ctg cct aaa tta cat gtt ggc gtt
!          179 180 181 182 183 184 185 186 187 188 189 190 191 192 193
!           V   K   Y   G   D   S   Q   L   S   P   T   V   E   R   W
    5280   gtt aaa tat ggc gat tct caa tta agc cct act gtt gag cgt tgg
!
!          194 195 196 197 198 199 200 201 202 203 204 205 206 207 208
!           L   Y   T   G   K   N   L   Y   N   A   Y   D   T   K   Q
    5325   ctt tat act ggt aag aat ttg tat aac gca tat gat act aaa cag
!
!          209 210 211 212 213 214 215 216 217 218 219 220 221 222 223
!           A   F   S   S   N   Y   D   S   G   V   Y   S   Y   L   T
    5370 gct ttt  tct agt aat tat gat tcc ggt gtt tat tct tat tta acg
!
!          224 225 226 227 228 229 230 231 232 233 234 235 236 237 238
!           P   Y   L   S   H   G   R   Y   F   K   P   L   N   L   G
    5415   cct tat tta tca cac ggt cgg tat ttc aaa cca tta aat tta ggt
!
!          239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
!           Q   K   M   K   L   T   K   I   Y   L   K   K   F   S   R
    5460 cag aag atg aaa tta act aaa ata tat ttg aaa aag ttt tct cgc
!
!          254 255 256 257 258 259 260 261 262 263 264 265 266 267 268
!           V   L   C   L   A   I   G   F   A   S   F   T   Y   S
    5505   gtt ctt tgt ctt gcg att gga ttt gca tca gca ttt aca tat agt
!
!          269 270 271 272 273 274 275 276 277 278 279 280 281 282 283
!           Y   I   T   Q   P   K   P   E   V   K   K   V   V   S   Q
    5550   tat ata acc ca acct aag ccg gag gtt aaa aag gta gtc tct cag
!
!          284 285 286 287 288 289 290 291 292 293 294 295 296 297 298
!           T   Y   D   F   D   K   F   T   I   D   S   S   Q   R   L
    5595   acc tat gat ttt gat aaa ttc act att gac tct tct cag cgt ctt
!
!          299 300 301 302  303 304 305 306 307 308 309 310 311 312 313
!           N   L   S   Y   R   Y   V   F   K   D   S   K   G   K   L
```

```
   5640   aat cta agc tat cgc tat gtt ttc aag gat tct aag gga aaa TTA
!                                                                PacI
!         314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
!          I   N   S   D   D   L   Q   K   Q   G   Y   S   L   T   Y
   5685   ATT AAT agc gac gat tta cag aac caa ggt tat tca ctc aca tat
!     PacI
!
!         329 330 331 332 333 334 335 336 337 338 339 340 341 342 343
!          i   I   D   L   C   T   V   S   I   K   K   G   G   N   S   N   E
!     iv                                                   M1  K
   5730   att gat tta tgt act gtt tcc att aaa aaa ggt aat tca aAT Gaa
!                                                       início de IV
!
!             344 345 346 347 348 349
!     i        I   V   K   C   N    .Fim de I
!     iv        L3  L  N5  V   I7   N   F  V10
   5775      att gtt aaa tgt aat TAA T TTT GTT
!     IV continua.....
   5800   ttc ttg atg ttt gtt tca tca tct tct ttt gct cag gta att gaa atg
   5848   aat aat tcg cct ctg cgc gat ttt gta act tgg tat tca aag caa tca
   5896   ggc gaa tcc gtt att gtt tct ccc gat gta aaa ggt act gtt act gta
   5944   tat tca tct gac gtt aaa cct gaa aat cta cgc aat ttc ttt att tct
   5992   gtt tta cgt gct aat aat ttc gat atg gtt ggt tca att cct tcc ata
   6040   att cag aag tat aat cca aac aat cag gat tat att gat gaa ttg cca
   6088   tca tct gat aat cag gaa tat gat gat aat tcc gct cct tct ggt ggt
   6136   ttc ttt gtt ccg caa aat gat aat gtt act caa act ttt aa att aat
   6184   aac gtt cgg gca aag gat tta ata cga gtt gtc gaa ttg ttt gta aag
   6232   tct aat act tct aaa tcc tca aat gta tta tct att gac ggc tct aat
   6280   cta tta gtt gtt TCT gca cct aaa gat att tta gat aac ctt cct caa
!                         ApaLI removido
   6328   ttc ctt tct act gtt gat ttg cca act gac cag ata ttg att gag ggt
   6376   ttg ata ttt gag gtt cag caa ggt gat gct tta gat ttt tca ttt gct
   6424   gct ggc tct cag cgt ggc act gtt gca ggc ggt gtt aat act gac cgc
   6472   ctc acc tct gtt tta tct tct gct ggt ggt tcg ttc ggt att ttt aat
   6520   ggc gat gtt tta ggg cta tca gtt cgc gca tta aag act aat agc cat
   6568   tca aaa ata ttg tct gtg cca cgt att ctt acg ctt tca ggt cag aag
   6616   ggt tct atc tct gtT GGC CAg at gtc cct ttt att act ggt cgt gtg
!                         MscI____
   6664   act ggt gaa tct gcc aat gta aat aat cca ttt cag acg att gag cgt
   6712   caa aat gta ggt att tcc atg agc gtt ttt cct gtt gca atg gct ggc
   6760   ggt aat att gtt ctg gat att acc agc aag gcc gat agt ttg agt tct
```

```
   6808   tct act cag gca agt gat gtt att act aat caa aa agt att gct aca
   6856   acg gtt aat ttg cgt gat gga cag act ctt tta ctc ggt ggc ctc act
   6904   gat tat aaa aac act tct caa gat tct ggc gta ccg ttc ctg tct aaa
   6952   atc cct tta atc ggc ctc ctg ttt agc tcc cgc tct gat tcc aac gag
   7000   gaa agc acg tta tac gtg ctc gtc aaa gca acc ata gta cgc gcc ctg
   7048   TAG cggcgcatt
!         Fim de IV
   7060   aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc
   7120   gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcGCCGGCt ttccccgtca
!                                                      NgoMI_
   7180   agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc
   7240   caaaaaactt gatttgggtg atggttCACG TAGTGggcca tcgccctgat agacggtttt
!                                    DraIII_____
   7300   tcgcccttTG ACGTTGGAGT Ccacgttctt taatagtgga ctcttgttcc aaactggaac
!               DrdI_____________
   7360   aacactcaac cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcgga
   7420   accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa
   7480   ctctctcagg gccaggcggt gaagggcaat CAGCTGttgc cCGTCTCact ggtgaaaaga
!                                       PvuII.     BsmBI.
   7540   aaaaccaccc tGGATCC    AAGCTT
!                   BamHI      HindIII ( )

!              Inserir portador de gene bla
   7563       gcaggtg gcacttttcg gggaaatgtg cgcggaaccc
   7600   ctatttgttt atttttctaa atacattcaa atatGTATCC gctcatgaga caataaccct
!                                          BciVI
   7660   gataaatgct tcaataatat tgaaaaAGGA AGAgt
!                                    RBS.?...
!         início do gene bla
   7695   ATG agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca ttt
   7746   tgc ctt cct gtt ttt gct cac cca gaa acg ctg aaa gta aaa gat gct
   7797   gaa gat cag ttg ggC gCA CGA Gtg ggt tac atc gaa ctg gat ctc aac agc
!                            BssSI...
!                        ApaLI removido
   7848   ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc
   7899   act ttt aaa gtt ctg cta tgt cat aca cta tta tcc cgt att gac gcc ggg
   7950   caa gaG CAA CTC GGT CGc cgg gcg cgg tat tct cag aat gac ttg gtt gAG
!             BcgI_____________                                          ScaI
   8001   TAC Tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa
!         ScaI_
   8052   tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt
```

```
    8103   ctg aca aCG AC Gga gga ccg aag gag cta acc gct ttt ttg cac aac atg
!                    PvuI_______
    8154   ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc
    8205   ata cca aac gac gag cgt gac acc acg atg cct gta gac atg cca aca acg
    8256   tTG CGC Aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa caa
!              FspI....
!
    8307   tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg
    8358   GCC ctt ccG GCt ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt
!          Bg1I_____________
    8409   gGG TCT Cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt
!           BsaI________
    8460   atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa cga aat
!                              AhdI_________
    8511   aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg TAA ctgt
!                                                                  paragem
    8560   cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa
    8620   ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt
    8680   cgttccactg tacgtaagac cccc
    8704   AAGCTT     GTCAC tgaa tggcgaatgg cgctttgcct
!          HinfIII    SalI..
!          (2/2)      HincII
    8740   ggtttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt
!
    8790   CCTGAGG
!          Bsu36I_
    8797           ccgat actgtcgtcg tcccctcaaa ctggcagatg
    8832   cacggttacg atgcgcccat ctacaccaac gtaacctatc ccattacggt caatccgccg
    8892   tttgttccca cggagaatcc gacgggttgt tactcgctca catttaatgt tgatgaaagc
    8952   tggctacagg aaggccagac gcgaattatt tttgatggcg ttcctattgg ttaaaaaatg
    9012   agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaATTTAAA
!                                                                  SwaI...
    9072   Tatttgctta tacaatcttc ctgtttttgg ggcttttctg attatcaacc GGGGTAcat
!                                                                  RBS?
    9131   ATG att gac atg cta gtt tta cga tta ccg ttc atc gat tct ctt gtt tgc
!          início do gene II
    9182   tcc aga ctc tca ggc aat gac ctg ata gcc ttt gtA GAT CTc tca aaa ata
!                                                        BglII...
    9233   gct acc ctc tcc ggc atg aat tta tca gct aga acg gtt gaa tat cat att
    9284   gat ggt gat ttg act gtc tcc ggc ctt tct cac cct ttt gaa tct tta cct
    9335   aca cat tact ca ggc att gca ttt aaa ata tat gag ggt tct aaa aat tt
```

9386 tat cct tgc gtt gaa ata aag gct tct ccc gca aaa gta tta cag ggt cat
9437 aat gtt ttt ggt aca acc gat tta tgc tct gag gct tta ttg ctt
9488 aat ttt gct aat tct ttg cct tgc ctg tat gat tta ttg gat gtt ! 9532
! gene II continua

Tabela 21 B: Sequência de MALIA3, condensada

LOCAL MALIA3 9532 CIRCULAR
ORIGEM

1 AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT
61 ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT
121 CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA
181 GTTGCATATT TAAAACATGT TGAGCTAGAG CACCAGATTC AGCAATTAAG CTCTAAGCCA
241 TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG
361 TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT
421 CAGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAGCA
481 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT
541 AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT
601 GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT
661 AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG
721 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT
781 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA
841 CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT
901 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG
961 AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC
1021 TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC
1081 GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT
1141 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT
1201 CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA
1261 GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAGT CTTTAGTCCT
1321 CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGTG CTGAGGGTGA
1381 CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA
1441 TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA
1501 ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT
1561 TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC
1621 TATTCTCACA GTGCACAGTC TGTCGTGACG CAGCCGCCCT CAGTGTCTGG GGCCCCAGGG
1681 CAGAGGGTCA CCATCTCCTG CACTGGGAGC AGCTCCAACA TCGGGGCAGG TTATGATGTA
1741 CACTGGTACC AGCAGCTTCC AGGAACAGCC CCCAAACTCC TCATCTATGG TAACAGCAAT
1801 CGGCCCTCAG GGGTCCCTGA CCGATTCTCT GGCTCCAAGT CTGGCACCTC AGCCTCCCTG
1861 GCCATCACTG GGCTCCAGGC TGAGGATGAG GCTGATTATT ACTGCCAGTC CTATGACAGC

1921 AGCCTGAGTC GCCTTTATGT CTTCGGAACT GGGACCAAGG TCACCGTCCT AGGTCAGCCC
1981 AAGGCCAACC CCACTGTCAC TCTGTTCCCG CCCTCCTCTG AGGAGCTCAA AGCCAACAAG
2041 GCCACACTAG TGTGTCTGAT CAGTGACTTC TACCCGGGAG CTGTGACAGT GGCCTGGAAG
2101 GCAGATAGCA GCCCCGTCAA GGCGGGAGTG GAGACCACCA CACCCTCCAA ACAAAGCAAC
2161 AACAAGTACG CGGCCAGCAG CTATCTGAGC CTGACGCCTG AGCAGTGGAA GTCCCACAGA
2221 AGCTACAGCT GCCAGGTCAC GCATGAAGGG AGCACCGTGG AGAAGACAGT GGCCCCTACA
2281 GAATGTTCAT AATAAACCGC CTCCACCGGG CGCGCCAATT CTATTTCAAG GAGACAGTCA
2341 TAATGAAATA CCTATTGCCT ACGGCAGCCG CTGGATTGTT ATTACTCGCG GCCCAGCCGG
2401 CCATGGCCGA AGTTCAATTG TTAGAGTCTG GTGGCGGTCT TGTTCAGCCT GGTGGTTCTT
2461 TACGTCTTTC TTGCGCTGCT TCCGGATTCA CTTTCTCTTC GTACGCTATG TCTTGGGTTC
2521 GCCAAGCTCC TGGTAAGGT TTGGAGTGGG TTTCTGCTAT CTCTGGTTCT GGTGGCAGTA
2581 CTTACTATGC TGACTCCGTT AAAGGTCGCT TCACTATCTC TAGAGACAAC TCTAAGAATA
2641 CTCTCTACTT GCAGATGAAC AGCTTAAGGG CTGAGGACAC TGCAGTCTAC TATTGCGCTA
2701 AAGACTATGA AGGTACTGGT TATGCTTTCG ACATATGGGG TCAAGGTACT ATGGTCACCG
2761 TCTCTAGTGC CTCCACCAAG GGCCCATCGG CTTCCCCCT GGCACCCTCC TCCAAGAGCA
2821 CCTCTGGGGG CACAGCGGCC CTGGGCTGCC TGGTCAAGGA CTACTTCCCC GAACCGGTGA
2881 CGGTGTCGTG GAACTCAGGC GCCCTGACCA GCGGCGTCCA CACCTTCCCG GCTGTCCTAC
2941 AGTCTAGCGG ACTCTACTCC CTCAGCAGCG TAGTGACCGT GCCCTCTTCT AGCTTGGGCA
3001 CCCAGACCTA CATCTGCAAC GTGAATCACA AGCCCAGCAA CACCAAGGTG GACAAGAAAG
3061 TTGAGCCCAA ATCTTGTGCG GCCGCTCATC ACCACCATCA TCACTCTGCT GAACAAAAAC
3121 TCATCTCAGA AGAGGATCTG AATGGTGCCG CAGATATCAA CGATGATCGT ATGGCTGGCG
3181 CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA TTTACTAACG
3241 TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT CTGTGGAATG
3301 CTACAGGCGT TGTAGTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA TGGGTTCCTA
3361 TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT TCTGAGGGTG
3421 GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT ATTCCGGGCT
3481 ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA AACCCGCTA
3541 ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTT CAGAATAATA
3601 GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT CAAGGACTG
3661 ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG TATGACGTT
3721 ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA GATCCATTCG
3781 TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT GCTGGCGGCG
3841 GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG agggtggtgg ctctgagggt ggcggttctg
3901 AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT GATTTTGATT
3961 ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT GAAAACGCGC
4021 TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT GCTGCTATCG
4081 ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT GGTGATTTTG
4141 CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT TTAATGAATA
4201 ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT TTTGTCTTTA
4261 GCGCTGGTAA ACCATATGAA TTTTCTATGG ATTGTGACAA AATAAACTTA TTCCGTGGTG
4321 TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG TTTGCTAACA

4381 TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT TATTATTGCG
4441 TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC TTAAAAAGGG
4501 CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTCTTGCT CTTATTATTG GGCTTAACTC
4561 AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTCACT TTGTTCAGGG
4621 TGTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC TCTCTGTAAA
4681 GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG ATTGGGATAA
4741 ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG CTCGTTAGCG
4801 TTGGTAAGAT TCAGGATAAA ATTGRAGCTG GGTGCAAAAT AGCAACTAAT CTTGATTTAA
4861 GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT CTTAGAATAC
4921 CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT TCCTACGATG
4981 AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT ACCCGTTCTT
5041 GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT AAATTAGGAT
5101 GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG CGTTCTGCAT
5161 TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT TTTGTCGGTA
5221 CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT GTTGGCGTTG
5281 TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT ACTGGTAAGA
5341 ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT TCCGGTGTTT
5401 ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA AATTTAGGTC
5461 AGAAGATGAA ATTAACTAAA ATATATTTGA AAAAGTTTTC TCGCGTTCTT TGTCTTGCGA
5521 TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG GAGGTTAAAA
5581 AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACTAT TGACTCTTCT CAGCGTCTTA
5641 ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT AGCGACGATT
5701 TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC ATTAAAAAAG
5761 GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT TGTTTCATCA
5821 TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT TGTAACTTGG
5881 TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAAAGG TACTGTTACT
5941 GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC TGTTTACGT
6001 GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA TAATCCAAAC
6061 AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA TGATAATTCC
6121 GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC TTTTAAAATT
6181 AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTAAA GTCTAATACT
6241 TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT TTCTGCACCT
6301 AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC AACTGACCAG
6361 ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA TTTTTCATTT
6421 GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG CCTCACCTCT
6481 GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTTAATG GCGATGTTTT AGGGCTATCA
6541 GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG TATTCTTACG
6601 CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT TACTGGTCGT
6661 GTGACTGGTG AATCTGCCAA AGTAAATAAT CCATTTCAGA CGATTGAGCG TCAAAATGTA
6721 GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT TCTGGATATT
6781 ACCACCAAGG CCGATAGTTT GAGTTCTTCT ATCCAGGCAA GTGATGTTAT TACTAATCAA

6841 AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT CGGTGGCCTC
6901 ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA ATCCCTTTA
6961 ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAG AAAGCACGTT ATACGTGCTC
7021 GTCAAAGCAA CATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT
7081 TACGCGCAGC GTGACCGCA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGTTCTT
7141 CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCGTCAA GCTCTAAATC GGGGGCTCCC
7201 TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTTGGGTGA
7261 TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC
7321 CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGAACA ACACTCAACC CTATCTCGGG
7381 CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA ACAGGATTTT
7441 CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG CCAGGCGGTG
7501 AAGGGCAATC AGCTGTTGCC CGTCTCACTG GTGAAAAGAA AAACCACCCT GGATCCAAGC
7561 TTGCCAGGTGG CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTTCTAA
7621 TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT CAATAATATT
7681 GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC TTTTTTGCGG
7741 CATTTTGCCT TCCTGTTTTT GCTCACCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG
7801 ATCAGTTGGG CGCACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG
7861 AGAGTTTTCG CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTAGTC
7921 ATACACTATT ATCCCGTATT GACGCCGGGC AAGAGCAACT CGGTCGCCGG GCGCGGTATT
7981 CTAGAATGA CTTGGTTGAG TACTCACCAG TCACAGAA GCATCTTACG GATGGCATGA
8041 GAGTAAGAGA ATTATGCAGT GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC
8101 TTCTGACAAC GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC
8161 ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC
8221 GTGACACCAC GATGCCTGTA GCAATGCCA CAACGTTGCG CAAACATTA ACTGGCGAAC
8281 TACTTACTCT AGCTTCCCGG CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG
8341 GACCACTTCT GCGCTCGGC CTTCCGGCTG GCTGGTTTAT TGCTGATAA TCTGAAGCCG
8401 GTGAGCGTGG GTCTCGCGGT ATCATTGCAG CACT GGGGCC AGATGTAAG CCCTCCCGTA
8461 TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG
8521 CTGAGATAGG TGCCTCACTG ATTAGCATT GGTAACTGTC AGACCAAGTT TACTCATATA
8581 TACTTTAGAC TGATTTAAAA CTTCATTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT
8641 TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTC GTTCCACTGT ACGTAAGACC
8701 CCCAAGCTTG TCGACTGAAT GGCGAATGGC GCTTTGCCTG GTTTCCGGA CCAGAAGCGG
8761 TGCCGGAAAG CTCGGCTGGAG TGCGATCTTC CTGAGGCCA TACTGTCGGTC TCCCCTCAA
8821 ACTGGCAGAT GCACGGTTAC GATGCGCCCA TCTACAACCAA CGTAACCTAT CCCATTACG
8881 TCAATCCGCC GTTTGTTCCC ACGGAGAATC CGACGGGTTG TTACTCGCTC ACATTTAATG
8941 TTGATGAAAG CTGGCTACAG GAAGGCCAGA CGCGAATTAT TTTTGATGCC GTTCCTATTG
9001 GTTAAAAAAT GAGCTGATTT AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGTT
9061 TACAATTTAA ATATTTGCTT ATACAATCTT CCTGTTTTTG GGGCTTTTCT GATTATCAAC
9121 CGGGGTACAT ATGATTGACA TGCTAGTTTT ACGATTACCG TTCATCGATT CTCTTGTTTG
9181 CTCAGACTC TCAGGCAATG ACCTGATAGC CTTTGTAGAT CTCTCAAAAA TAGCTACCCT
9241 CTCCGGCATG AATTTATCAG CTAGAACGGT TGAATATCAT ATTGATGGTG ATTTGACTGT

9301 CTCCGGCCTT TCTCACCCTT TTGAATCTTT ACCTACACAT TACTCAGGCA TTGCATTTAA
9361 AATATATGAG GGTTCTAAAA ATTTTTATCC TTGCGTTGAA ATAAAGGCTT CTCCCGCAA
9421 AGTATTACAG GGTCATAATG TTTTGGTAC AACCGATTTA GCTTTATGCT CTGAGGCTTT
9481 ATTGCTTAAT TTTGCTAATT CTTTGCCTTG CCTGTATGAT TTATTGGATG TT

Tabela 22: Iniciadores utilizados em amplificação por RACE:

Cadeia pesada	
HuCm-FOR (1ª PCR)	5′-TGG AAG AGG CAC GTT CTT TTC TTT-3′
HuCm-Interno(2ª PCR)	5′ CTT TTC TTT GTT GCC GTT GGG GTG-3′
Cadeia leve kappa	
HuCkFor (1ª PCR)	5′-ACA CTC TCC CCT GTT GAA GCT CTT-3′
HuCkForAscI(2ª PCR)	**5′ -ACC GCC TCC ACC GGG CGC GCC TTA TTA ACA CTC TCC CCT GTT GAA GCT CTT -3′**
Cadeia leve lambda	
HuClambdaFor (1ª PCR)	
HuCL2-FOR	5′-TGA ACA TTC TGT AGG GGC CAC TG-3′
HuCL7-FOR	5′-AGA GCA TTC TGC AGG GGC CAC TG-3′
HuClambdaForAscI (2ª PCR)	
HuCL2-FOR-ASC	**5′ -ACC GCC TCC ACC GGG CGC GCC TTA TTA TGA ACA TTC TGT AGG GGC CAC TG -3**
HuCL7-FOR-ASC	**5′ -ACC GCC TCC ACC GGG CGC GCC TTA TTA AGA GCA TTC TGC AGG GGC CAC TG -3**
GeneRAcer 5′ Iniciadores proporcionados com o kit (Invitrogen)	
5′A 1ª PCR	5′CGACTGGAGCACGAGGACACTGA 3′
5′NA 2ª pCR	5′GGACACTGACATGGACTGAAGGAGTA-3′

Tabela 23: ON utilizados na Captura de cadeias leves kappa utilizando O método CJ e *BsmAI*

Todos os ON são escritos de 5' a 3'.

REdaptors (6)		
ON_20SK15012		gggAggATggAgAcTgggTc
ON_20SK15L12		gggAAgATggAgAcTgggTc
ON_20SK15A17		gggAgAgTggAgAcTgAgTc
ON_20SK15A27		gggTgccTggAgAcTgcgTc
ON_20SK15A11		gggTggcTggAgAcTgcgTc
ON_20SK15B3	gggAgTcTggAgAcTgggTc	
Pontes (6)		
kapbril012	gggAggATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
kapbril L12	gggAAgATggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
kapbrilA17	gggAgAgTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
kapbril A27	gggTgccTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
kapbrilA11	gggTggcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
kapbrilB3	gggAgTcTggAgAcTgggTcATcTggATgTcTTgTgcAcTgTgAcAgAgg	
Entensor (5' bioetinilado)		
kapextl bio	ccTcTgTcAcAgTgcAcAAgAcATccAgATgAcccAgTcTcc	
Iniciadores		
kaPCRtl	ccTcTgTcAcAgTgeAcAAgAc	
kapfor	5'-aca ctc tcc cct gtt gaa gct ctt-3'	

Tabela 24: programa PCR para amplificação de ADN kappa

95 °C	5 minutos
95 °C	15 segundos
65 °C	30 segundos
72 °C	1 minuto
72 °C	7 minutos
4 °C	manter

Reagentes (reacção 100 uL):	
Molde	50 ng
tampão de PCR 10x turbo	1x
turbo Pfu	4U
dNTPs	cada 200 µM
kaPCRtl	300 nM
kapfor	300 nM

Tabela 25: h3401-h2 capturado Via CJ com BsmAI

```
!   1 2  3 4 5 6 7 8 9 10 11 12 13 14 15
!   S A   Q D I Q M   T Q   S   P   A   T L S
    aGT GCA Caa gac atc cag atg acc cag tct cca gcc acc ctg tct
! ApaLI...                          a gcc acc ! L25, L6. L20, L2, L16. A11
! Extensor.........................Ponte...

! 16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
! V   S   P   G   E   R   A   T   L   S   C  R  A  S  Q
 gtg tct cca ggg gaa agg gcc acc ctc tcc tgc agg gcc agt cag

! 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
! S  V  S  N N  L  A  W  Y  Q  Q  K  P  G  Q
agt gtt agt aac aac tta gcc tgg tac cag cag aaa cct ggc cag

! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
! V  P  R  L  L I Y  G  A  S  T  R  A  T  D
gtt ccc agg ctc atc tat ggt gca cc acc agg gcc act gat

! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
! I P  A  R  F  S  G  S  G  S  G  T  D  F  T
atc cca gcc agg ttc agt ggc agt ggg tct gga aca gac ttc act

! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
! L T  I  S  R  L  E P  E  D  F  A  V  Y  Y
ctc acc atc agc aga ctg gag cct gaa gat ttt gca gtg tat tac

! 91 92 93 94 95 96 97 98 99 100 101 012 103 104 105
! C  Q  R  Y  G  S  S  P  G  W  T  F  G  Q  G
tgt cag gcc tat ggt agc tca ccg ggg tgg acg ttc ggc caa ggg

!106 107 108 109 110111 112 113 114 115 116 117 118 119 120
! T  K  V  E  I  K  R  T  V  A  A  P  S  V  F
acc aag gtg gaa act aaa act gtg gct gca cca tct gtg ttc

! 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
!  I  F  P  P  S  E  D  Q  L  K  S  G  T  A   S
atc tc ccg cca tct gat gag cag ttg aaa tct gga tct gcc tct
```

```
!136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
! V  V  C  L  L  N  N  F  Y   P  R  E  A  K   V
gtt gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta

! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!  Q  W  K  V  D   N  A   L  Q  S  G  N  S  Q  E
cag tgg aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag

! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!  S  V  T E  Q   D   S   D   S   T   Y  S  L  S
agt gtc aca gac cag gac agc aag gac agc acc tac agc ctc agc

! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
! S  T  L  T  L  S  K  A  D   Y   E   K  H  K  V
agc acc ctg acg ctg agc aa gca gac tac gag aaa cac aaa gtc

!196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!  Y  A  C  E  V  T  H  Q  G  L  S  S  P  V  T
tac gcc tgc gaa gtc acc cat cag ggc ctg agc tcg cct gtc aca

!211 212 213 214 215 216 217 218 219 220 221 222 223
!  K  S  F  N  K  G  E  C  K   G  E  F  A
aag agc ttc aac aaa gga gag tgt aag ggc gaa ttc gc.....
```

Tabela 26: h3401-d8 KAPPA capturado com CJ e BsmAI

```
!   1 2  3 4 5 6 7 8 9 10 11 12 13 14 15
!    S A   Q D I Q M   T   Q   S P   A T   L S
    agt GCA Caa gac atc cag atg acc cag tct cca gcc acc ctg tct
! ApaLI... Extensor                      a gcc acc ! L25, L6. L20, L2, L16. A11
!                            A GC ACC CTG TCT ! L2

! 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
!   S  P  G  E  R  A  T  L  S  C  T  A  S  Q
    gtg tct cca ggg gaa agg gcc acc ctc tcc tgc agg gcc agt cag
!GTG TCT CCA GGG GAA AGA GCC ACC CTC ACC CTC TCC TGC ! L2

! 31 32 33 34 35 36 27 38 39 40 41 42 43 44 45
! N  L  L  S  N  L  A  W  Y  Q  Q  K  P  G Q
aat ctt ctc agc aat tta gcc tgg tac cag aaa cct ggc cag

! 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
! A  P R  L  L  I  Y G A  S  T  G  A  I  G
gct ccc agg ctc CTC atc tat ggt gtc gca tcc acc ggg gcc att ggt

! 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
! IPARFSGSGSGTEFT
atc cca gcc agg ttc agt ggc agt ggg tct gga aca gac ttc act

! 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
!L  T  I  S  R  L  E  P E D  F  A  V  Y  F
ctc acc atc agc aga ctg gag cct gaa gat ttt gca gtg tat tac

! 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
!C  Q  Q  Y  G  T  S  P  P  T  F  G  G  G  T
tgt cag cag tat ggt acc tca ccc act ttc ggc gga ggg acc

!106 107 108 109 110111 112 113 114 115 116 117 118 119 120
! K  V  E  I  K  R  T  V  A  A  P  S  V  F  I
aag gtg gag act aaa cga act gtg gct gca cca tct gtg ttc atc

! 121 122 123 124 124 126 127 128 129 130 131 132 133 134 135
!   F  P P  S  E  D  Q  L  K  S  G   T  A  S   V
ttc ccg cca tct gat gag cag ttg aaa tct gga tct gcc tct gtt
```

```
!136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
!  V  C  L  L  N  N  F  Y  P  R  E  A  K  V  Q
gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag

! 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!  W  K  V  D  N  A   L   Q  S  G  N  S  Q  E  S
tgg aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt

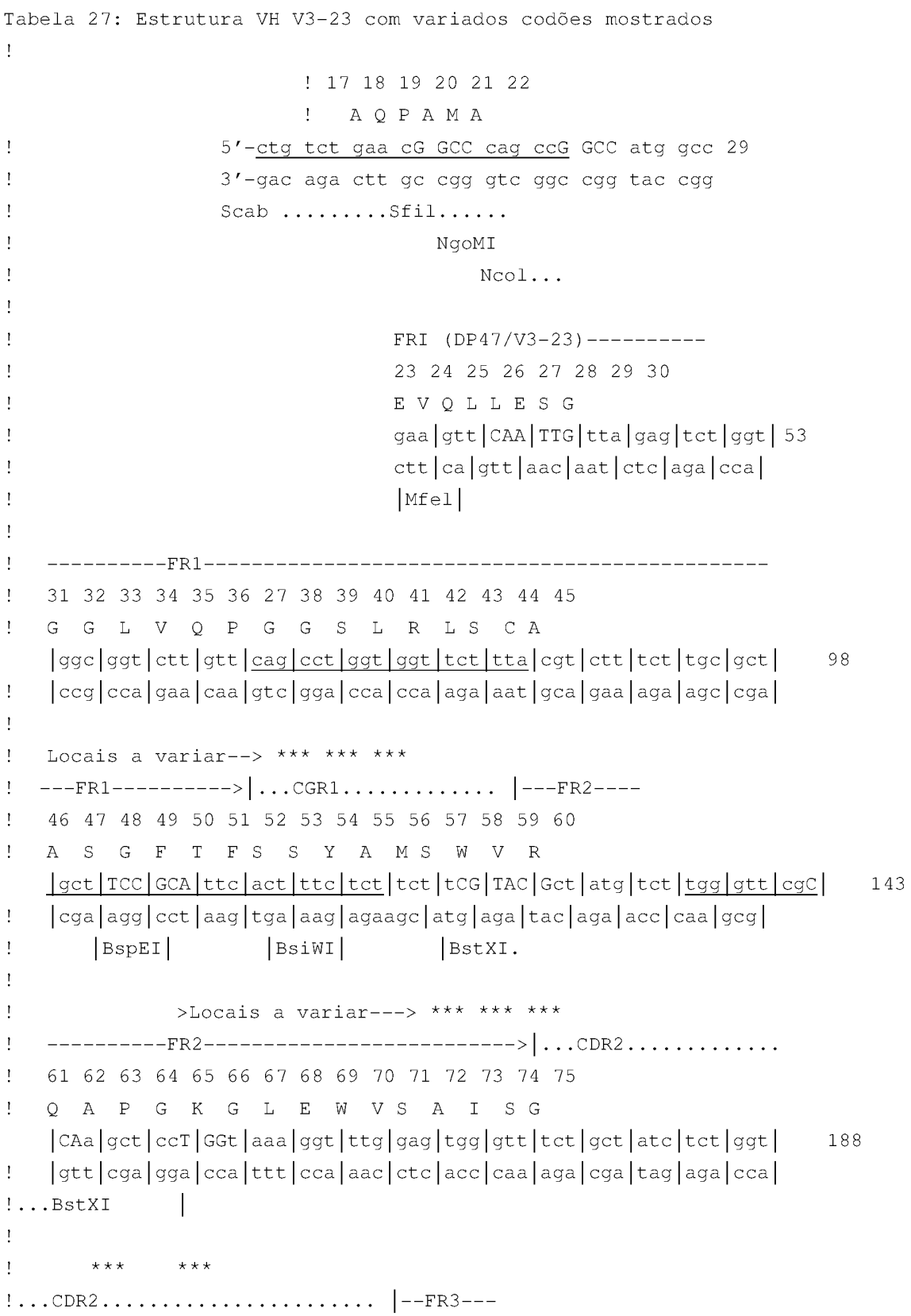
! 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!  V  T  E  Q  D  N  K  D  S  T  Y  S  L  S   S
gtc aca gac cag gac agc aag gac agc acc tac agc ctc agc agc

! 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
!  T  L  T  L  S  K  V  D  Y  E  K  H  E  V   Y
acc ctg acg ctg agc aaa gca gac tac gag aaa cac aaa gtc tac

!196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!  A  C  E  V  T  H  Q  G  L  S  S  P  V  T  K
gcc tgc gaa gtc acc cat cag ggc ctg agc tcg cct gtc aca aag

!211 212 213 214 215 216 217 218 219 220 221 222 223
! S  F  N  K  G  E  C  K  G  E  F  A  V
agc ttc aac aaa gga gag tgt aag ggc gaa ttc gtt t
```

Tabela 27: Estrutura VH V3-23 com variados codões mostrados



```
!   76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
!   S  G  G  S  T  Y  Y  A  D  S  V  K  G  R F
    |tct|ggt|ggc|agt|act|tac|tat|gct|gac|tcc|gtt|aaa|ggt|gcg|    233
!   |aga|cca|ccg|tca|tga|atg|ata|cga|cga|ctg|agg|caa|ttt|cca|gcg|aag|
!
!   ----------FR3-----------------------------------------------
!   91 92 93 94 95 96 97 98 99 100 101 012 103 104 105
!   T  I  S  R D  N  S  K  N  T  L  Y  L  Q  M
    |act|atc|TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|    278
!   |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|
!         |XbaI|
!
!   ---FR3-----------------------------------------------------|
!   106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!     N  S  L  R  A  E  D  T  A  V  Y  Y  C  A  K
    |aac|agC|TTA|AGg|gct|gag|gac|aCT|GCA|gtc|tac|tat|tgc|gct|aaa|    323
!   |ttg|tcg|aat|tcc|cga|ctc|ctg|tga|cgt|cag|atg|ata|acg|cga|ttt|
!      |AflII|    |PstI|
!
!   .....CDR3.....  |----FR4-----------------
! 121 122 123 124 124 126 127 128 129 130 131 132 133 134 135
!   D Y E G T G Y A F G I W G Q G
    |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ata|tgg|ggt|caa|ggt|    368
!   |ctg|ata|ctt|cca|tga|cca|ata|cga|aag|ctg|tat|acc|cca|gtt|cca|
!                       |NdeI|
!
!   ------------Fr4---------->|
!   136 137 138 139 140 141 142
!    T  M  V  T  V   S  S
    |act|atG|GTC|ACC|gtc|tct|agt-       389
!   |tga|tac|cag|tgg|cag|aga|tca-
!      |BstEII|
!
!                143 144 145 146 147 148 149 150 151 152
!                 A  S  T  K  G  P  S  V  F  P
                 gcc tcc acc aaG GGC CCa tcg GTC TTC ccc-3' 419
!                cgg agg tgg ttc ccg ggt agc cag aag ggg-5'
!                       Bsp1201.   BbsI...(2/2)
!                       ApaI....
```

```
(SFPRMET) 5'-ctg tct gaa cG GCC cag ccG-3'
(TOPFR1A) 5'-ctg tct gaa cG GCC cag ccG GCC atg gcc-
   gaa|att|CAA|TTG|tta|gag|tct|ggt|-
   |ggc|ggt|ctt|gtt|cag|cct|ggt|ggt|tct|tta-3'
(BOTFR1B)         3'-caa|gtc|gga|cca|cca|aga|aat|gca|faa|aga|agc|cga|-
       |cga|agg|cct|aag|tga|aag-5'! cadeia de fundo
(BOTFR2)   3'-acc|caa|gcg|-
       |gtt|cga|gga|cca|ttt|cca|aac|ctc|acc|caaaga|-5' ! cadeia de fundo
(BOTFR3) 3'-a|cgactg|agg|caa|ttt|cca|gcg|aag|-
       |tga|tag|aga|tct|ctg|ttg|aga|ttc|tta|tga|gag|atg|aac|gtc|tac|-
   |ttg|tcg|aat|tcc|cga|ctc|ctg|tga-5'
(F06)      5'-gC|TTA|AGg|gct|gag|gac|aCT|GCA|Gtc|tac|tat|tgc|gct|aaa|-
   |gac|tat|gaa|ggt|act|ggt|tat|gct|ttc|gaC|ATA|TGg|ggt|c-3'
(BOTFR4)   3'-cga|aag|ctg|tat|acc|caa|gtt|cca|-
       |tga|tac|cag|tgg|cag|aga|tca|-
          cgg agg tgg ttc ccg ggt agc cag aag ggg-5'! cadeia de fundo
(BOTPRCPRIM)      3'-gg ttc ccg ggt agc cag aag ggg-5'
!
!diversidade CDR1
!
(ON-vgC1)  5'-|gct|TCC|GGA|ttc|act|ttc|tct|<1>|TAC|<1>|atg|<1>|
!                            CDR1..............6859
       |tgg|gtt|cgC|CAa|gct|ccT|GG-3'
!
!<1> permanence para uma mistura equimolar de {ADEFGHIKLMNPQRSTVWY};sem C
!                        (isto não é uma Sequência)
!
! diversidade CDR2
!
(ON-vgC2)  5'-ggt|ttg|gag|tgg|gtt|tct|<2>|atc|<2>|<3>|-
                            CDR2.............
              |tct|ggt|ggc|<1>|act|<1>|tat|gct|gac|tcc|aaa|gg-3'
!             CDR2............................
!<1> é uma mistura equimolar de {ADEFGHIKLMNPQRSTVWY};sem C
!<2> é uma mistura equimolar de {YRWVGS}; sem ACDEFHIKLMNPQT

!<3> é uma mistura equimolar de {ps}; sem ADEFGHIKLMNPQRSTVWY
```

Tabela 28: Enchimento utilizado em VH

1 TCCGGAGCTT CAGATCTGTT TGCCTTTTTG TGGGGTGGTG CAGA TCGCGT TACGGAGATC
61 GACCGACTGC TTGAGCAAAA GCCACGCTTA ACTGCTGATC AGGCATGGGA TGTTATTCGC
121 CAAACCAGTC GTCAGGATCT TAACCTGAGG CTTTTTTTAC CTACTCTGCA AGCAGCGACA
181 TCTGGTTTGA CACAGAGCGA TCGCGTCGT CAGTTGGTAG AAACATTAAC ACGTTGGGAT
241 GGGATCAATT TGCTTAATGA TGATGGTAAA ACCTGGCAGC AGCCAGGCTC TGCCATCCTG
301 AACGTTTGGC TGACCAGTAT GTTGAAGCGT ACCGTAGTGG CTGCCGTACC TATGCCATTT
361 GATAAGTGGT ACAGCGCCCAG TGGCTACGAA ACAACCCAGG ACGGCCCAAC TGGTTCGCTG
421 AATATAAGTG TTGGAGCAAA AATTTTGTAT GAGGCGGTGC AGGGAGACAA ATCACCAATC
481 CCACAGGCGG TTGATCTGTT TGCTGGGAAA CCACAGCAGG AGGTTGTGTT GGCTGCGCTG
541 GAAGATACCT GGGAGACTCT TTCCAAACGC TATGGCAATA ATGTGAGTAA CTGGAAAACA
601 CCTGCAATGG CCTTAACGTT CCGGGCAAAT AATTTCTTTG GTGTACCGCA GGCCGCAGCG
661 GAAGAAACGC GTCATCAGGC GGAGTATCAA AACCGTGGAA CAGAAAACGA TATGATTGTT
721 TTCTCACCAA CGACAAGCGA TCGTCCTGTG CTTGCCTGGG ATGTGGTCGC ACCCGGTCAG
781 AGTGGGTTTA TTGCTCCCGA TGGAACAGTT GATAAGCACT ATGAAGATCA GCTGAAAATG
841 TACGAAAATT TTGGCCGTAA GTCGCTCTGG TTAACGAAGC AGGATGTGAA GGCGCATAAG
901 GAGTCGTCTA GA

Tabela 29: Sequência de ADN de pCES5

```
! pCES5 6680 bases - pCes4 com Enchimentos em CDR1-2 e CDR3 2000.12.13
!
!Ngene = 6680
! RE Útil (corte MAnoLi menos do que 3 vezes) 2000.06.05
!
!Não-cortantes
!Acc651 Ggtacc    Afel AGCgct    AvrII Cctagg
!BsaBi GATNNnnatc bsiWICgtacg    BsmFI Nnnnnnnnnnnnnnngtcc
!BsrFI Tgtaca  BstAPI GcANNNNntgc BstBI TTcgaa
!BstZ171GTAtac BtrI CACgtg       EcII36I GAGctc
!EcoRV GATatc  Fsel GGCCGGcc  K[nl GGTACc
!Mscl TGGcca   Nrul TCGcga       Nsil ATGCAt
!Pacl TTAATtaa    Pmel GTTTaaac  Pmll CACgtg
!PpuMl TGgwccy    PshAIGACNNnngtc   SacI GAGCCTc
!SacII CCGcgg     Sbfl CCTGCAgg  SexAI Acceggt
!Sgdl GcGATcgc    SnaBl TACgta   Spel Actagt
!Sphl GCATGc      Sse83871 CCTGCAgg Stul AGGcct
!Swalk ATTTaaat   Xmal Cccggg
!
!cortantes
!Enzimas que cortam mais do que  3 vezes.
!AlwNl CAGNNNctg             5
!Bsgl ctgcac             4
!BsrFl Rccggy             5
!Earl CTCTTCNnnn          4
!Faul nNNNNNNGCGGG        10
!
!Enzimas que cortam de 1 a 3 vezes.
!
!EcoO1091 RGgnccy         3   7   2636    4208
!BssSI Ctcgtg         1   12
!-"= Cacgag           1   1703
!BspHI Tcatga             3   43  148 1156
!AatII GACGTc             1   65
!BcuVU GTATCCNNNNNN       2   140 1667
!Eco571CTGAAG             1   301
!-"- cttcag           2   1349
!Aval Cycgrg              3   319 2347    6137
!BsiHKI GWGCCWc           3   401 2321    4245
```

```
!HgiAI GWGCWc         3   401 2321   4245
!Bcgl gcannnnnntcg        1   461
!Scal AGTact          1   505
!Pvul CGATcg          3   616 3595   5926
!Fspl TGCgca          2   763 5946
!BgII GCCNNNNnggc     3   864 2771   5952
!BpmI CTGGAG      1   898
!-"- ctccag               1   4413
!Bsal GGTCTCNnnnn             1   916
!Ahdl GACNNNnngtc         1   983
!EamII051 GACNNNnngtc     1983
!Drdl GACNNNNnngtc            3   1768   6197   6579
!Sapl gaagagc         1   1998
!Pvull CAGctg         3   2054   3689   5896
!PflMI  CCANNNNntgg           3   2233   3943   3991
!HindIII Aagctt           1   2235
!ApaLi Gtgcac             1   2321
!BspMI Nnnnnnnnngcaggt            1   2328
!-"-       ACCTGCNNNNn            2   3460
!Pstl CTGCAg              1   2335
!Accl GTmkac          2   2341   2611
!Hincll GTYrac            2   2341   3730
!SaII Gtcgac          1   2341
!Tlil Ctcgac  1   2347
!Bbsl gtcttc          2   2383   4219
!Blpl GCtnagc         1   2580
!Espl GCtnagc         1   2580
!SgrAl CRccggyg           1   2648
!Agel Accggt              2   2649   4302
!Ascl GGcgcgcc    1   2689
!BssHII Gcgcgc        1   2690
!Sfil GGCCNNNNnggcc       1   2270
!Naek GCCggc          2   2776   6349
!NgoMiV Gccggc        2   2776   6349
!Btgl Ccrygg          3   2781   3553   5712
!Dsal Ccrygg          3   2781   3553   5712
!Ncol Ccatgg   `       1   2781
!Styl Ccwwgg          3   2781   4205   4472
!Mfel Caattg          1   2795
!BspEL Tccgga         1   2861
!BgIII Tgatca         1   2956
```

```
!Bsu361 CCtnagg          3   3004   4143   4373
!Xcml CCANNNNNnnnntgg    1   3215
!Mlul Acgccgt         1   3527
!Hpal GTTaac          1   3730
!XbaI Tctaga      1   3767
!
!AflII Cttaag         1   3811
!BsmL NGcattc         1   3821
!-"-GAATGCN               1   4695
!RsrII CGgwccg        1   3827
!Nhel Gctagc      1   4166
!BstEII Ggtnacc           1   4182
!BsmB1 CGTCTCNnnnn            2   4188   6625
!-"-   Nnnnnngagacg          1   6673
!Apal GGGCCc          1   4209
!BanII GRGCYc         3   4209   4492   6319
!Bsp1201 Gggccc       1   4209
!PspOMl Gggccc    1   4209
!BseRl NNnnnnnnnnctcctc       1   4226
!-"-   GAGGAGNNNNNNNNNN  1   4957
!EcoNI CCTNNnnnagg            1   4278
!PflFI GACNnngtc          1   4308
!TthIIIIGACNnngtc     1   4308
!Kasl Ggcgcc          2   4327   5967
!BstXL CCANNNNNntgg           1   4415
!NotI GCggccgc        1   4507
!EagI Cggccg          1   4508
!BamHL Ggatcc         1   5169
!BspDL ATcgat         1   5476
!NdeI CAtatg          1   5672
!EcoRI Gaattc     1   5806
!PsiI TTAtaa   1   6118
!DraIII CACNNNgtg         1   6243
!BsaAI YACgtr         1   6246
!-----------------------------------------------------------------------
    1   gacgaaaggg cCTCGTGata cgcctatttt tataggttaa tgtcatgata ataatggttt
!              BssSI.(1/2)
    61  cttaGACGTC aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt
!          AatII
    121 tctaaataca ttcaaatatG TATCCgctca kgagacaata accctgataa atgcttcaat
!              BcVI..(1 of 2)
```

```
   181 aataatgaaa aaggaagagt
!Base # 201 to 1061 = gene ApR de pUC119 com doma de locais RE removidos
!
!      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
!    fM  S I Q H F R  V A  L  I  P F F  A
201 atg agc att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg
!
!      16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
!      A  F  C  L P  V  F  A  H  P E  T  L  V  K
246 gca ttt lgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa
!
!      31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
!      C  K  D  A  E  D  Q  L  G  A  R  V G  Y I
291 gta aaa gat gct gaa gat cag ttg ggt gcc cga gtg ggt tac atc
!
!      46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
!      E L D  L  N  S  G  K  I  L  E  S  F R P
366 gaa ctg gat ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc
!
!      61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
!      E E  R  F  P  M  M  S  T  F  K  V  L  L C
   381 gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt
!
!      76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
!      G A V L  S  R  I  D  A  G  Q  E  Q  L  G
426 ggc gcg gta tta tcc cgt att gac gcc ggg caa gaG CAa ctc ggT
!                          Bcgl........
!
!      91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
!      R R I H  Y  S  Q  N  D  L  V  E  Y  S  P
471 CGc cgc ata cac tat tct cag aat gac ttg gtt gAG TAC Tca cca
!..BcgL......                          Scal....
!
!      106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!      V  T  E  K  H   L  T  D   G  M  T  V  R  E  L
   516 gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta
!
!      121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
!      C  S  A  A  I  T  M  S  F  N  T  A  A  N  L
   561 tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta
!
```

```
!      136 137 138 139 140 141 142 143 144 145 146 147 148 148 150
!      L  L  T  T  I  G  G  P  K  E  L  T  A  F  L
   606 ctt ctg aga aCG ATC Gaa gga ccg aag gag cta acc gct ttt ttg
!            Pvul...(1/2)
!
!      151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!      H  N  M  G  D  H  V  T  R  L  D  R  W  E  P
   651 cac aac atg ggg gat cat gta act cgc att gat cgt tgg gaa ccg
!
!      166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!      E  L  N  E  A  I  P  N  D  E  R  D  T  T  M
   696 gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg
!
!      181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
!      P  V  A  M  A  T  T  L  R  K  L  L  T  G  E
   741 cct gta GCA ATG gca aca acg tTG CGC Aaa cta tta act ggc gaa
!         BsrDl..(1/2)             Dspl....(1/2)
!      196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!      L  L  T  L  A  S  R  Q  Q  L  I  D  W  M  E
   786 cta ctt act cta gct tcc cgg caa caa tta ata gac tgg atg gag
!
!      211 212 213  214215 216 217 218 219 220 221 222 223 224 225
!      A  D  K  V  A  G  P  L  L  R  S  A  L  P  A
   831 gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct
!
!      226 227 228 229 230  231232  233  234  235  236  237   238 239 240
!      G  W  F  I  A  D  K  S  G  A  G  E  R  G  S
   876 ggc tgg ttt att gct gat aaa tCT GGA Gcc ggt gag cgt gGG TCT
!         Bpml.....(1/2)        Bsal.....
!
!      241 242 243 244 246 246 247 248 249 250 251 252 253 254 255
!      R  G  I  I  A  A  L  G  P  D  K  P  S  R
   921 Cgc ggt atC ATT GCa gca ctg ggg cca gat ggt aag ccc tcc cgt
!Bsal.....        BsrDL...(2/2)
!
!      256 257 258 259 260 261 262 263 264 264 266 267 268 269 270
!      I  V  V  I  Y  T  T  G  S  Q  A  T  M  D  E
   966 atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa
!            Ahdl.........
!
```

```
!      271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
!      R  N  R  Q  I  A  E  I  G  A  S  L  I  K  H
   1011   cga aat aga cag atc gct gag ata ggt gcc tca ctg att aag cat
!
!      286 287
!      W
   1056   tgg taa
   1062                          ctgtcagaccaagtttact
   1081   catatatact ttagattgat ttaaaacttc atttttaatttaaaaggatc taggtgaaga
   1141   tccttttttga taatctcatg accaaaatcc attaacgtga gttttcgttc cactgagcgt
   1201   cagaccccgt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct
   1261   gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc
   1321   taccaactct ttttccgaag gtaactggct tcaggacagc gcagatacca aatactgtcc
   1381   ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc
   1441   tcgctcgct aatcctgtta ccagtggctg ctggcagtgg cgatagtcg tgtcttaccg
   1501   ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggtt
   1561   cgtgcataca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg
   1621   agacattgaga aagcgccacg cttcccgaag ggagaaggc ggacagGTAT CCggtaagcg
!                           BciVI..(2 of 2)
   1681   gcagggtcgg aacaggagag cgCACGAGgg agcttccagg gggaaacgcc tggtatcttt
!             BssSI.(2/2)
   1741   atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag
   1801   gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt
   1861   gctggccttt tgctcACATG Ttctttcctg cgttatcccc tgattctgtg gataaccgta
!                  PciL...
   1921   ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgac cgcagcgagt
   1981   cagtgagcga ggaagcgGaa GAGCgcccaa tacgcaaacc gcctctcccc gcgcgttggc
!             SapI...
   2041   cgattcatta atgCAGCTGg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca
!  2101   acgcaatTAA TGTgagttag ctcactcatt aggcacccca ggcTTTACAc tttatgcttc
!             ..-35..    Plac                          ..-10.
   2161   cggctcgtat ggttgtgga attgtgagcg gataacaatt tcacaCAGGA AACAGCTATG
!                               M13Rev_seq_Iniciador
   2221   ACcatgatta cgCCAAGCTT TGGagccttt ttttggaga ttttcaac
!             PflMI.......
!                 Hind3.
! sinal::Elemento de ligação::CLeve
!
!      1   2   3   4   5   6   7   8   9   10  11  12  13  14  15
!      fM      K   K   L   L   F   A   I   P   L   V   V   P   F   Y
   2269   gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat
```

```
!
!                Elemento de ligação.....................Fim de FR4
!      16  17  18  19     20  21  22  23  24  25  26  27  28  29  30
!      SHSA   QVQLVDLEIK
   2315   TCT CAC Agt gca Cag gtc caa CTG CAG GTC GAC CTC GAG atc aaa
!            ApLI.....     Pstl...    Xhol...
!                          BspMI...
!                             SaLL...
!                             Accl...(1/2)
!                          Hincll.(1/2)
!
! Domínios Vleve podiam ser clonados como fragmentos ApaLI-Xhol.
!Segmentos VL-CL(kappa) podem ser clonados como fragmentos ApaLI-Ascl. <----
!
!      Ckappa-------------------------------------------
!      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
!       R  G  T  V  A  A  P  S  V  F  I  F  P   P  S
   2359   cgt gga act gtg gct gca caa tct GTC TTC atc ttc ccg cca tct
!               Bbsl...(1/2)
!
!      46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
!       D  E  Q  L  K  S   G  T  A  S   V   V  C  L  L
   2404   gat gag cag ttg aaa tct gga act gcc tct gtt gtg tgc ctg ctg
!
!      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
!       N  N  F  Y  P  R  E   A   K  V  Q  W  K  V   D
   2449   aat aac ttc tat ccc aga gag gcc aaa gta cag tgg aag gtg gat
!
!      76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
!       N  A  L  Q  S  G   N  S  Q  E  S   V  T  E   Q
   2494   aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc aca gag cag
!
!      91  92  93  94  95  96  97  98  99  100 101 102 103 104 105
        D  S  K  D  S  T  Y  S  L  S  S  T  L  T  L
   2539   gac agc aag gac agc acc tac agc ctc agc agc acc ctg acG CTG
!                                    EspI...
!
!      106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!       S  K  A  D  Y  E  K  H  K  V  Y  A  C  E  V
   2584   AGC aaa gca gac tac gag aaa cac aaa GTC TAC gcc tgc gaa gtc
!...Espl.....                       Accl....(2/2)
```

```
!
!      121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
!       T   H   Q   G   L   S   S   P   V   T   K   S   F   N   R
   2629   acc cat cag ggc ctg agt tcA CCG GTg aca aag agc ttc aac agg
!                Agcl....(1/2)
!
!      136 137 138 139 140
!      G   E   C   .   .
   2674   gga gag tgt taa taa GG CGCGCCaatt
!                   Ascl......
!                   BssHII.
!
   2701   ctatttcaag gagacagtca ta
!
!PelB::3-23(cheias)::CH1:: gene de fusão III
!
!      1   2   3   4   5   6   7   8   9   10  11  12  13  14  15
!       M   K   Y   L   L   P   T   A   A   A   G   L   L   L   L
   2723   atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc
!
!---------------------------
!
!      16  17  18  19  20  21  22
!       A   A   Q   P   A   M   A
   2768   gcG GCC cag ccG GCC atg gcc
!      Sfil...............
!             NgoMiV..(1/2)
!                 Ncol....
!
!                   FR1(DP47/V3-23------------
!                   23  24  25  26  27  28  29  30
!                    E   V   Q   L   L   E   S   G
   2789                gaa|gtt|Caa|TTG|tta|gag|tct|ggt|
!                          |Mfel|
!
!   ------------FR1---------------------------------------------
!      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
!       G   G   L   V   Q   P   G   G   S   L   R   L   S   C   A
   2813   |ggc|ggt|ctt|gtt|cag|cct|ggt|ggt|tct|tta|cgt|ctt|tct|tgc|gct|
!
!      -----FR1-----
```

```
!       46  47  48
!       A   S   G
        2858    |gct|TCC|GGA|
!           BspEI|
!
!       Enchimento para CDR1,FR2, e CDR2--------------------------->
!       Existem codões de paragem neste enchimento.
        2867                                        gcttcAGATC Tgtttgcctt
!                                                   BgIII..
        2887    tttgtggggt ggtgcagatc gcgttacgga gatcgaccga ctgcttgagc aaaagccacg
        2947 cttaactgcT GATCAggcat gggatgttat tcgccaaacc agtcgtcagg atcttaacct
!           BcII...
        3007    gaggcttttt ttacctactc tgcaagcagc gacatctggt ttgacacaga gcgatccgcg
        3067    tcgtcagttg gtagaaacat taacacgtt ggatggcatc aatttgctta atgatgatgg
        3127    taaaacctgg cagcagccag gctctgccat cctgaacgtt tggctgacca gtatgttgaa
        3187    gcgtaccgta gtggctgccg tacctatgCC Atttgataag TGGtacagcg ccagtggcta
!                                 Xcml........
        3247    cgaaacaacc caggacggcc caactggttc gctgaatata agtgttggag caaaaatttt
        3307    gtatgaggcg gtgcagggag acaaatcacc aatcccacag gcggttgatc tgtttgctgg
        3367    gaaaccacag caggaggttg tgttggctgc gctggaagat acctgggaga ctctttccaa
        3427    acgctatggc aataatgtga gtaactggaa aacagcctgca atggccttaa
cgttccgggc
        3487    aaataatttc tttggtgtac cgcaggccgc agcggaagaa ACGCGTcatc
aggcaggagta
!                       Mlul..
        3547    tcaaaccgt ggaacagaaa acgatatgat tgttttctca ccaacgacaa gcgatcgtcc
        3607    tgtgcttgcc tgggatgtgg tcgcacccgg tcagagtggg tttattgctc ccgatggaac
        3667    agttgataag cactatgaag atcagctgaa aatgtacgaa aattttggcc gtaagtcgct
!                       Pvull.
        3727    ctgGTTAACg aagcaggatg tggaggcgca taaggagtcg
!           Hpal..
!           HincLL(2/2)
!
!   ------FR3-----------------------------------------------
!       4   5   6   7   8   9   10  11  12  13  14  15  16
!       93  94  95  96  97  98  99  100 101 102 103 104 105
!       S   R   D   N   S   K   N   T   L   Y   L   Q   M
        3767    |TCT|AGA|gac|aac|tct|aag|aat|act|ctc|tac|ttg|cag|atg|
!               |Xbal|
!
!   --FR3---------------------------------------------->|
```

```
!   17  18  19  20
!   106 107 108 109
!   N   S   L   s   l   s   I   r   s   g
    3806    |aac|agC|TTA|AG t ctg agc att CGG TCC G
!               |Aflll|    Rsrll...
!
!       q   h   s   p   t   .
    3834    gg caa cat tct cca aac tga ccagacga cacaaacggc
    3782    ttacgctaaa tcccgcgcat gggatggtaa agaggtggcg tctttgctgg cctggactca
    3932    tcagatgaag gccaaaatt ggcaggagtg gacacagcag gcagcgaaac aagcactgac
    3992    catcaactgg tactatgctg atgtaaacgg caatattggt tatgttcata ctggtgctta
    4052    tccagatcgt caatcaggcc atgatccgcg attacccgtt cctggtacgg gaaaatggga
    4112    ctggaaaggg ctattgcctt ttgaaatgaa ccctaaggtg tataacccccc ag
    4164        aa GCTAGC ctgcggcttc
!               Nhel..
!
    4182    G|GTC|ACC|                      gtc tca agc
!       |BstEII|
!

!       136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
!       A   S   T   K   G   P   S   V   F   P   L   A   P   S   S
    4198    gcc tcc acc aag ggc cca tcg gtc ttc ccc ctg gca ccc tcc tcc
!
!       151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!       K   S   T   S   G   G   T   A   A   L   G   C   L   V   K
    4243    aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc ctg gtc aag
!
!       166 167 163 169 170 171 172 173 174 175 176 177 178 179 180
!       D   Y   F   P   E   P   V   T   V   S   W   N   S   G   A
    4288    gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg aac tca ggc gcc
!
!       181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
!       L   T   S   G   V   H   T   F   P   A   V   L   Q   S   S
    4333    ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta cag tcc tca
!
!       196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!       G   L   Y   S   L   S   S   V   V   T   V   P   S   S   S
    4378    gga ctc tact cc ctc agc agc gta gtg acc gtg ccc tcc agc agc
!
!       211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
```

```
!      L  G  T  Q  T  Y  I  C  N  V  N  H  K  P  S
   4423   ttg ggc acc tac atc tgc aac gtg aat cac aag ccc agc
!
!      226 227 228 229 230 231 232 233 234 235 236 237 238
!      N  T  K  V  D  K  K  V  E  P  K  S  C
   4468   aac acc aag gtg gac aag AAA GTT GAG CCC AAA TCT TGT
!               ON-TQHCforw..............
!
!            Poly His Elemento de ligação
!         139 140141  142 143 144 145 146 147 148 149 150
!         A  A  A  H  H  H  H  H  H  G  A  A
   4507       GCG GCC GCa cat cat cat cac cat cac ggg gcc gca
!         NotI......
!         EagI....
!
!      151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!      E  Q  K  L  I  S  E  E  D  L  N  G  A  A   .
   4543   gaa caa aaa ctc atc tca gaa gag gat gtc aat ggg gcc gca tag
!
!      III Maduro--------------------------->...
!      166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!      T  V  E  S  C  L  A  K  P  H  T  E  N  S  F
   4588   act gtt gaa agt tgt tta gca aaa cct cat aca gaa aat tca ttt
!
!      181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
!      T  N  V  W  K  D  D  K  T  L  D  R  Y  A  N
   4633   act aac gtc tgg aaa gac gac aaa act tta gat cgt tac gct aac

!      196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!      Y  E  G  C  L  W  N  A  T  G  V  V  V  C  T
   4678 tat gag ggc tgt ctg tgG AAT GCt aca ggc gtt gtg gtt tgt act
!               BsmI....
!
!      211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
!      G  D  E  T  Q  C  Y  G  T  W  V  P  I  G  L
   4723   gct atc cct gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct
!
!      241 242 243 244 245 246 247 248 249 250 251 252 253 254 255
!      E  G  G  G  S  E  G  G  G  T  K  P  P  E  Y
   4813   gag ggt ggc ggt tct gag ggt ggc ggt act aaa cct cct gag tac
!
```

```
!       271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
!       G   T   Y   P   P   G   T   E   Q   N   P   A   N   P   N
    4903   ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct aat
!
!       286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
!       P   S   L   E   E   S   Q   P   L   N   T   F   M   F   Q
    4948   cct tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag
!                    BseRI..(2/2)
!
!       301 302 303 304 305 306 307 308 309 310 311 312 313 314 315
!       N   N   R   F   R   N   R   Q   G   A   L   T   C   Y   T
    4993   aat aat agg ttc cga aat agg cag ggt gca tta act gtt tat acg
!
!       316 317 318 319 320 321 322 323 324 325 326 327 328 329 330
!       G   T   V   T   Q   G   T   D   P   V   K   T   Y   Y   Q
    5038   ggc act gtt act caa ggc act gac ccc gtt aaa act tat tac cag
!
!       331 332 333 334 335 336 337 338 339 340 341 342 343 344 345
!       Y   T   P   V   S   S   K   A   M   Y   D   A   Y   W   N
    5083   tac act cct gta tca tca aaa gcc atg tat gac gct tac tgg aac
!
!       346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
!       G   K   F   R   D   C   A   F   H   S   F   N   E   D
    5128   ggt aaa ttc aga gac tgc gct ttc cat tct ggc ttt aat gaG GAT
!                                       BamHI..
!
!       361 362 363 364 365 366 367 368 369 370 371 372 373 374 375
!       P   F   V   C   E   Y   Q   G   Q   S   S   D   L   P   Q
    5173   CCa ttc gtt tgt gaa tat caa ggc caa tcg tct gAC CTG Cct caa
! BamHI...                              BspMI...(2/2)
!
!       376 377 378 379 380 381 382 383 384 385 386 387 388 389 390
!       P   P   V   N   A   G   G   G   S   G   G   G   S   G   G
    5218   cct cct gtc aat gct ggc ggc ggc tct ggt ggtggt tct ggt ggc
!
!       391 392 393 394 395 396 397 398 399 400 401 402 403 404 405
!       G   S   E   G   G   G   S   E   G   G   G   S   E   G   G
    5263   ggc tct gag ggt ggc ggc tct gag ggt ggc ggt tct gac ggt ggc
!
!       406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
!       G   S   E   G   G   G   S   G   G   G   S   G   S   G   D
```

```
    5308   ggc tct gag ggt ggc ggt tcc ggt ggc ggc tcc ggt tcc ggt gat
!
!      421 422 423 424 425 426 427 428 429 430 431 432 433 434 435
!      F  D  Y  E  K  M  A  N  A  N  K  G  A  M  T
    5353   ttt gat tat gaa aaa atg gca aac gct aat aag ggg gct atg acc
!
!      436 437 438 439 440 441 442 443 444 445 446 447 448 449 450
!      E  N  A  D  E  N  A  L  Q  S  D  A  K  G  K
    5398   gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc aaa
!
!      451 452 453 454 455 456 457 458 459 460 461 462 463 464 465
!      L  D  S  V  A  T  D  Y  G  A  A  I  D  G  F
    5443 ctt gat tct gtc gct act gat tac ggt gct gct ATC GAT ggt ttc
!                                    BspDI..
!
!      466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
!      I  G  D  V  S  G  L  A  N  G  N  G  A  T  G
    5488   att ggt tcc ggc ctt gct aat ggt aat ggt gct act ggt
!
!      481 482 483 484 485 486 487 488 489490  491 492 493 494 495
!      D  F  A  G  S  N  S  Q  M  A  Q  V  G  D  G
    5533   gat ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt
!
!      496 497 498 499 500 501 502 503 504 505 506 507 508 509 510
!      D  N  S  P  L  M  N  N  F  R  Q  Y  L  P  S
    5578   gat aat tca cct tta atg aat aat ttc cgt caa tat tta cct tct
!
!      511 512 513 514 515 516 517 518 519 520 521 522 523 524 525
!      L  P  Q  S  V  E  C  R  P  Y  V  F  G  A  G
    5623   ttg cct cag tcg gtt gaa tgt cgc cct tat gtc ttt ggc gct ggt
!
!      526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
!      K  P  Y  E  F  S  I  D  C  D  K  I  N  L  F
5668   aaa cCA TAT Gaa ttt tct att gat tgt gac aaa ata aac tta ttc
!      NdeI...
!
!      541 542 543 544 545 546 547 548 549 550 551 552 553 554 555
!      R  G  V  F  A  F  L  L  Y  V  A  T  F  M  Y
    5713   cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat
!
```

```
!       556 557 558 559 560561  562 563 564 565 566 567 568 569 570
!       V  F  S  T  F  A  N  I  L  R  N  K  E  S
        5758   gta ttt tcg acg ttt gct aac ata ctg cgt aat aag gag tct taa
!

!       571
!
    5803   taa GAATTC
!          EcoRI.
    5812   actggccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc
    5871   gccttgcag acatcccccc ttcgccagct ggcgtaatag cgaagaggcc cgcacCGATC
!                               Pvul..
    5931   Gcccttccca acagtTGCGC Agcctgaatg gcgaatGGCG CCtgatgcgg tattttctcc
!...Pvul...(3/3)  fspl...(2/2)      Kasl...(2/@)
    5991   ttacgcatct gtgcggtatt tcacaccgca tataaattgt aaacgttaat attttgttaa
    6051   aattcgcgtt aaatttttgt taaatcagct catttttaa ccaataggcc gaaatcggca
    6111   aaatcccTTA TAAatcaaaa gaatagcccg agatagggtt gagtgttgtt ccagtttgga
!          Psil...
    6171   acaagagtcc actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc
    6231   agggcgatgg ccCACtacGT Gaaccatcac ccaaatcaag ttttttgggg tcgaggtgcc
!                 Dralll....
    6291   gtaaagcact aaatcggaac cctaaaggga gcccccgatt tagacttga cggggaaaGC
!                                          NgoMIV..
    6351   CGGCgaacgt ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg
!   ..NgoMIV.(2/2)
    6411   caagtgtagc ggtcacgvtg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac
    6471   agggcgcgta ctatggttgc tttgacgggt gactctcag tacaatctgc tctgatgccg
    6531   catagttaag ccagccccga cacccgccaa cacccgctga cgcgccctga cgggcttgtc
    6591   tgctccgcc atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga
    6651   ggttttcacc gtcatcaccg aaacgcgcga
```

Tabela 30: Oligonucleótidos utilizados para clonar diversidade CDR1/2

Todas as Sequências são de 5' a 3'.

1) ON_CD1Bsp, 30 bases

A c c T c A c T g g c T T c c g g A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

T T c A c T T T c T c T

19 20 21 22 23 24 25 26 27 28 29 30

2) ON_Br12, 42 bases

A g A A A c c c A c T c c A A A c c

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

T T T A c c A g g A g c T T g g c g

19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

A A c c c A

37 38 39 40 41 42

3) ON_CD2Xba. 51 bases

g g A A g g c A g T g A T c T A g A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

g A T A g T G A A g c g A c c T T T

19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36

A A c g g A g T c A g c A T A

37 38 39 40 41 42 43 44 45 46 47 48 49 50 51

4) ON_BotXba, 23 bases

g g A A g g c A g T g A T c T A g A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

g A T A g

19 20 21 22 23

Tabela 31: Oligonucleótidos de Ponte/Extensor

ON_Lam1aB7(rc) GTGCTGACTCAGCCACCCTC.20
ON_Lam2aB7(rc) .....GCCCTGACTCAGCCTGCCTC. 20
ON_Lam31B7(rc) .....GAGCTGACTCAGG.ACCCTGC 20
ON_Lam3rB7(rc) .....GAGCTGACTCAGCCACCCTC. 20
ON_LamHf1cBrg(rc) CCTCGACAGCGAAGTGCACAGAGCGTCTTGACTCAGCC.....38
ON_LamHf1cExt CCTCGACAGCGAAGTGCACAGAGCGTCTTG..... 30
ON_LamHf2b2Brg CCTCGACAGCGAAGTGCACAGAGCGCTTTGACTCAGCC..... 38
(rc)
ON_LamHf2b2Ext CCTCGACAGCGAAGTGCACAGAGCGCTTTG..... 30
ON_LamHf2dBrg(rc) CCTCGACAGCTAAGTGCACAGAGCGCTTTGACTCAGCC.... 38
ON_LamHf2dExt CCTCGACAGCGAAGTGCACAGAGCGCTTTG..... 30
ON_LamHf31Brg(rc) CCTCGACAGCGAAGTGCACAGAGCGAATTGACTCAGCC.....38
ON_LamHf31Ext CCTCGACAGCGAAGTGCACAGAGCGAATTG..... 30
ON_LamHf3rBrg(rc) CCTCGACAGCGAAGTGCACAGTACGAATTGACTCAGCC.....38
ON_LamHf3rExt CCTCGACAGCGAAGTGCACAGTACGAATTG..... 30
ON_lamPlePCR CCTCGACAGCGAAGTGCACAG..... 21
Consenso

Tabela 32: Oligonucleótidos utilizados para fazer ADN cs localmente em cadeia dupla

Adaptadores (8)
H43HF3.1?02#1 5'-cc gtg tat tac tgt gcg aga g-3'
H43.77.97.1-03#2 5'-cc gtg tat tac tgt gcg aga g-3'
H43.77.97.323#22 5'-cc gta tat tac tgt gcg aaa g-3'
H43.77.97.330#23 5'-ct gtg tat tac tgt gcg aaa g-3'
H43.77.97.439#44 5'-ct gtg tat tac tgt gcg aga c-3'
H43.77.97.551#48 5'-cc atg tat tac tgt gcg aga c-3'

Tabela 33: Pares de Ponte/Extensor

Pontes (2)
H43.XABr1
5'ggtgtagtgaTCTAGtgacaactctaagaatactctctacttgcagatgaacagCTTtAGgg ctgaggacaCTGCAGtctactattgtgcgaga-3'
H43.XABr2
5'ggtgtagtgaTCTAGtgacaactctaagaatactctctacttgcagatgaacagCTTtAGgg ctgaggacaCTGCAGtctactattgtgcgaaa-3'
Extensor
H43.XAExt
5'ATAgTAgAcTgcAgTgTccTcAgcccTTAAgcTgTTcATcATcTgAAgTAgAgAgTATTcTTAg AgTTgTcTcTAgATcAcTAcAcc-3'

Tabela 34: Iniciadores de PCR

Iniciadores	
H43.XAPCR2	gactgggTgTAgTgATcTAg
Hucmnest	cttttctttgttgccgttggggtg

Tabela 35

: programa de PCR para amplificação de cadeia pesada de CDR3

	ADN	
95 graus C	5 minutos	
95 graus C	20 segundos	
65 graus C	30 segundos	repetir 20x
72 graus C	1 minuto	
72 graus C	7 minutos	
4 graus C	manter	

Reagentes (reacção 100 uL):	
Molde	Mistura de ligaçaõ de 5 uL
tampão de PCR 10x turbo	1x
Taq	5U
dNTPs	cada 200 µM
MgCl2	2mM
H43.XAPCR2-biotina	400 nM
Hucmnest	200 nM

```
! Tabela 36: Sequência Anotada de CJR DY3F7(CJr-Ao5) 10251 bases
!
! Não-cortantes
!
!BclI Tgatca             BsiWI Cgtacg          BssSI Cacgag
!BstZ17I GTAtac          BtrI CACgtg           EcoRV GATatc
!FseI GGCCGGcc           HpaI GTTaac           MluI Acgcgt
!PmeI GTTTaaac           PmlI CACgtg           PpuMI RGgwccy
!RsrII CGgwccg           SapI GCTCTTC          SexAI Accwggt
!SgfI GCGATcgc           SgeAI CRccggyg        SphI GCATGc
!StuI AGGcct             XmaI Cccggg
!
!cortantes
!
!Enzimas que cortam de 1 a 4 vezes e outras características
! do alvo
!Fim dos genes II e X          829
!Início do gene V                        843
!BsrGI Tgtaca            1              1021
!BspMI Nnnnnnnnngcaggt   3              1104   5997   9183
!-"-   ACCTGCNNNNn       1              2281
!Fim do gene V                 1106
! Início do gene  VII                    1108
!BsaBI GATNNnnatc        2              1149   3967
! Início do gene IX                     1208
!Fim do gene  Vii                1211
!SnaBI TACgta            2              1268   7133
!BspHI Tcatga            3              1299   6085   7093
!Início do gene  VIII                   1301
!Fim do gene IX                  1304
!Fim do gene  VIII               1522
! Início do gene  III                    1578
!EagI Cggccg             2              1630   8905
!XbaI Tctaga             2              1643   8436
!KasI Ggcgcc             4              1650   8724   9039   9120
!BsmI GAATGCN            2              1769   9065
!BseRI GAGGAGNNNNNNNNNN  2              2031   8516
!-"-   NNnnnnnnnnctcctc  2              7603   8623
!aLWni cagnnnCTG         3              2210   8072   8182
!BspDI ATcgat            2              2520   9883
!NdeI CAtag              3              2716   3796   9847
```

!Fim do gene III		2846				
! Início do gene VI			2848			
!AfeI AGCgct	1		3032			
!Fim do gene VI		3187				
!Início do gene I			3189			
!EarI CTCTTCNnnn	2		4067	9274		
!-"- Nnnnngaagag	2		6126	8953		
!PacITTAATtaa	1		4125			
!Início do gene IV			4213			
!Fim do gene I		4235				
!BsmFI Nnnnnnnnnnnnnngtccc	2		5068	9515		
!MscI TGGcca	3		5073	7597	9160	
!PsiI TTAtaa	2		5349	5837		
!Fim do gene IV		5493				
!Início ori			5494			
!NgoMIV Gccgcc	3		5606	8213	9315	
!BanII GRGCYc	4		5636	8080	8606	8889
!DraIII CACNNNgtg	1		5709			
!DrdI GACNNNNnngtc	1		5752			
!AvaI Cycgrg	2		5818	7240		
!PvuII CAGctg	1		5953			
!BsmBI CGTCTCNnnnn	3		5964	8585	9271	
!Fim de região ori			5993			
!BamHI Ggatcc	1		5994			
!HindIII Aagcc	3		6000	7147	7384	
!BciVI GTATCCNNNNNN	1		6077			
!Início bla			6138			
!Eco57I CTGAAG	2		6237	7716		
!SpeI Actagt	1		6257			
!BcgI gcannnnnntcg	1		6398			
!ScaI AGTact	1		6442			
!PvuI CGATcg	1		6553			
!FspI TGCgca	1		6700			
!BglI GCCNNNNnggc	3		6801	8208	8976	
!BsaI GGTCTCNnnnn	1		6853			
!AhdI GACNNNnngtc	1		6920			
!Eam1105I GACNNNnngtc	1		6920			
!End bla			6998			
!AccI GTmkac	2		7153	8048		
!HincII GTY rac	1		7153			
!SalI Gtcgac	1		7153			

!ChoI Ctcgag	1		7240		
!início da região PlacZ			7246		
!Fim da região PlacZ		7381			
!PflMI CCANNNNntgg	1		7382		
!RBSi			7405		
!Início da seq de sinal M13-iii para LC			7818		
!ApaLI Gtgcac	1		7470		
! Fim da seq de sinal M13-iii		7471			
!início da cadeia leve kappa L20:JK1			7472		
!PflFI GACNnngtc	3		7489	8705	9099
!SbfI CCTGCAgg	1		7542		
!PstI CTGCAg	1		7543		
!KpnI GGTACc	1		7581		
!XcmI CCANNNNNnnnntgg	2		7585	9215	
!NsiI ATGCAt	2		7626	9503	
!BsgI ctgcac`	1		7809		
!BbsI gtcttc	2		7820	8616	
!BlpI GCTnagc	1		8017		
!EspI GCTnagc	1		8017		
!Eco019I RGgnccy	2		8073	8605	
!Ecl136I GAGctc	1		8080		
!SacI GAGcTc	1		8080		
!Fim da cadeia leve		8122			
!AscI GGcgcgcc	1		8126		
!BssHII Gcgcgc	1		8127		
!RBS2			8147		
!SfiI GGCCNNNNNnggcc	1		8207		
!NcoI Ccatgg	1		8218		
!Início 3-23,FR1			8226		
!MfeI Caattg	1		8232		
!BspEI Tccgga	1		8298		
!Início CDR1			8316		
!Statt FR2			8331		
!BstCi CCANNNNNntgg	2		8339	8812	
!EcoNI CCTNNnnnagg	2		8346	8675	
!Início FR3			8373		
!XbaI Tctaga	2		8436	1643	
!AflII Cttaag	1		8480		
!Início CDR3			8520		
!AatII GaCGTc	1		8556		
!Início FR4			8562		

```
!PshAI GACNNnngtc            2             8573   9231
!BstEII Ggtnacc              1             8579
!Início CH1                                8595
!ApaI GGGCCc                 1             8606
!Bsp120I Gggccc              1             8606
!PspOMI Gggccc               1             8606
!AgeI Accggt                 1             8699
!Bsu 36I CCtnagg             2             8770   9509
!Fim de CH1                          8903
!NotI GCggccgc               1             8904
!Início do alvo His6                       8913
!Início do alvo cMyc                       8931
!Amber codão                               8982
!NheI Gctagc                 1             8985
!Início de Domínio 3M13 III                8997
!NruI TCGcga                 1             9106
!BstBI TTcgaa                1             9197
!EcoRI Gaattc                1             9200
!XcmI CCANNNNNnnnntgg        1             9215
!BstAPI GCANNNNntgc          1             9337
!SacII CCGCgg                1             9365
!Fim da âncora IIIstump                    9455
!AvrII Cctagg                1             9462
! terminal trp                             9470
!SwaI ATTTaaat               1             9784
!Início do gene  II                        9850
!BglII Agatct                1             9936
!-----------------------------------------------------------------
    1 aat gct act act att agt aga att gat gcc acc ttt tca gct cgc gcc
!   gene ii continuou
   49 cca aat gaa aat ata gct aaa cag gtt att gac cat ttg cga aat gta
   97  tct aat ggt caaact aaa tct act cgt tcg cag aat tgg gaa tca act
  145 gtt aTa tgg aat gaa act tcc aga cac cgt act tta gtt gca tat tta
  193 aaa cat gtt gag cta cag caT TaT att cag caa tta agc tct aag cca
  241 tcc gca aaa atg acc tct tat caa aag gag caa tta aag gta ctc tct
  289 aat cct gac ctg ttc gag ttt gct tcc ggt ctg gtt cgc ttt gaa gct
  337 cga att aaa acg cga tat ttg aac tct ttc ggg ctt cct ctt aat ctt
  385 ttt gat gca atc cgc ttt gct tct gac tat aat agt cag ggt aaa gac
  433 ctg att ttt gat tta tgg tca ttc tcg ttt tct gaa ctg ttt aaa gca
  481 ttt gag ggg gat tca ATG aat att tat gac gat tcc gca gta ttg gac
!                       Início do gene  x, ii continua
```

```
   529 gct atc cag tct aaa cat ttt act att acc ccc tct ggc aaa act tct
   577 ttt gca aaa gcc tct cgc tat ttt ggt ttt tat cgt cgt ctg gta aac
   625 gag ggt tat gat agt gtt gct ctt act atg cct cgt aat tcc ttt tgg
   673 cgt tat gta tct gca tta gtt gaa tgt ggt att cct aaa tct caa ctg
   721 atg aat ctt tct acc tgt aat aat gtt gtt ccg tta gtt cgt ttt att
   769 aac gta gat ttt tct caa cgt cct gac tgg tat aat gag cca gtt
   817 ctt aaa atc gca TAA
!                Fim de  X & II
   832 ggtaattca ca
!
!         M1            E5                Q10               T15
   843 ATG att aaa gtt gaa att aaa cca tct caa gcc caa ttt act act cgt
!      Início do gene  V
!
          S17    S20                  P25               E30
   891 tct ggt gtt tct cgt cag ggc aag cct tat tca ctg aat gag cag ctt
!
!             V35                  E40               V45
   939 tgt tac gtt gat ttg ggt aat gaa tat ccg gtt ctt gtc aag att act
!
!         D50                  A55               A60
   987 ctt gat gaa ggt cag cca gcc tat gcg cct ggt cTG TAC Acc gtt cat
!                BsrGI...
!         L65                  V70        S75                          R80
  1035   ctg tcc tct ttc aaa gtt ggt cag ttc ggt tcc ctt atg att gac cgt
!
!                       P85    K87 Fim de V
  1083 ctg cgc ctc gtt ccg gct aag TAA C
!
  1108   ATG gag cag gtc gcg gat ttc gac aca att tat cag gcg atg
!      Início do gene  VII
!
  1150 ata caa atc tcc gtt gta ctt tgt ttc gcg ctt ggt ata atc
!
!             VII e IX sobrepostos.
!             .....S2    V3  L4  V5                S10
  1192 gct ggg ggt caa agA TGA gt gtt tta gtg tat tct ttT gcc tct ttc gtt
!                   Fim de  VII
!                   | Início IX
!      L13    W15           G20    T25    E29
  1242 tta ggt tgg tgc ctt cgt agt ggc att acg tat ttt acc cgt tta atg gaa
```

```
!
   1293 act tcc tc
!
!         ....paragem de IX,IX e VIII sobrepostos por quatro bases
   1301 ATG aaa aag tct tta gtc ctc aaa gcc tct gta gcc gtt gct acc ctg
!      Início da sequência de sinal de viii.
!
   1349 gtt ccg atg ctg tct ttc gct gct gag ggt gac gat ccc gca aaa gcg
!                            VIII maduro--->
   1397 gcc ttt aac tcc ctg caa gcc tca gcg acc gaa tat atc ggt tat gcg
   1445 tgg gcg atg gtt gtt gtc att
   1466   gtc ggc gca act atc ggt atc aag ctg ttt aag
!
!Bases 1499-1539 são promotores prováveis para iii
   1499 aaa ttc acc tcg aaa gca !  1515
!      ..............-35 ..
!
   1517       agc tga taaaccgat acaattaag gctccttttg
!                    ......-10  ...
!
   1552 gagccttttt ttt GGAGAt ttt ! S.D. letra maiúscula, pode ser 9 Ts
!
!   <------ sequência de sinal III ------------------------------------>
!             M   K    K   L   L   F     A   I   P   L   V    V   P   F
   1574 caac gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc ! 1620
!
!        Y    S   G   A   A   E    S    H  L   D   G   A
   1620 tat tct ggc gCG GCC Gaa tca caT CTA GAc ggc gcc
!                EagI......        XbaI....
!
!Domínio 1 ------------------------------------------------------------
!          A    E  T   V   E   S   C   L   A
   1656   gct gaa act gtt gaa agt tgt tta gca
!

!          K   S   H   T   E   I   S   F   T   N   V   W   K   D   D   K  T
   1683   aaA Tcc cat aca gaa aat tca ttt aCT AAC GTC Tgg AAa GAC GAC AAA Act
!
!        L   D   R   Y   A   N   Y   E   G   S   L   W   N   A   T   G   V
   1734 tta gat cgt tac gct aac tat gag ggC tgt ctg tgg AAT GCt aca ggc gtt
!                                    BsmI...
```

```
!
!         V  V    C    T    G    D    E    T    Q    C    Y    G    T    W    V    P    I
    1785 gta gtt tgt act ggt GAC GAA ACT CAG TGT TAC GGT ACA TGG GTT cct att
!
!         G    L    A    I    P    E    N
    1836 ggg ctt gct atc cct gaa aat
!
!   L1 Elemento de ligação--------------------------------
              E    G    G    G    S    E    G    G    G    S
    1857   gag ggt ggt gcc tct gag ggt ggc gtgt tct
!
!           E    G    G    G    S    E    G    G    G    T
    1887   gag ggt ggc ggt tct gag ggt ggc ggt act
!
!Domínio 2 ------------------------
    1917   aaa cct cct gag tac ggt gat ac acct att ccg ggc tat act tat atc aac
    1968   cct ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct
    2019   aat cct tct ctt GAG GAG tct cag cct ctt aat act ttc atg ttt cag aat
!                        BseRI..
    2070   aat agg ttc cga aat agg cag ggg gca tta act gtt tat acg ggc act
    2118   gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act cct
    2166   gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttC AGA
!                                                                AlwNI
    2214 GAC TGc gct ttc cat tct ggc ttt aat gaG gat TTa ttT gtt tgt gaa
!       AlwNI
    2262 tat caa ggc caa tcg tct gac ctg cct ca acct cct gtc aat gct
!
    2307 ggc ggc ggc tct
!   Início L2---------------------------------------------------------------
    2319 ggt ggt ggt tct
    2331 ggt ggc ggc tct
    2343 gag ggt ggt ggc tct gag gga ggc ggt tcc
    2373 ggt ggt ggc tct ggt ! end L2
!
!muitas Sequências publicadas defagos derivados de M13 têm um elemento de
ligação maior do que o aqui mostrado por repetições do motivo EGGGS mais duas
vezes.
!
!Domínio 3 ------------------------------------------------------------
!         S    G    D    F    D    Y    E    K    M    A    N    A    N    K    G    A
    2388 tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat aag ggg gct
```

```
!
!        M   T   E   N   A   D   E   N   A   L   Q   S   D   A   K   G
    2436 atg acc gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc
!
!          K   L   D   S   V   A   T   D   Y   G   A   A   M   D   G   F
    2484   aaa ctt gat tct gtc gct act gat tac ggt gct gct atc gat ggt ttc
!
!          I   G   D   V   S   G   L   A   N   G   N   G   A   T   G   D
    2532   att ggt gac gtt tcc ggc ctt gct aat ggt aat ggt gct act ggt gat
!
!        F   A   G   S   N   S   Q   M   A   Q   V   G   D   G   D   N
    2580 ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt gat aat
!
!        S   P   L   M   N   N   F   R   Q   Y   L   P   S   L   P   Q
    2628 tca cct tta atg aat aat ttc cgt caa tat tta cct tcc ctc cct caa
!
!        S   V   E   C   R   P   F   V   F   G   A   G   K   P   Y   E
    2676 tcg gtt gaa tgt cgc cct ttc gtc ttt Ggc gct ggt aaa cca tat gaa
!
!          F   S   I   D   C   D   K   I   N   L   F   R
    2724   ttt tct att gat tgt gac aaa ata aac tta ttc cgt
!                              Fim de Domínio 3
!
!          G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V   F140
    2760   ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat gta ttt
           Início do segmento transmembranar
!
!        S   T   F   A   N   I   L
    2808 tct acg ttt gct aac ata ctg
!
!        R   N   K   E   S
    2829 cgt aat aag gag tct TAA ! paragem de iii
!       âncora intracelular.
!
!             M1  P2   V   L   L5   G   I    P   L   L10  L   R   F    L  G15
    2847   tc ATG cca gtt ctt ttg ggt att ccg tta tta ttg cgt ttc ctc ggt
!          Início VI
!
    2894   ttc ctt ctg gta act ttg ttc ggc tat ctg ctt act ttt ctt aaa aag
    2942   ggc ttc ggt aag ata gct att gct att tca ttg ttt ctt gct ctt att
    2990   att ggg ctt aac tca att ctt gtg ggt tat ctc tct gat att agc gct
```

```
   3038   caa tta ccc tct gac ttt gtt cag ggt gtt cag tta att ctc ccg tct
   3086   aat gcg ctt ccc tgt ttt tat gtt att ctc tct gta aag gct gct att
   3134   ttc att ttt gac gtt aa caa aaa atc gtt tct tat ttg gat tgg gat
!
!                   M1  A2  V3     F5                  L10          G13
   3182   aa TAA t ATG gct gtt tat ttt gtc act ggc aaa tta ggc tct gga
!             Fim de  VI     Início do gene  I
!
!           K   T   L   V   S   V   G   K   I   Q   D   K   I   V   A
   3228   aag acg ctc gtt agc gtt ggt aag att cag gat aaa att gta gct
!
!           G   C   K   I   A   T   N   L   D   L   R   L   Q   N   L
   3273   ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa aac ctc
!
!           P   Q   V   G   R   F   A   K   T   P   R   V   L   R   I
   3318   ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga ata
!
!           P   D   K   P   S   I   S   D   L   L   A   I   G   R   G
   3363   ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt
!
!           N   D   S   Y   D   E   N   K   N   G   L   L   V   L   D
   3408   aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat
!
!           E   C   G   T   W   F   N   T   R   S   W   N   D   K   E
   3453   gag tgc ggt act tgg ttt aat acc cgt tct tgg aat gat aag gaa
!
!           R   Q   P   I   I   D   W   F   L   H   A   R   K   L   G
   3498   aga cag ccg att att gat tgg ttt cta cat gct cgt aaa tta gga
!
!           W   D   I   I   F   L   V   Q   D   L   S   I   V   D   K
   3543   tgg gat att att ttt ctt gtt cag gac tta tct att gtt gat aaa
!
!       Q   A   R   S   A   L   A   E   H   V   V   Y   C   R   R
3588   cag gcg cgt tct gca tta gct gaa cat gtt gtt tat tgt cgt cgt
!
!           L   D   R   I   T   L   P   F   V   G   T   L   Y   S   L
   3633   ctg gac aga att act tta cct ttt gtc ggt act tta tat tct ctt
!
!           I   T   G   S   K   M   P   L   P   K   L   H   V   G   V
   3678   att act ggc tcg aaa atg cct ctg cct aaa tta cat gtt ggc gtt
!
```

```
!          V   K   Y   G   D   S   Q   L   S   P   T   V   E   R   W
    3723   gtt aaa tat ggc gat tct caa tta agc cct act gtt gag cgt tgg
!
!          L   Y   T   G   K   N   L   Y   N   A   Y   D   T   K   Q
    3768   ctt tat act ggt aag aat ttg tat aac gca tat gat act aaa cag
!
!        A   F   S   S   N   Y   D   S   G   V   Y   S   Y   L   T
    3813 gct ttt tct agt aat tat gat tcc ggt gtt tat tct tat tta acg
!
!        P   Y   L   S   H   G   R   Y   F   K   P   L   N   L   G
    3858 cct tat tta tca cac ggt cgg tat ttc aaa cca tta aat tta ggt
!
!          Q   K   M   K   L   T   K   I   Y   L   K   K   F   S   R
    3903   cag aag atg aaa tta act aaa ata tat ttg aaa aag ttt tct cgc
!
!          C   L   C   L   A   I   G   F   A   S   A   F   T   Y   S
    3948   gtt ctt tgt ctt gcg att gga ttt gca tca gca ttt aca tat agt
!
!          Y   I   T   Q   P   K   P   E   V   K   K   V   V   S   Q
    3993   tat ata acc caa cct aag ccg gag gtt aaa aag gta gtc tct cag
!
!          T   Y   D   F   D   K   F   T   I   D   S   S   Q   R   L
    4038   acc tat gat ttt gat  ttc act att gac tct tct cag cgt ctt
!
!          N   L   S   Y   R   Y   V   F   K   D   S   K   G   K   L
    4083   aat cta agc tat cgc tat gtt ttc aag gat tct aag gga aaa TTA
!                                                                PacI
!
!        I   N   S   D   D   L   Q   K   Q   G   Y   S   L   T   Y
    4128 ATT AAt agc gac gat tta cag aag caa ggt tat tca ctc aca tat
!       PacI
!       I  I   D   L   C   T   V   S   I   K   K   G   N   S   N   E
!   iv                                                          M1  K
    4173  att gat tta tgt act gtt tcc att aaa aaa ggt aat tca aAT Gaa
!                                                      Início de IV
!   I   IVKCN.Fim de I
!   iv L3  L  N5  V  I7  N  F  V10
    4218   att gtt aaa tgt aat ATT T TTT GTT
!   IV continuou.....
    4243 ttc ttg atg ttt gtt tca tca tct tct ttt gct cag gta att gaa atg
    4291   aat aat tcg cct ctg cgc gat ttt gta act tgg tat tca aag caa tca
```

```
    4339   ggc gaa tcc gtt att gtt tct ccc gat gta aaa ggt act gtt act gta
    4387   tat tca tct gac gtt aaa cct gaa aat cta cgc aat ttc ttt att tct
    4435   gtt tta cgt gcA aat aat ttt gat atg gA ggt tcT aAC cct tcc atT
    4483   att cag aag tat aat cca aac aat cag gat tat att gat gaa ttg cca
    4531   tca tct gat aat cag gaa tat gat gat aat tcc gct cct tct ggt ggt
    4579   ttc ttt gtt ccg caa aat gat aat gtt act caa act tt aaa att aat
    4627   aac gtt cgg gca aag gat tta ata cga gtt gtc gaa ttg ttt gta aag
    4675   tct aat act tct aaa tcc tca aat gta tta tct att gac ggc tct aat
    4723   cta tta gtt gtt agt gcT cct aaa gat att tta gat aac ctt cct caa
    4771   ttc ctt tcA act gtt gat ttg cca act gac cag ata ttg att gac ggt
    4819   ttg ata ttt gag gtt cag caa ggt gat gct tta gat ttt tca ttt gct
    4867   gct ggc tct cag cgt ggc act gtt gca ggc ggt gtt aat act gac cgc
    4915   ctc acc tct gtt tta tct tct gct ggt ggt tcg ttc ggt att ttt aat
    4963   ggc gat gtt tta ggg cta tca gttcgc gca tta aag act aat agc cat
    5011   tca aa ata ttg tct gtg cca cgt att ctt acg ctt tca ggt cag aag
    5059   ggt tct atc tct gtT GGC CAg aat gtc cct ttt att act ggt cgt gtg
!                          MacI....
    5107 act ggt gaa tct gcc aat gta aat aat cca ttt cag acg att gag cgt
    5155   caa aat gta ggt att tcc atg agc gt ttt cct gtt gca atg gct ggc
    5203   ggt aat att gtt ctg gat att acc agc aag gcc gat agt ttg agt tct
    5251   tct act cag gca agt gat gtt att act aat caa aga agt att gct aca
    5299   acg gtt aat ttg cgt gat gga cag act ctt tta ctc ggt ggc ctc act
    5347   gat tat aaa aac act tct caG gat tct ggc gta ccg ttc ctg tct aaa
    5395   atc cct tta atc ggc ctc ctg ttt agc tcc cgc tct gat tcT aac gag
    5443   gaa agc acg tta tac gtg ctc gtc aaa gca acc ata gta cgc gcc ctg
    5491   TAG CGGCGCATT
!      Fim de  IV
    5503   aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgacca gcgccctagc
    5563   gcccgctcct ttcgcttct tcccttcctt tctcgcaacg ttcGCCGGCt ttccccgtca
!                                          NgoMI.
    5623   agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc
    5683   caaaaaactt gatttgggtg atggttCACG TAGTGggcca tcgccctgat agacggtttt
!                              DraIII....
    5743   tcgccctttG ACGTTGGAGT Ccacgttctt taatagtgga ctcttgttcc aaactggaac
!                     DrdI......
    5803   aacactccac cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcgga
    5863   accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa
    5923   ctctctcagg gccaggcggt gaagggcaat CAGCTGttgc cCGTCTCact ggtgaaaaga
!                                    PvuII.     BsmBI.
    5983   aaaaccaccc tGGATCC AAGCTT
!                     BamHI      HindIII (1/2)
```

```
!                    inserir transportador do gene bla
    6006       gcaggtg gcacttttcg gggaaatgtg cgcggaaccc
    6043    ctatttgttt atttttctaa atacattcaa atatGTATCC gctcatgaga caataaccct
!                   BciVI
    6103    gataatgct tcaataatat tgaaaAGGA AGAgt
!                           RBS.?...
!       Início gene bla
    6138    ATG agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca ttt
    6189    tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aa gta aaa gat gct
    6240    gaa gat cag ttg ggC gcA CTA GTg ggt tac atc gaa ctg gat ctc aac agc
!                           SpeI....
!                   ApaLI & BssSI removido
    6291    ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc
    6342    act tt aaa gtt ctg cta tgt GGC GcG Gta tta tcc cgt att gac gcc ggg
    6393    caa gaG CAA CTC GGT CGc cgC ATA cAC tat tct cag aat gac ttg gtt gAG
!                        BcgI..................                      ScaI
    6444    TAC Tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa
!           ScaI.
    6495    tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt
    6546    ctg aca aCG ATC Gga gga ccg aag gag cta acc gct ttt ttg cac aac atg
!              PvuI....
    6597    ggg gat cat gta act cgc ctt gat cgt  tgg gaa ccg gag ctg aat gaa gcc
    6648    ata cca aac gac gag cgt gac acc acg tag cct gta gca atg Gca aca acg
    6699    tTG CGC Aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa caa
!                  FspI....
!
    6750 tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg
    6801    GCC ctt ccG GCt ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt
!           BglI........
    6852    gGG TCT Cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt
!           BsaI....
    6903    atc gta gtt atc tac acG ACg ggg aGT Cag gca act atg gat gaa cga aat
!                        AhdI..................
    6954    aga cag atc gct gag ata ggt gcc tca crg att aag cat tgg TAA ctgt
!                                                                 paragem
    7003    cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa
    7063    ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt
    7123    cgttccactg tacgtaagac cccc
    7147    AAGCTT GTCGAC tgaa tggcgaatgg cgctttgcct
!           HindIIISalI..
!           (2/2)  HincII
```

```
    7183   ggtttccggc accagaagcg gtgccggaaa actggctgga gtgcgatctt
!
!   Início do invólucro de apresentação de Fab, o Fab DSR-A05, seleccionado para
se ligar a um antigénio de proteína.
!
    7233 CCTGAcG CTCGAG
!          xBau36I XhoI..
!
!   o promoter PlacZ está no bloco seguinte
    7246                          cgcaacgc aattaatgtg agttagctca
    7274       ctcattaggc accccaggct ttacactta tgcttccggc tcgtatgttg
    7324       tgtggaatg tgagcggata acaatttcac acaggaaaca gctatgacca
    7374       tgattacgGG AagcttTGGa gccttttt tggagatttt caac
!                    PflMI......
!                        Hind3. (Existem 3)
! sequência de sinal de Gene iii:
!              1   2   3   4   5   6   7   8   9   10  11  12  13  14  15
!           M   K   K   L   L   F   A   I   P   L   V   V   P   F   Y
    7418   gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat
!
!          16  17  18         Início de cadeia leve (L20:JK1)
!           S   H   S   A   Q   D   I   Q   M   T   Q   S   P   A
    7463   tct cag aGT GCA Caa gac atc cag atg acc cag tct cca gcc
!                    ApaLI...
!                    Sequência fornecida pelo Extensor.........
!
!              T   L   S   L
    7505       acc ctg tct ttg
!
!           S   P   G   E   R   A   T   L   S   C   R   A   S   Q   G
    7517   tct cca ggg gaa aga gcc acc ctc tcc tgc agg gcc agt cag Ggt
!
!           V   S   S   Y   L   A   W   Y   Q   Q   K   P   G   Q   A
    7562   gtt agc agc tac tta gcc tgg tac cag cag aaa cct ggc cag gct
!
!           P   R   L   L   I   Y   D   A   S   S   R   A   T   G   I
    7607   ccc agg ctc ctc atc tat gAt gca tcc aAc agg gcc act ggc atc
!
!           P   A   R   F   S   G   S   G   P   G   T   D   F   T   L
    7652   cca gCc agg ttc agt ggc agt ggg Cct ggg aca gac ttc act ctc
!
```

```
!          T   I   S   S   L   E   P   E   D   F   A   V   Y   Y   C
   7697   acc atc agc agC ctA gag cct gaa gat ttt gca gtT tat tac tgt
!
!          Q   Q   R   S   W   H   P   W   T   F   G   Q   G   T   R
   7742   cag cag CGt aAc tgg cat ccg tgg ACG TTC GGC CAA GGG ACC AAG
!          V   E   I   K   R   T   V   A   A   P   S   V   F   I   F
   7787   gtg gaa atc aaa cga act gtg gCT GCA Cca tct gtc ttc atc ttc
!                                      BsgI....
!          P   P   S   D   E   Q   L   K   S   G   T   A   S   V   V
   7832   ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt gtg
!
!          C   L   L   N   N   F   Y   P   R   E   A   K   V   Q   W
   7877   tgc ctg ctg aat aac ttc tat ccc aga gag gcc aa gta cag tgg
!
!          K   V   A   D   N   L   Q   S   G   N   S   Q   E   S   V
   7922   aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc
!
!          T   E   R   D   S   K   D   S   T   Y   S   L   S   S   T
   7967   aca gag cgg gac agc aag gac agc acc tac agc ctc agc agc acc
!
!          L   T   L   S   K   A   D   Y   E   K   H   K   V   Y   A
   8012   ctg acG CTG AGC aaa gca gac tac gag aaa cac aaa gtc tac gcc
!             EspI.....
!
!          C   E   V   T   H   Q   G   L   S   S   P   V   T   K   S
   8057   tgc gaa gtc acc cat cag ggc ctG AGC TCg ccc gtc aca aag agc
!                        SacI....
!
!          F   N   R   G   E   C   .   .
   8102   ttc aac agg gga gag tgt taa taa
!
   8126   GGCGCG CCaattctat ttcaa GGAGA cagtcata
!         AscI.....             RBS2.
!
!      sequência de sinal PelB -----(22 codões)------->
!      1   2   3   4   5   6   7   8   9   10  11  12  13  14  15
!    M  K   Y   L   L   P  T  A   A   A   G   L   L   L   L
   8160   atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc
!
!      ...PelB sinal --------------> Início VH, FR1------------->
!      16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
```

```
!      A   A   Q   P   A   M   A   E   V   Q   L   L   E   S   G
   gcG GCC cag ccG GCC atg gcc gaa gtt CAA TTG tta gag tct ggt
!         SfiI.........       MfeI...
!                   NcoI....
!
!      31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
!           G   G   L   V   Q   P   G   G   S   L   R   L   S   C   A
   8250   ggc ggt ctt gtt cag cct ggt ggt tct tta cgt ctt tct tgc gct
!
!      ...FR1------------>   CDR1-------------> FR2-------->
!      46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
!           A   S   G   F   T   F   S   T   Y   E   M   R   W   V   R
   8295   gct TCC GGA ttc act ttc tct act tac gag atg cgt tgg gtt cgC
!             BspEI..                                          BstXI...
!
!         FR2-------------------------------> CDR2-------------->
!      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
!           Q   A   P   G   K   G   L   E   W   V   S   Y   I   A   P
   8340   CAa gct cct GGt aaa ggt ttg gag tgg gtt tct tat atc gct cct
!  BstCI.........
!
!      ...CDR2---------------------------------> FR3---->
!      76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
!           S   G   G   D   T   A   Y   A   D   S   V   K   G   R   F
   8385   tct ggt ggc gat act gct tat gct gac tcc gtt aaa ggt cgc ttc
!
!         91  92  93  94  95  96  97  98  99  100 101 102 103 104 105
!          T   I   S   R   D   N   S   K   N   T   L   Y   L   Q   M
   8430   act atc TCT AGA gac aac tct aag aat act ctc tac ttg cag atg
!                XbaI...
!                Fornecido pelo Extensor ----------------------------
!
!         ---------------------------------------------FR3----------->
!         106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!          N   S   L   R   A   E   D   T   A   V   Y   Y   C   A   R
   8475   aac agC TTA AGc gct gag gac act gca gtc tac tat tgt gcg agg
!         AflII...
!         de Extensor ---------------------------->
!
!         CDR3-------------------------------------------------> FR4-->
!         121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
```

```
!          R   L   D   G   Y   I   S   Y   Y   Y   G   M   D   V   W
   8520   agg ctc gat ggc tat att tcc tac tac tac ggt atg GAC GTC tgg
!                                       AatII..
!
!          126 137 138 139 140 141 142 143 144 145
!           G   Q   G   T   T   V   T   V   S   S
   8565   ggc caa ggg acc acG GTC ACC gtc tca agc
!             BstEII...
!
!             CH1 de IgG1---------->
!          A   S   T   K   G   P   S   V   F   P   L   A   P   S   S
   8595   gcc tcc acc aag ggc cca tcg gtc ttc ccc ctg gca ccc tcc tcc
!
!          K   S   T   S   G   G   T   A   A   L   G   C   L   V   K
   8640   aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc ctg gtc aag
!
!          D   Y   F   P   E   P   V   T   V   S   W   N   S   G   A
   8685   gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg aac tca ggc gcc
!          L   T   S   G   V   H   T   F   P   A   V   L   Q   S   S
   8730   ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta gac tCC TCA
!                                                           Bsu36I....
!
!          G   L   Y   S   L   S   S   V   V   T   V   P   S   S   S
   8775   GGa ctc tac tcc ctc agc agc gta gtg acc gtg ccc tcc agc agc
!  Bsu36I....
!          L   G   T   Q   T   Y   I   C   N   V   N   H   K   P   S
   8820   ttg ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc agc
!
!       N   T   K   V   D   K   K   V   E   P   K   S   C   A   A
8865   aac acc aag gtg gac aag aaa gtt gag ccc aaa tct tgt GCG GCC
!                                                       NotI.....
!
!          A   H   H   H   H   H   H   G   A   A   E   Q   K   L   I
   8910   GCa cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc atc
!  ..NotI.... H6 tag..................... Myc-Tag..................
!
!          S   E   E   D   L   N   G   A   A   q   A   S   S   A
   8955   tca gaa gag gat ctg aat ggg gcc gca tag GCT AGC tct gct
!         Myc-Tag.................      ..NheI...
!                              Amber
!
```

```
!   III'stump
!
!Domínio 3 de III ------------------------------------------------
!
!        S   G   D   F   D   Y   E   K   M   A   N   A   N   K   G   A
    8997 agt ggc gac ttc gac tac gac aaa atg gct aat gcc aac aaa GGC GCC
!       tcc  t   t   t   t   t   a   g       a   c   t   t   g   g   t   !W.T.
!                                                KasI...(2/4)
!
!          M  T   E   N   A   D   E   N   A   L   Q   S   D   A   K   G
    9045 atG ACT GAG AAC GCT GAC GAG aat gct ttg caa agc gat gcc aag ggt
!             c   a   t   c   t   a   c g c a g tct  c   t   a     c !W.T.
!
!           K   L   D   S   V   A   T   D   Y   G   A   A   I   D   G   F
    9093    aag tta gac agc gTC GCG Acc gac tat GGC GCC gcc ATC GAc ggc ttt
!           a c t  t tct         t   t   t   c   t   t   t         t   t   c
    !W.T.
!                    NruI...        KasI...(3/4)
!
!           I   G   D   V   S   G   L   A   N   G   N   G   A   T   G   D
    9141    atc ggc gat gtc agt ggt tTG GCC Aac ggc aac gga gcc acc gga gac
!           t   t   c   t   tcc  c ct   t   t   t   t   t   t   t   t   t!W.T.
!                    MscI....(3/3)
!           F   A   G   S   N   S   Q   M   A   Q   V   G   D   G   D   N
    9189    ttc GCA GGT tcG AAT TCt cag atg gaC CAG GTT GGA GAT GGg gac aac
!           t   t   c   t       c   a       t   a   c   t   c   t   t   t  !W.T.
!       BspMI.. (2/2)                      XcmI.........................
!              EcoRI...
!           S   P   L   M   N   N   F   R   Q   Y   L   P   S   L   P   Q
    9237    agt ccg ctt atg aac aac ttt aga cag tac ctt ccg tct ctt ccg cag
!           tca   t t a     t   t   c c t   a   t t a   t   c   c   t   a  !W.T.
!
!           S   V   E   C   R   P   F   V   F   S   A   G   K   P   Y   E
    9285    agt gtc gag tgc cgt cca ttc gtt ttc tct gcc ggc aag cct tac gag
!           tcg t   a   t   c   t   t   c   t agc   t   t   a   a   t   a !W.T.
!
!         F   S   I   D   C   D   K   I   N   L   F   R
    9333 ttc aGC Atc gac TGC gat aag atc aat ctt ttC CGC
!          t tct   t   t   t   c   a   a   c t a   c   t !W.T.
!       BstAPI.......                  SacII...
!                                          Fim do Domínio 3
```

```
!
!          G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V   F
    9369   GGc gtt ttc gct ttc ttg cta tac gtc gct act ttc atg tac gtt ttc
!          t   c   t   g   t c t t a   t   t   c   c   t     t   a   t
    !W.T.
!       Início do segmento transmembranar
!
!          S   T   F   A   N   I   L   R   N   K   E   S
    9417   aGC ACT TTC GCC AAT ATT TTA  Cgc aac aaa gaa agc
!          tct  g   t   t   c    a c g   t   t   g   g  tct  !W.T.
!                              Âncora intracelular.
!
!          .   .
    9453   tag tga tct CCT AGG
!                 AvrII..
!
    9468   aag ccc gcc taa tga gcg ggc ttt ttt ttt ct ggt
!       Trp terminal
!
! Fim do invólucro de Fab
!
    9503   ATGCAT CCTGAGG ccgat actgtcgtcg tccccctcaaa ctggcagatg
!       NsiI.. Bsu36I.(3/3)
    9551   cacggttacg atgcgcccat ctacaccaac gtgacctatc ccattacggt caatccgccg
    9611   tttgttccca cggagaatcc gacgggttgt tactcgctca catttaatgt tgatgaaagc
    9671   tggctacagg aaggccagac gcgaattatt tttgatggcg ttcctattgg ttaaaaaatg
    9731   agctgattta acaaaaattt aaTgcgaatt ttaacaaaat attaacgttt acaATTTAAA
!                                                                SwaI...
    9791   Tatttgctta tacaatcttc ctgtttttgg ggcttttctg attatcaacc GGGGTAcat
    9850   ATG att gac atg cta gtt tta cga tta ccg ttc atc gat tct ctt gtt tgc
!          Início do gene  II
    9901   tcc aga ctc tca ggc aat gac ctg ata gcc ttt gtA GAT CTc tca aaa ata
!                                              BglII...
    9952   gct acc ctc tcc ggc atT aat tta tca gct aga acg gtt gaa tat cat att
    10003  gat ggt gat ttg act gtc tcc ggc ctt tct cac cct ttt gaa tct tta cct
    10054  aca cat tact ca ggc att gca ttt aaa ata tat gag ggt tct aaa aat ttt
    10105  tat cct tgc gtt gaa ata aag gct tct ccc gca aaa gta tta cag ggt cat
    10105  tat cct tgc gtt gaa ata aag gct tct ccc gca aaa gta tta cag ggt cat
    10156  aat gtt ttt ggt aca acc gat tta gct tta tgc tct gag gct tta ttg ctt
    10207  aat ttt gct aat tct ttg cct tgc ctg tat gat tta ttg gat gtt  !
! gene II continua
!------------------------------Fim da Tabela   ---------------------------------
```

```
!Tabela 37: seq de ADN w.t. M13 de gene iii
!
!          1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
!         fM   K   K   L   L   F   A   I   P   L   V   V   P   F   Y
   1579   gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat
!      Sequência de sinal ............................
!
!          16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
!          S   H   S   A   E   T   V   E   S   C   L   A   K   P   H
   1624   tct cac tcc gct gaa act gtt gaa agt tgt tta gca aaa ccc cat
! Sequência de sinal> Domínio 1------------------------------
!
!          31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
!          T   E   N   S   F   T   N   V   W   K   D   D   K   T   L
   1669   aca gaa aat tca ttt act aac gtc tgg aaa gac gac aaa act tta
!         Domínio 1----------------------------
!
!          46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
!          D   R   Y   A   N   Y   E   G   C   L   W   N   A   T   G
   1714   gat cgt tac gct aac tat gag ggt tgt ctg tgG AAT GCt aca ggc
!                                              BsmI...
!         Domínio 1-------------------------------------------
!
!          61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
!          V   V   V   C   T   G   D   E   T   Q   C   Y   G   T   W
   1759   gtt gta gtt tgt act ggt gac gaa act cag tgt tac ggt aca tgg
!      Domínio 1-------------------------------------------
!
!          76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
!          V   P   I   G   L   A   I   P   E   N   E   G   G   G   S
   1804   gtt cct acc ggg ctt gct atc cct gaa aat gag ggt ggt ggc tct
!      Domínio 1-------------------------------> Elemento de ligação 1---------
---
!
!          91  92  93  94  95  96  97  98  99  100 101 102 103 104 105
!          E   G   G   G   S   E   G   G   G   S    E   G   G   G   T
   1849   gag ggt ggc ggt tct gag ggt ggc ggt tct gag ggt ggc ggt act
!      Elemento de ligação 1---------------------------------------------->
!
!         106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
!          K   P   P   E   Y   G   D   T   P   I   P   G   Y   T   Y
```

```
   1894   aaa cct cct gag tac ggt gat ac acct att ccg ggc tat act tat
!      Domínio 2-----------------------------------------
!
!          121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
!           I   N   P   L   D   G   T   Y   P   P   G   T   E   Q   N
   1939   atc aac cct ctc gac ggc act taT CCG CCt ggt act gag caa aac
!                      EcilI....
!   Domínio 2-------------------------------------------
!          136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
!           P   A   N   P   N   P   S   L   E   E   S   Q   P   L   N
   1984   ccc gct aat cct aat cct tct ctt GAG GAG tct cag cct ctt aat
!                                      BseRI...
!   Domínio 2-------------------------------------------
!
!          151 152 153 154 155 156 157 158 159 160 161 162 163 164 165
!           T   F   M   F   Q   N   N   R   F   R   N   R   Q   G   A
   2029   act ttc atg ttt cag aat aat agg ttc cga aat agg cag ggg gca
!   Domínio 2-------------------------------------------
!
!          166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
!           L   T   V   Y   T   G   T   V   T   Q   G   T   D   P   V
   2074   tta act gtt tat acg ggc act gtt act caa ggc act gac ccc gtt
!   Domínio 2-------------------------------------------
!
!          181 182 183 184 185 186 187 188 189 190 191 192 193 194 195
!           K   T   Y   Y   Q   Y   T   P   V   S   S   K   A   M   Y
   2119   aaa act tat tac cag tac act cct gta tca tca aaa gcc atg tat
!   Domínio 2-------------------------------------------
!
!          196 197 198 199 200 201 202 203 204 205 206 207 208 209 210
!           D   A   Y   W   N   G   K   F   R   D   C   A   F   H   S
   2164   gac gct tac tgg aac ggt aaa ttC AGa gaC TGc gct ttc cat tct
!                              AlwNI.........
!   Domínio 2-------------------------------------------
!
!          211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
!           G   F   N   E   D   P   F   V   C   E   Y   Q   G   Q   S
   2209   ggc ttt aat gaG GAT CCa ttc gtt tgt gaa tat caa ggc caa tcg
!                    BamHI...
!   Domínio 2-------------------------------------------
!
```

```
!          226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
!           S   D   L   P   Q   P   P   C   N   A   G   G   G   S   G
    2254   tct gac ctg cct ca acct cct gtc aat gct ggc ggc ggc tct ggt
!   Domínio 2----------------------------------> Elemento de ligação 2--------
!
!          241 242 243 244 245 246 247 248 249 250 251 252 253 254 255
!           G   G   S   G   G   G   S   E   G   G   G   S   E   G   G
    2299   ggt ggt tct ggt ggc ggc tct gag ggt ggt ggc tct gag ggt ggc
!   Elemento de ligação 2-------------------------------------------
!
!          256 257 258 259 260 261 262 263 264 265 266 267 268 269 270
!           G   S   E   G   G   G   S   E   G   G   G   S   G   G   G
    2344   ggt tct gag ggt ggc ggc tct gag gga ggc ggt tcc ggt ggt ggc
!   Elemento de ligação 2-------------------------------------------
!
!          271 272 273 274 275 276 277 278 279 280 281 282 283 284 285
!           S   G   S   G   D   F   D   Y   E   K   M   A   N   A   N
    2389   tct ggt tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat
!   Elemento de ligação 2>   Domínio 3---------------------------------------
!
!          286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
!           K   G   A   M   T   E   N   A   D   E   N   A   L   Q   S
    2434   aag ggg gct atg acc gaa aat gcc gat gaa aac gcg cta cag tct
!   Domínio 3-------------------------------------------
!
!          301 302 303 304 305 306 307 308 309 310 311 312 313 314 315
!           D   A   K   G   K   L   D   S   V   A   T   D   Y   G   A
    2479   gac gct aaa ggc aaa ctt gat tct gtc gtc act gat tac ggt gct
!   Domínio 3-------------------------------------------
!
!          316 317 318 319 320 321 322 323 324 325 326 327 328 329 330
!           A   I   D   G   F   I   G   D   V   S   G   L   A   N   G
    2524   gct atc gat ggt ttc att ggt gac gtt tcc ggc ctt gct aat ggt
!   Domínio 3-------------------------------------------
!
!          331 332 333 334 445 336 337 338 39 340 341 342 343 344 345
!           N   G   A   T   G   D   F   A   G   S   N   S   Q   M   A
    2569   aat ggt gct act ggt gat ttt gct ggc tct aat tcc caa atg gct
!   Domínio 3-------------------------------------------
!
!          346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
```

```
!          Q   V   G   D   G   D   N   S   P   L   M   N   N   F   R
   2614   caa gtc ggt gac ggt gat aat tca cct tta atg aat aat ttc ctg
!  Domínio 3---------------------------------------------
!
!         361 362 363 364 365 366 367 368 369 370 371 372 373 374 375
!          Q   Y   L   P   S   L   P   Q   S   V   E   C   R   P   F
   2659   caa tat tta cct tcc ctc cct cca tcg gtt gaa tgt cgc cct ttt
!  Domínio 3---------------------------------------------
!
!         376 377 378 379 380 381 382 383 384 385 386 387 388 389 390
!          V   F   S   A   G   K   P   Y   E   F   S   I   D   C   D
   2704   gtc ttt agc gct ggt aaa cca tat gaa ttt tct att gat tgt gac
!  Domínio 3---------------------------------------------
!
         391 392 393 394 395 396 397 398 399 400 401 402 403 404 405
          K   I   N   L   F   R   G   V   F   A   F   L   L   Y   V
   2749   aaa ata aac tta ttc cgt ggt gtc ttt gtc ttt gcg ttt ctt tta tat gtt
!  Domínio 3---------------------------------------------
!
         406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
!          A   T   F   M   Y   V   F   S   T   F   A   N   I   L   R
   2794   gcc acc ttt atg tat gta ttt tct acg ttt gct aac ata ctg cgt
!  Segmento Transmembranar ------------------------------------> ICA-
!
!         421 422 423 424 425
!          N   K   E   S
   2839   aat aag gag tct taa ! 2853
!  ICA-------------------->     ICA=âncora intracelular
!
!------------ Fim de Tabela ----------------------------------------------
```

Tabela 38: âncora III Madura total M13-III

Derivado de âncora com ADN codificado

```
!
!              1   2   3
!              A   A   A
       1       GCG gcc gca
!              NotI.......
!
!          4   5   6   7   8   9  10  11  12  13  14  15  16  17
!          H   H   H   H   H   H   G   A   A   E   Q   K   L   I
    10     cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc atc
!
!          18  19  20  21  22  23  24  25  26  27  28  29
!          S   E   E   D   L   N   G   A   A   .   A   S
    52     tca gaa gag gat ctg aat ggg gcc gca Tag GCT AGC
!                                                      NheI...
!
!          30  31  32  33  34  35  36  37  38  39
!           D   I   N   D   D   R   M   A   S   T
       88  GAT ATC aac gat gat cgt atg  gct tct act
!   (ON_G37bot)    [RC] 5'-c aac gat gat cgt atg gcG Cat Gct gcc gag aca g-3'
!          EcoRV..
!          local de clivagem de Enterocinase.
!
!   início do maduro III (recodificado)     Domínio 1 --->
!              40  41  42  43
!              A   E   T   V
    118       |gcG|gaG|acA|gtC|
!               t   a   t   t  !   W.T.
!
!       44  45  46  47  48  49  50  51  52  53  54  55  56  57  58
!        E   S   C   L   A   K   P   H   T   E   N   S   F   T   N
130     |gaa|TCC|tgC|CTG|GCC|AaG|ccT|caC|acT|gaG|aat|AGT|ttC|aCA|Aat|
!            agt  t   t   a   a   a   c   t   a   a  tca  t   t   c   ! W.T.
!                    MscI...
!
!       59  60  61  62  63  64  65  66  67  68  69  70  71  72  73
!        V   W   K   D   D   K   T   L   D   R   Y   A   N   Y   E
175     |gtg|TGG|aaG|gaT|gaT|aaG|aaC|CtT|gAT|CGA|TaT|gcC|aaT|taC|gaA|
!       c        a   c   c   a  t t a     t   c   t   c   t   g   !  W.T.
!                          BspDI...
```

```
!
!      74  75  76  77  78  79  80  81  82  83  84  85  86  87  88
!       G   C   L   W   N   A   T   G   V   V   V   C   T   G   D
   220 |ggC|tgC|Tta|tgg|aat|gcC|ACC|GGC|GtC|gtT|gtC|TGC|ACG|ggC|gaT|
!       t   t c g           t   a     t   a   t   t   t   t   c !   W.T.
!                          SgrAI.......        BsgI....
!
!      89  90  91  92  93  94  95  96  97  98  99 100 101 102 103
!       E   T   Q   C   Y   G   T   W   V   P   I   G   L   A   I
265   |gaG|acA|caA|tgC|taT|ggC|ACG|TGg|gtG|ccG|atA|gGC|TTA|GCC|atA|
!       a   t   g   t   c   t   a     t   t   t g c   t   t   c   !   W.T.
!               PmlI....                BlpI....
!
!   Domínio 1-------> Elemento de ligação 1----------------->
!      104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
        P   E   N   E   G   G   G   S   E   G   G   G   S   E   G
310    |cgG|gaG|aaC|gaA|ggC|ggC|ggT|AGC|gaA|ggC|ggT|ggC|AGC|gaA|ggC|
!       t   a   t   g   t   t   c tct g   t   c   t tct   g   t  !W.T.
!   Elemento de ligação 1------------------------------> Domínio 2------------->
!      119 120 121 122 123 124 125 126 127 128 1299 130 131 132 133
!       G   G   S   E   G   G   G   T   K   P   P   E   Y   G   D
355    |ggT|GGA|TCC|gaA|ggA|ggT|ggA|acC|aaG|ccG|ccG|gaA|taT|ggC|gaC|
         c   t   t   g   t   c   t   t   a   t   t   g   c   t   t!W.T.
!      BamHI..(2/2)
!
!      134 235 136 137 138 139 140 141 142 143 144 145 146 147 148
        T   P   I   P   G   Y   T   Y   I   N   P   L   D   G   T
400    acT|ccG|atA|CCT|GGT|taC|acC|taC|atT|aaT|ccG|Tta|gaT|gga|acC|
!       a   t   t   g   t   t   t   c   c   t   c   c   c   c   t!  W.T.
!             SexAI.....
!
!      149 150 151 152 153 154 155 156 157 158 159 160 161 162 163
        Y   P   P   G   T   E   Q   N   P   A   N   P   N   P   S
445   |taC|cct|ccG|ggC|acC|gaA|caG|aaT|ccT|gcC|aaC|ccG|aaC|ccA|AGC|
!      T   G   t   t   t   g   a   c   c   t   t   t   t   t   tct ! W.T.
!                                                    HindIII...
!
!      164 165 166 167 168 169 170 171 172 173 174 175 176 177 178
!       L   E   E   S   Q   P   L   N   T   F   M   F   Q   N   N
490    |TTA|ggA|gaA|AGC|caA|ccG|TtA|aaC|acC|ttT|atg|ttC|caA|aaC|aaC|
        c t   G   G tct   g   t c t   t   t   c       t   g   t   t !W.T.
```

```
!HindIII.
!
!      179 180 181 182 183 184 185 186 187 188 189 190 191 192 193
!       R   F   R   N   R   Q   G   A   L   T   V   Y   T   G   T
535    |Cgt|ttT|AgG|aaC|CgT|ccA|gGT|GCT|CtT|acC|gTG|TAC|AcT|ggA|acC|
!      a g   c c a   t a g   g   g   a t a   t   t   t   g   c   t! W.T.
!                        HgiAI...               BsrGI...
!
!      194 195 196 197 198 199 200 201 202 203 204 205 206 207 208
!       V   T   Q   G   T   D   P   V   K   T   Y   Y   Q   Y   T
580    |gtC|acC|caG|GGT|ACC|gaT|ccT|gtC|aaG|acC|taC|taT|caA|taT|acC|
!       t   t   a   c   t   c   c   t   a   t   t   c   g   c   t  !W.T.
!             KpnI...
!      209 210 211 212 213 214 215 216 217 218 219 220 221 222 223
!       P   V   S   S   K   A   M   Y   D   A   Y   W   N   G   K
625    |ccG|gtC|TCG|AGt|aaG|gcT|atg|taC|gaT|gcC|taT|tgg|aaT|ggC|aag|
!       t   a   a tca   a   c       t   c   t   c       c   t   a  !W.T.
!      BsaI...
!                 XhoI....
!
!      224 225 226 227 228 229 230 231 232 233 234 235 236 237 238
!       F   R   D   C   A   F   H   S   G   F   N   E   D   P   F
670    |ttT|CgT|gaT|tgT|gcC|ttT|caC|AGC|ggT|ttC|aaC|gaa|gac|CCt|ttT|
!      C   A a   C   c   t   c   t tct   c   t   t   G   T   a   c  !W.T.
!
!      239 240 241 242 243 244 245 246 247 248 249 250 251 252 253
!       V   C   E   Y   Q   G   Q   S   S   D   L   P   Q   P   P
715    |gtC|tgC|gaG|taC|caG|ggT|caG|AGT|AGC|gaT|TtA|ccG|caG|ccA|CCG|
!       t   t   a   t   a   c   a tcg tct   c c g   t   a   t   t  !W.T.
!   DrdI....                                              AgeI....
!
!   Domínio 2 -------->  Elemento de ligação 2---------------------------->
!      254 255 256 257 258 259 260 261 262 263 264 265 266 267 268
!       V   N   A   G   G   G   S   G   G   G   S   G   G   G   S
760    |GTT|AAC|gcG|ggT|ggT|ggT|AGC|ggC|ggA|ggC|AGC|ggC|ggT|ggT|AGC|
!       c   t   t   c   c   c tct   t   t   t tct   t   c   c tct  !W.T.
!AgeI....
!          HpaI....
!          HincII.
!
```

```
!      Elemento de ligação 2------------------------------------------>
Domínio 3 -------->
!      269 270 271 272 273 274 275 276 277 278 279 280 281 282 283
!       E   G   G   G   S   E   G   G   G   S   G   G   G   S   G
805   |ggA|ggC|ggA|ggT|AGC|gaA|ggA|ggT|ggc|AGC|ggA|ggC|ggT|AGC|ggC|
!      g   t   t   c   tct   g   t   c   t tct   g   t   c tct   t  !W.T.
!
!      ----------------- Domínio 3 ---------------------------------->
!      284 285 286 287 288 289 290 291 292 293 294 295 296 297 298
!       S   G   D   F   D   Y   E   K   M   A   N   A   N   K   G
850   |AGT|ggC|gac|ttc|gac|tac|gag|aaa|atg|gct|aat|gcc|aac|aaa|GGC|
!      tcc  t   t   t   t   t   a   g       a   c   t   t   g   g   !  W.T.
!                                                      KasI....
!
!      299 300 301 302 303 304 305 306 307 308 309 310 311 312 313
!       A   M   T   E   N   A   D   E   N   A   L   Q   S   D   A
895   |GCC|atg|act|gag|aac|gct|gac|gaG|AAT|GCA|ctg|caa|agt|gat|gCC|
!      t       c   a   t   c   t   a   c   g   a   g  tct   c   t   !W.T.
!KasI....              BsmI....             StyI...
!
!      314 315 316 317 318 319 320 321 322 323 324 325 326 327 328
!       K   G   K   L   D   S   V   A   T   D   Y   G   A   A   I
940   |AAG|GGt|aag|tta|gac|agc|gTC|GCc|Aca|gac|tat|ggT|GCt|gcc|atc|
!      a   c   a  c t   t tct     t   t   t   c           t   !W.T.
!StyI.....          PflFI....
!
!      329 330 331 332 333 334 335 336 337 338 339 340 341 342 343
!       D   G   F   I   G   D   V   S   G   L   A   N   G   N   G
985   |gac|ggc|ttt|atc|ggc|gat|gtc|agt|ggt|ctg|gct|aac|ggc|aac|gga|
!       t   t   c   t   t   c   t tcc   c c t     t   t   t   t   !W.T.
!
!      344 345 346 347 348 349 350 351 352 353
!       A   T   G   D   F   A G   S   N   S
1030  |gcc|acc|gga|gac|ttc|GCA|GGT|tcG|AAT|TCt|
!        t   t   t   t   t   t   c   t     c!W.T.
!                                BstBI...
!                                   EcoRI...
!                      bspMI..
!
!      354 355 356 357 358 359 360 361 362 363
!       Q   M   A   Q   V   G   D   G   D   N
```

```
1060   cag atg gcC CAG GTT GGA GAT GGg gac aac
!      a       t   a   c   t   c   t   t   t!W.T.
!             XcmI..............
!
!      364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379
!       S   P   L   M   N   N   F   R   Q   Y   L   P   S   L   P   Q
1090   agt ccg ctt atg aac aac ttt aga cag tac ctt ccg tct ctt ccg cag
!     tca    t t a     t   t   c  c t   a   t t a   t   c   c   t   a!W.T.
!
!      380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
!       S   V   E   C   R   P   F   V   F   S   A   G   K   P   Y   E
1138   agt gtc gag tgc cgt cca ttc gtt ttc tct gcc ggc aag cct tac gag
!     tcg   t   a   t   c   t   t   c   t  agc   t   t   a   a   t   a  ! W.T.
!
!      Domínio 3 ---------------------------------------------------->
!      396 397 398 399 400 401 402 403 404 405 406 407
!       F   S   I   D   C   D   K   I   N   L   F   R

1186   ttc aGC Atc gac TGC gat aag atc aat ctt ttC CGC
!      t   tct   t   t   t   c   a   a   c t a     t
!      BstAPI.......                     SacII...
!
!      segmento transmembranar-------------->
!      408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423
!       G   V   F   A   F   L   L   Y   V   A   T   F   M   Y   V   F
1222   GGc gtt ttc gct ttc ttg cta tac gtc gct act ttc atg tac gtt ttc
!       t   c   t   g   t c t t a   t   t   c   c   t     t   a   t!W.T.
!
!      424 425 426 427 428 429 430  431 432 433 434 435
!       S   T   F   A   N   I   L   R   N   K   E   S
1270   aGC ACT TTC GCC AAT ATT TTA  Cgc aac aaa gaa agc
!      tct   g   t   t   c   a c g   t   t   g  g  tct! W.T.
!                                 Âncora intracelular.
!
!                .   .
   1306          tag tga tct CCT AGG
!                AvrII..
!
   1321   aag ccc gcc taa tga gcg ggc ttt ttt ttt ct ggt
!         |Trp terminal                        |
!
```

```
!Fim de invólucro Fab
!----------------------------------<Fim de Tabela------------------------------


Tabela 39: On para fazer eliminações em III
!On para utilização com NheI
!
N
(ON_G29bot)                              5'-c gTT gAT ATc gcT Agc cTA Tgc-3' !
22
!   este é o complemento reverso de 5'-gca tag gct agc gat atc aac g-3'
!                                  NheI... scab.........
(On_G104top)   5'-g|ata|ggc|tta|gcT|aGC|ccg|gag|aac|gaa|gg-3'     !
30
!              Scab.............NheI...  104 105 106 107 107
(ON_G236top)   5'-c|ttt|cac|agc|ggt|ttc|GCT|AGC|gag|cct|ttt|gtc|tgc-3'   !
37
!                                             NheI...236 237 238 239 240
(ON_G236tCS)   5'-c|ttt|cac|agc|ggt|ttc|GCT|AGC|gac|cct|ttt|gtc|Agc-
!                                             NheI...236 237 238 239 240
               gag|tac|cag|ggt|c-3'                                          !
50

!ON para utilização com SphI G CAT Gc
(ON_X37bot)       5'-gAc TgT cTc ggc Agc ATg cgc CAT Acg ATc ATc gTT g-3'   !
37
!              N   D   D   R   M   A   H   A
! (ON_X37bot) =[RC]   5'-c aac gat gat cgt atg gcG Cat Gct gcc gag aca gtc-3'
!                                                SphI...Scab..............
(ON_X104top)   5'-g(ON_X37bot)gtG ccg|ata|ggc|ttG|CAT|GCa|ccg|gag|aac|gaa|gg-3'!
36
!          Scab............................SphI....    104 105 106 107 108
(ON_X236top)   5'-c|ttt|cac|agc|ggt|ttG|CaT|gCa|gac|cct|ttt|gtc|tgc-3'      !
37
!                                  SphI....   236 237 238 239 240
(ON_X236tCS)   5'-c|ttt|cac|agc|ggt|ttG|CaT|gCa|gac|cct|ttt|gtc|Agc-
!                                     NheI... 236    237 238 239 240
                     gag|tac|cag|ggt|c-3'
50
```

Tabela 40: Titulantes de fagos e concentração de uma selecção com uma biblioteca de Fab humanas à base de DY3F31

	Entrada (cfu total)	Saída (cfu total)	Razão entrada/saída
R1-ox seleccionado em phOx-BSA	$4,5 \times 10^{12}$	$3,4 \times 10^{5}$	$7,5 \times 10-^{8}$
R2-estrep seleccionado em esferas de estrep	$9,2 \times 10^{12}$	3×10^{8}	$3,3 \times 10^{-5}$

Tabela 41: Frequência de ELISA positivos em bibliotecas de Fab humanas à base de DY3F31

	Anti-M13 HRP	9E10/RAMHRP	Anti-CK/CL Gar-HRP
R2-ox (com indução de IPTG) R2-ox (sem IPTG)	18/44 13/4	10/44 ND	10/44 ND
R3-estrep (com IPTG)	39/44	38/44	36/44
R3-estrep (sem IPTG)	33/44	ND	ND

LISTAGEM DE SEQUÊNCIAS

<110> LADNER, ROBERT C.
COHEN, EDWARD H.
NASTRI, HORACIO G.
ROOKEY, KRISTIN L.
HOET, RENE
HOOGENBOOM, HENDRICUS R. J. M.

<120> NOVOS MÉTODOS DE BIBLIOTECAS DE CONSTRUÇÃO COMPREENDENDO ELEMENTOS APRESENTADOS E/OU EXPRESSOS DE UMA FAMÍLIA DIVERSA DE PÉPTIDOS, POLIPÉPTIDOS OU PROTEÍNAS E AS NOVAS BIBLIOTECAS

<130> DYAX/002 CIP2

<140> 10/045,674
<141> 2001-10-25

<150> 06/198,069
<151> 2000-04-17

<150> 09/837,306
<151> 2001-04-17

<160> 635

<170> PatentIn Ver. 2.1

<210> 1
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 1
catgtgtatt actgtgc 17

<210> 2
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 2
cacatccgtg cttcttgcac ggatgtggca cagtaataca catg 44

<210> 3
<211> 18
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 3
gtgtattaga ctgctgcc 18

<210> 4
<211> 43
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 4
ggcagcagtc taatacacca catccgtgtt cttcacggat 43

<210> 5
<211> 47
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 5
cacatccgtg tttgttacac ggatgtggtg tcttacagtc cattctg 47

<210> 6
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 6
cagaatggac tgtaagacac 20

<210> 7
<211> 43
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 7
atcgagtctc actgagccac atccgtggtt ttccacggat gtg 43

<210> 8
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 8
gctcagtgag actcgat 17

<210> 9
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (10)..(24)
<223> A, T, C, G, outro ou desconhecido

<400> 9
cacgaggagn nnnnnnnnnn nnnn 24

<210> 10
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 10
atgaccgaat tgctacaag 19

<210> 11
<211> 46
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 11
gactcctcag cttcttgctg aggagtcctt gtagcaattc ggtcat 46

<210> 12
<211> 6
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: 6 His tag

<400> 12
His His His His His His
1 5

<210> 13
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (6)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 13
gtctcnnnnn 10

<210> 14
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 14
nnnnnngaga c 11

<210> 15
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (11)..(24)
<223> A, T, C, G, outro ou desconhecido

<400> 15
cacggatgtg nnnnnnnnnn nnnn 24

<210> 16
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(14)
<223> A, T, C, G, outro ou desconhecido

<400> 16
nnnnnnnnnn nnnncacatc cgtg 24

<210> 17
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 17
gtgtattact gtgc 14

<210> 18
<211> 34
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 18
cacatccgtg cacggatgtg gcacagtaat acac 34

<210> 19
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 19
gtgtattaga ctgc 14

<210> 20
<211> 34
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 20
gcagtctaat acaccacatc cgtgcacgga tgtg 34

<210> 21
<211> 34
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 21
cacatccgtg cacggatgtg gtgtcttaca gtcc 34

<210> 22
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 22
ggactgtaag acac 14

<210> 23
<211> 34
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 23
gagtctcact gagccacatc cgtgcacgga tgtg 34

<210> 24
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 24
gctcagtgag actc 14

<210> 25
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 25
gtgtattact gtgc 14

<210> 26
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 26
gtatattact gtgc 14

<210> 27
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 27
gtgtattact gtaa 14

<210> 28
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 28
gtgtattact gtac 14

<210> 29
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 29
ttgtattact gtgc 14

<210> 30
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 30
ttgtatcact gtgc 14

<210> 31
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 31
acatattact gtgc 14

<210> 32
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 32
acgtattact gtgc 14

<210> 33
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 33
atgtattact gtgc 14

<210> 34
<211> 101
<212> ADN
<213> Homo sapiens

<400> 34

agggtcacca tgaccaggga cacgtccatc agcacagcct acatgabcga gctgagcagg 60
ctgagatctg acgacacggc cgtgtattac tgtgcgagag a 101

<210> 35
<211> 98
<212> ADN
<213> Homo sapiens

<400> 35

agagtcacca ttaccaggga cacatccgcg agcacagcct acatggagct gagcagcctg 60
agatctgaag acacggctgt gtattactgt gcgagaga 98

<210> 36
<211> 98
<212> ADN
<213> Homo sapiens

<400> 36

agagtcacca tgaccaggaa cacctccata agcacagcct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcgagagg 98

<210> 37
<211> 98
<212> ADN
<213> Homo sapiens

<400> 37

agagtcacca tgaccacaga cacatccacg agcacagcct acatggagct gaggagcctg 60
agatctgacg acacggccgt gtattactgt gcgagaga 98

<210> 38
<211> 98
<212> ADN
<213> Homo sapiens

<400> 38

agagtcacca tgaccgagga cacatctaca gacacagcct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcaacaga 98

<210> 39
<211> 98
<212> ADN
<213> Homo sapiens

<400> 39

agagtcacca ttaccaggga caggtctatg agcacagcct acatggagct gagcagcctg 60
agatctgagg acacagccat gtattactgt gcaagata 98

<210> 40
<211> 98
<212> ADN
<213> Homo sapiens

<400> 40

agagtcacca tgaccaggga cacgtccacg agcacagtct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcgagaga 98

<210> 41
<211> 98
<212> ADN
<213> Homo sapiens

<400> 41

agagtcacca ttaccaggga catgtccaca agcacagcct acatggagct gagcagcctg 60
agatccgagg acacggccgt gtattactgt gcggcaga 98

<210> 42
<211> 98
<212> ADN
<213> Homo sapiens

<400> 42

agagtcacga ttaccgcgga cgaatccacg agcacagcct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcgagaga 98

<210> 43
<211> 98
<212> ADN
<213> Homo sapiens

<400> 43

agagtcacga ttaccgcgga caaatccacg agcacagcct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcgagaga 98

<210> 44
<211> 98
<212> ADN
<213> Homo sapiens

<400> 44

agagtcacca taaccgcgga cacgtctaca gacacagcct acatggagct gagcagcctg 60
agatctgagg acacggccgt gtattactgt gcaacaga 98

<210> 45
<211> 100
<212> ADN
<213> Homo sapiens

<400> 45

aggctcacca tcaccaagga cacctccaaa aaccaggtgg tccttacaat gaccaacatg 60
gaccctgtgg acacagccac atattactgt gcacacagac 100

<210> 46
<211> 100
<212> ADN
<213> Homo sapiens

<400> 46

aggctcacca tctccaagga cacctccaaa agccaggtgg tccttaccat gaccaacatg 60
gaccctgtgg acacagccac atattactgt gcacggatac 100

<210> 47
<211> 100
<212> ADN
<213> Homo sapiens

<400> 47

aggctcacca tctccaagga cacctccaaa aaccaggtgg tccttacaat gaccaacatg 60
gaccctgtgg acacagccac gtattactgt gcacggatac 100

<210> 48
<211> 98
<212> ADN
<213> Homo sapiens

<400> 48

cgattcacca tctccagaga caacgccaag aactcactgt atctgcaaat gaacagcctg 60
agagccgagg acacggctgt gtattactgt gcgagaga 98

<210> 49
<211> 100
<212> ADN
<213> Homo sapiens

<400> 49

cgattcacca tctccagaga caacgccaag aactccctgt atctgcaaat gaacagtctg 60
agagctgagg acacggcctt gtattactgt gcaaaagata 100

<210> 50
<211> 98
<212> ADN
<213> Homo sapiens

<400> 50

cgattcacca tctccaggga caacgccaag aactcactgt atctgcaaat gaacagcctg 60
agagccgagg acacggccgt gtattactgt gcgagaga 98

<210> 51
<211> 98
<212> ADN
<213> Homo sapiens

<400> 51

cgattcacca tctccagaga aaatgccaag aactccttgt atcttcaaat gaacagcctg 60
agagccgggg acacggctgt gtattactgt gcaagaga 98

<210> 52
<211> 98
<212> ADN
<213> Homo sapiens

<400> 52

agattcacca tctcaagaga tgattcaaaa aacacgctgt atctgcaaat gaacagcctg 60
aaaaccgagg acacagccgt gtattactgt accacaga 98

<210> 53
<211> 98
<212> ADN
<213> Homo sapiens

<400> 53

cgattcacca tctccagaga caacgccaag aactccctgt atctgcaaat gaacagtctg 60
agagccgagg acacggcctt gtatcactgt gcgagaga 98

<210> 54
<211> 98
<212> ADN
<213> Homo sapiens

<400> 54

cgattcacca tctccagaga caacgccaag aactcactgt atctgcaaat gaacagcctg 60
agagccgagg acacggctgt gtattactgt gcgagaga 98

<210> 55
<211> 98
<212> ADN
<213> Homo sapiens

<400> 55

cggttcacca tctccagaga caattccaag aacacgctgt atctgcaaat gaacagcctg 60
agagccgagg acacggccgt atattactgt gcgaaaga 98

<210> 56
<211> 98
<212> ADN
<213> Homo sapiens

<400> 56

cgattcacca tctccagaga caattccaag aacacgctgt atctgcaaat gaacagcctg 60
agagctgagg acacggctgt gtattactgt gcgaaaga 98

<210> 57
<211> 98
<212> ADN
<213> Homo sapiens

<400> 57

cgattcacca tctccagaga caattccaag aacacgctgt atctgcaaat gaacagcctg 60
agagctgagg acacggctgt gtattactgt gcgagaga 98

<210> 58
<211> 98
<212> ADN
<213> Homo sapiens

<400> 58

cgattcacca tctccagaga caattccaag aacacgctgt atctgcaaat gaacagcctg 60
agagctgagg acacggctgt gtattactgt gcgaaaga 98

<210> 59
<211> 98
<212> ADN
<213> Homo sapiens

<400> 59

cgattcacca tctccagaga caattccaag aacacgctgt atctgcaaat gaacagcctg 60
agagccgagg acacggctgt gtattactgt gcgagaga 98

<210> 60
<211> 100
<212> ADN
<213> Homo sapiens

<400> 60

cgattcacca tctccagaga caacagcaaa aactccctgt atctgcaaat gaacagtctg 60
agaactgagg acaccgcctt gtattactgt gcaaaagata 100

<210> 61
<211> 98
<212> ADN
<213> Homo sapiens

<400> 61

cgattcacca tctccagaga caatgccaag aactcactgt atctgcaaat gaacagcctg 60
agagacgagg acacggctgt gtattactgt gcgagaga 98

<210> 62
<211> 98
<212> ADN
<213> Homo sapiens

<400> 62

agattcacca tctcaagaga tggttccaaa agcatcgcct atctgcaaat gaacagcctg 60
aaaaccgagg acacagccgt gtattactgt actagaga 98

<210> 63
<211> 98
<212> ADN
<213> Homo sapiens

<400> 63

cgattcacca tctccagaga caattccaag aacacgctgt atcttcaaat gaacagcctg 60
agagccgagg acacggccgt gtattactgt gcgagaga 98

<210> 64
<211> 98
<212> ADN
<213> Homo sapiens

<400> 64

agattcacca tctccagaga caattccaag aacacgctgt atcttcaaat gggcagcctg 60
agagctgagg acatggctgt gtattactgt gcgagaga 98

<210> 65
<211> 98
<212> ADN
<213> Homo sapiens

<400> 65

agattcacca tctccagaga caattccaag aacacgctgt atcttcaaat gaacagcctg 60
agagctgagg acacggctgt gtattactgt gcgagaga 98

<210> 66
<211> 98
<212> ADN
<213> Homo sapiens

<400> 66

agattcacca tctcaagaga tgattcaaag aactcactgt atctgcaaat gaacagcctg 60
aaaaccgagg acacggccgt gtattactgt gctagaga 98

<210> 67
<211> 98
<212> ADN
<213> Homo sapiens

<400> 67

aggttcacca tctccagaga tgattcaaag aacacggcgt atctgcaaat gaacagcctg 60
aaaaccgagg acacggccgt gtattactgt actagaca 98

<210> 68
<211> 98
<212> ADN
<213> Homo sapiens

<400> 68

cgattcacca tctccagaga caacgccaag aacacgctgt atctgcaaat gaacagtctg 60
agagccgagg acacggctgt gtattactgt gcaagaga 98

<210> 69
<211> 98
<212> ADN
<213> Homo sapiens

<400> 69

agattcacca tctccagaga caattccaag aacacgctgc atcttcaaat gaacagcctg 60
agagctgagg acacggctgt gtattactgt aagaaaga 98

<210> 70
<211> 98
<212> ADN
<213> Homo sapiens

<400> 70

cgagtcacca tatcagtaga caagtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgcgg acacggccgt gtattactgt gcgagaga 98

<210> 71
<211> 98
<212> ADN
<213> Homo sapiens

<400> 71

cgagtcacca tgtcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgtgg acacggccgt gtattactgt gcgagaaa 98

<210> 72
<211> 98
<212> ADN
<213> Homo sapiens

<400> 72

cgagttacca tatcagtaga cacgtctaag aaccagttct ccctgaagct gagctctgtg 60
actgccgcgg acacggccgt gtattactgt gcgagaga 98

<210> 73
<211> 98
<212> ADN
<213> Homo sapiens

<400> 73

cgagtcacca tatcagtaga caggtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgcgg acacggccgt gtattactgt gccagaga 98

<210> 74
<211> 98
<212> ADN
<213> Homo sapiens

<400> 74

cgagttacca tatcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
actgccgcag acacggccgt gtattactgt gccagaga 98

<210> 75
<211> 98
<212> ADN
<213> Homo sapiens

<400> 75

cgagttacca tatcagtaga cacgtctaag aaccagttct ccctgaagct gagctctgtg 60
actgccgcgg acacggccgt gtattactgt gcgagaga 98

<210> 76
<211> 98
<212> ADN
<213> Homo sapiens

<400> 76

cgagtcacca tatcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgcgg acacggctgt gtattactgt gcgagaga 98

<210> 77
<211> 98
<212> ADN
<213> Homo sapiens

<400> 77

cgagtcacca tatccgtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgcag acacggctgt gtattactgt gcgagaca 98

<210> 78
<211> 98
<212> ADN
<213> Homo sapiens

<400> 78

cgagtcacca tatcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgctgcgg acacggccgt gtattactgt gcgagaga 98

<210> 79
<211> 98
<212> ADN
<213> Homo sapiens

<400> 79

cgagtcacca tatcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgctgcgg acacggccgt gtattactgt gcgagaga 98

<210> 80
<211> 98
<212> ADN
<213> Homo sapiens

<400> 80

cgagtcacca tatcagtaga cacgtccaag aaccagttct ccctgaagct gagctctgtg 60
accgccgcag acacggccgt gtattactgt gcgagaga 98

<210> 81
<211> 98
<212> ADN
<213> Homo sapiens

<400> 81

caggtcacca tctcagccga caagtccatc agcaccgcct acctgcagtg gagcagcctg 60
aaggcctcgg acaccgccat gtattactgt gcgagaca 98

<210> 82
<211> 96
<212> ADN
<213> Homo sapiens

<400> 82

cacgtcacca tctcagctga caagtccatc agcactgcct acctgcagtg gagcagcctg 60
aaggcctcgg acaccgccat gtattactgt gcgaga 96

<210> 83
<211> 98
<212> ADN
<213> Homo sapiens

<400> 83

cgaataacca tcaacccaga cacatccaag aaccagttct ccctgcagct gaactctgtg 60
actcccgagg acacggctgt gtattactgt gcaagaga 98

<210> 84
<211> 98
<212> ADN
<213> Homo sapiens

<400> 84

cggtttgtct tctccttgga cacctctgtc agcacggcat atctgcagat ctgcagccta 60
aaggctgagg acactgccgt gtattactgt gcgagaga 98

<210> 85
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (3)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 85
gcnnnnnnng c 11

<210> 86
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 86
caynnnnrtg 10

<210> 87
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (6)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 87
gagtcnnnnn n 11

<210> 88
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 88
nnnnnngaga c 11

<210> 89
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 89
gaannnnttc 10

<210> 90
<211> 90
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência Nucleotídica 3-23 FR3 Sintética

<220>
<221> CDS
<222> (1)..(90)

<220>
<221> base_modificada
<222> (3)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (9)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (12)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (21)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (30)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (36)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (51)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (57)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (60)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (69)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (72)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (75)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (78)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (87)
<223> A, T, C or G

<400> 90

acn ath wsn mgn gay aay wsn aar aay acn ytn tay ttn car atg aay 48
Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn
1 5 10 15

wsn ttr mgn gcn gar gay acn gcn gtn tay tay tgy gcn aar 90
Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys
20 25 30

<210> 91
<211> 30
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência Proteica 3-23 FR3 Sintética

<400> 91

Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn
1 5 10 15

Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys
20 25 30

<210> 92
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 92
agttctccct gcagctgaac tc 22

<210> 93
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 93
cactgtatct gcaaatgaac ag 22

<210> 94
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 94
ccctgtatct gcaaatgaac ag 22

<210> 95
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 95
ccgcctacct gcagtggagc ag 22

<210> 96
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 96
cgctgtatct gcaaatgaac ag 22

<210> 97
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 97
cggcatatct gcagatctgc ag 22

<210> 98
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 98
cggcgtatct gcaaatgaac ag 22

<210> 99
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 99
ctgcctacct gcagtggagc ag 22

<210> 100
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 100
tcgcctatct gcaaatgaac ag 22

<210> 101
<211> 63
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 101

cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agg 63

<210> 102
<211> 45
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 102
caagtagaga gtattcttag agttgtctct agacttagtg aagcg 45

<210> 103
<211> 54
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 103
cgcttcacta agtctagaga caactctaag aatactctct acttgcagct gaac 54

<210> 104
<211> 54
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 104
cgcttcacta agtctagaga caactctaag aatactctct acttgcaaat gaac 54

<210> 105
<211> 54
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 105
cgcttcacta agtctagaga caactctaag aatactctct acttgcagtg gagc 54

<210> 106
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 106
cgcttcacta agtctagaga c 21

<210> 107
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sintética

<400> 107
acatggagct gagcagcctg ag 22

<210> 108
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 108
acatggagct gagcaggctg ag 22

<210> 109
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 109
acatggagct gaggagcctg ag 22

<210> 110
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 110
acctgcagtg gagcagcctg aa 22

<210> 111
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 111
atctgcaaat gaacagcctg aa 22

<210> 112
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 112
atctgcaaat gaacagcctg ag 22

<210> 113
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 113
atctgcaaat gaacagtctg ag 22

<210> 114
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 114
atctgcagat ctgcagccta aa 22

<210> 115
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 115
atcttcaaat gaacagcctg ag 22

<210> 116
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 116
atcttcaaat gggcagcctg ag 22

<210> 117
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 117
ccctgaagct gagctctgtg ac 22

<210> 118
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 118
ccctgcagct gaactctgtg ac 22

<210> 119
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 119
tccttacaat gaccaacatg ga 22

<210> 120
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 120
tccttaccat gaccaacatg ga 22

<210> 121
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 121
acatggagct gagcagcctg ag 22

<210> 122
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 122
ccctgaagct gagctctgtg ac 22

<210> 123
<211> 54
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 123
cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaac 54

<210> 124
<211> 60
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 124
cgcttcactc agtctagaga taacagtaaa aatactttgt acttgcagct gagcagcctg 60

<210> 125
<211> 60
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 125
cgcttcactc agtctagaga taacagtaaa aatactttgt acttgcagct gagctctgtg 60

<210> 126
<211> 52
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 126
tcagctgcaa gtacaaagta tttttactgt tatctctaga ctgagtgaag cg 52

<210> 127
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 127
cgcttcactc agtctagaga taac 24

<210> 128
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 128
ccgtgtatta ctgtgcgaga ga 22

<210> 129
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 129
ctgtgtatta ctgtgcgaga ga 22

<210> 130
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 130
ccgtgtatta ctgtgcgaga gg 22

<210> 131
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 131
ccgtgtatta ctgtgcaaca ga 22

<210> 132
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 132
ccatgtatta ctgtgcaaga ta 22

<210> 133
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 133
ccgtgtatta ctgtgcggca ga 22

<210> 134
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 134
ccacatatta ctgtgcacac ag 22

<210> 135
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 135
ccacatatta ctgtgcacgg at 22

<210> 136
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 136
ccacgtatta ctgtgcacgg at 22

<210> 137
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 137
ccttgtatta ctgtgcaaaa ga 22

<210> 138
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 138
ctgtgtatta ctgtgcaaga ga 22

<210> 139
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 139
ccgtgtatta ctgtaccaca ga 22

<210> 140
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 140
ccttgtatca ctgtgcgaga ga 22

<210> 141
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 141
ccgtatatta ctgtgcgaaa ga 22

<210> 142
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 142
ctgtgtatta ctgtgcgaaa ga 22

<210> 143
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 143
ccgtgtatta ctgtactaga ga 22

<210> 144
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 144
ccgtgtatta ctgtgctaga ga 22

<210> 145
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 145
ccgtgtatta ctgtactaga ca 22

<210> 146
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 146
ctgtgtatta ctgtaagaaa ga 22

<210> 147
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 147
ccgtgtatta ctgtgcgaga aa 22

<210> 148
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 148
ccgtgtatta ctgtgccaga ga 22

<210> 149
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 149
ctgtgtatta ctgtgcgaga ca 22

<210> 150
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 150
ccatgtatta ctgtgcgaga ca 22

<210> 151
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 151
ccatgtatta ctgtgcgaga 20

<210> 152
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 152
ccgtgtatta ctgtgcgaga 21

<210> 153
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 153
ctgtgtatta ctgtgcgaga g 21

<210> 154
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 154
ccgtgtatta ctgtgcgaga g 21

<210> 155
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 155
ccgtatatta ctgtgcgaaa g 21

<210> 156
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 156
ctgtgtatta ctgtgcgaaa g 21

<210> 157
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 157
ctgtgtatta ctgtgcgaga c 21

<210> 158
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 158
ccatgtatta ctgtgcgaga c 21

<210> 159
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 159
ccatgtatta ctgtgcgaga 20

<210> 160
<211> 94
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 160

ggtgtagtga tctagtgaca actctaagaa tactctctac ttgcagatga acagctttag 60
ggctgaggac actgcagtct actattgtgc gaga 94

<210> 161
<211> 94
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 161

ggtgtagtga tctagtgaca actctaagaa tactctctac ttgcagatga acagctttag 60
ggctgaggac actgcagtct actattgtgc gaaa 94

<210> 162
<211> 85
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 162

atagtagact gcagtgtcct cagcccttaa gctgttcatc tgcaagtaga gagtattctt 60
agagttgtct ctagatcact acacc 85

<210> 163
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 163
ggtgtagtga tctagagaca ac 22

<210> 164
<211> 55
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 164
ggtgtagtga aacagcttta gggctgagga cactgcagtc tactattgtg cgaga 55

<210> 165
<211> 55
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 165
ggtgtagtga aacagcttta gggctgagga cactgcagtc tactattgtg cgaaa 55

<210> 166
<211> 46
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 166
atagtagact gcagtgtcct cagcccttaa gctgtttcac tacacc 46

<210> 167
<211> 46
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 167
ggtgtagtga aacagcttaa gggctgagga cactgcagtc tactat 46

<210> 168
<211> 26
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 168
ggtgtagtga aacagcttaa gggctg 26

<210> 169

<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 169
agttctccct gcagctgaac tc 22

<210> 170
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 170
cactgtatct gcaaatgaac ag 22

<210> 171
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 171
ccctgtatct gcaaatgaac ag 22

<210> 172
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 172
ccgcctacct gcagtggagc ag 22

<210> 173
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 173
cgctgtatct gcaaatgaac ag 22

<210> 174
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 174
cggcatatct gcagatctgc ag 22

<210> 175
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 175
cggcgtatct gcaaatgaac ag 22

<210> 176
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 176
ctgcctacct gcagtggagc ag 22

<210> 177
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 177
tcgcctatct gcaaatgaac ag 22

<210> 178
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 178
acatggagct gagcagcctg ag 22

<210> 179
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 179
acatggagct gagcaggctg ag 22

<210> 180
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 180
acatggagct gaggagcctg ag 22

<210> 181
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 181
acctgcagtg gagcagcctg aa 22

<210> 182
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 182
atctgcaaat gaacagcctg aa 22

<210> 183
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 183
atctgcaaat gaacagcctg ag 22

<210> 184
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 184
atctgcaaat gaacagtctg ag 22

<210> 185
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 185
atctgcagat ctgcagccta aa 22

<210> 186
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 186
atcttcaaat gaacagcctg ag 22

<210> 187
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 187
atcttcaaat gggcagcctg ag 22

<210> 188
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 188
ccctgaagct gagctctgtg ac 22

<210> 189
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 189
ccctgcagct gaactctgtg ac 22

<210> 190
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 190
tccttacaat gaccaacatg ga 22

<210> 191
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 191
tccttaccat gaccaacatg ga 22

<210> 192
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 192
ccgtgtatta ctgtgcgaga ga 22

<210> 193
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 193
ctgtgtatta ctgtgcgaga ga 22

<210> 194
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 194
ccgtgtatta ctgtgcgaga gg 22

<210> 195
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 195
ccgtgtatta ctgtgcaaca ga 22

<210> 196
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 196
ccatgtatta ctgtgcaaga ta 22

<210> 197
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 197
ccgtgtatta ctgtgcggca ga 22

<210> 198
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 198
ccacatatta ctgtgcacac ag 22

<210> 199
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 199
ccacatatta ctgtgcacgg at 22

<210> 200
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 200
ccacgtatta ctgtgcacgg at 22

<210> 201
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 201
ccttgtatta ctgtgcaaaa ga 22

<210> 202
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 202
ctgtgtatta ctgtgcaaga ga 22

<210> 203
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 203
ccgtgtatta ctgtaccaca ga 22

<210> 204
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 204
ccttgtatca ctgtgcgaga ga 22

<210> 205
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 205
ccgtatatta ctgtgcgaaa ga 22

<210> 206
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 206
ctgtgtatta ctgtgcgaaa ga 22

<210> 207
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 207
ccgtgtatta ctgtactaga ga 22

<210> 208
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 208
ccgtgtatta ctgtgctaga ga 22

<210> 209
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 209
ccgtgtatta ctgtactaga ca 22

<210> 210
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 210
ctgtgtatta ctgtaagaaa ga 22

<210> 211
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 211
ccgtgtatta ctgtgcgaga aa 22

<210> 212
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 212
ccgtgtatta ctgtgccaga ga 22

<210> 213
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 213
ctgtgtatta ctgtgcgaga ca 22

<210> 214
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 214
ccatgtatta ctgtgcgaga ca 22

<210> 215
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 215
ccatgtatta ctgtgcgaga aa 22

<210> 216
<211> 90
<212> ADN
<213> Homo sapiens

<400> 216

caggtgcagc tggtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcctgcaagg cttctggata caccttcacc 90

<210> 217
<211> 90
<212> ADN
<213> Homo sapiens

<400> 217

caggtccagc ttgtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtt 60
tcctgcaagg cttctggata caccttcact 90

<210> 218
<211> 90
<212> ADN
<213> Homo sapiens

<400> 218

caggtgcagc tggtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcctgcaagg cttctggata caccttcacc 90

<210> 219
<211> 90
<212> ADN
<213> Homo sapiens

<400> 219

caggttcagc tggtgcagtc tggagctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcctgcaagg cttctggtta cacctttacc 90

<210> 220
<211> 90
<212> ADN
<213> Homo sapiens

<400> 220

caggtccagc tggtacagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtc 60
tcctgcaagg tttccggata caccctcact 90

<210> 221
<211> 90
<212> ADN
<213> Homo sapiens

<400> 221

cagatgcagc tggtgcagtc tggggctgag gtgaagaaga ctgggtcctc agtgaaggtt 60
tcctgcaagg cttccggata caccttcacc 90

<210> 222
<211> 90
<212> ADN
<213> Homo sapiens

<400> 222

caggtgcagc tggtgcagtc tggggctgag gtgaagaagc ctggggcctc agtgaaggtt 60
tcctgcaagg catctggata caccttcacc 90

<210> 223
<211> 90
<212> ADN
<213> Homo sapiens

<400> 223

caaatgcagc tggtgcagtc tgggcctgag gtgaagaagc ctgggacctc agtgaaggtc 60
tcctgcaagg cttctggatt cacctttact 90

<210> 224
<211> 90
<212> ADN
<213> Homo sapiens

<400> 224

caggtgcagc tggtgcagtc tggggctgag gtgaagaagc ctgggtcctc ggtgaaggtc 60
tcctgcaagg cttctggagg caccttcagc 90

<210> 225
<211> 90
<212> ADN
<213> Homo sapiens

<400> 225

caggtgcagc tggtgcagtc tggggctgag gtgaagaagc ctgggtcctc ggtgaaggtc 60
tcctgcaagg cttctggagg caccttcagc 90

<210> 226
<211> 90
<212> ADN
<213> Homo sapiens

<400> 226

gaggtccagc tggtacagtc tggggctgag gtgaagaagc ctggggctac agtgaaaatc 60
tcctgcaagg tttctggata caccttcacc 90

<210> 227
<211> 90
<212> ADN
<213> Homo sapiens

<400> 227

cagatcacct tgaaggagtc tggtcctacg ctggtgaaac ccacacagac cctcacgctg 60
acctgcacct tctctgggtt ctcactcagc 90

<210> 228
<211> 90
<212> ADN
<213> Homo sapiens

<400> 228

caggtcacct tgaaggagtc tggtcctgtg ctggtgaaac ccacagagac cctcacgctg 60
acctgcaccg tctctgggtt ctcactcagc 90

<210> 229
<211> 90
<212> ADN
<213> Homo sapiens

<400> 229

caggtcacct tgaaggagtc tggtcctgcg ctggtgaaac ccacacagac cctcacactg 60
acctgcacct tctctgggtt ctcactcagc 90

<210> 230
<211> 90
<212> ADN
<213> Homo sapiens

<400> 230

gaggtgcagc tggtggagtc tgggggaggc ttggtccagc ctgggggggtc cctgagactc 60
tcctgtgcag cctctggatt caccctttagt 90

<210> 231
<211> 90
<212> ADN
<213> Homo sapiens

<400> 231

gaagtgcagc tggtggagtc tgggggaggc ttggtacagc ctggcaggtc cctgagactc 60
tcctgtgcag cctctggatt caccctttgat 90

<210> 232
<211> 90
<212> ADN
<213> Homo sapiens

<400> 232

caggtgcagc tggtggagtc tgggggaggc ttggtcaagc ctggagggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 233
<211> 90
<212> ADN
<213> Homo sapiens

<400> 233

gaggtgcagc tggtggagtc tgggggaggc ttggtacagc ctgggggggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 234
<211> 90
<212> ADN
<213> Homo sapiens

<400> 234

gaggtgcagc tggtggagtc tgggggaggc ttggtaaagc ctgggggggtc ccttagactc 60
tcctgtgcag cctctggatt cactttcagt 90

<210> 235
<211> 90
<212> ADN
<213> Homo sapiens

<400> 235

gaggtgcagc tggtggagtc tgggggaggt gtggtacggc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt cacctttgat 90

<210> 236
<211> 90
<212> ADN
<213> Homo sapiens

<400> 236

gaggtgcagc tggtggagtc tgggggaggc ctggtcaagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 237
<211> 90
<212> ADN
<213> Homo sapiens

<400> 237

gaggtgcagc tgttggagtc tgggggaggc ttggtacagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt cacctttagc 90

<210> 238
<211> 90
<212> ADN
<213> Homo sapiens

<400> 238

caggtgcagc tggtggagtc tgggggaggc gtggtccagc ctgggaggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 239
<211> 90
<212> ADN
<213> Homo sapiens

<400> 239

caggtgcagc tggtggagtc tgggggaggc gtggtccagc ctgggaggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 240
<211> 90
<212> ADN
<213> Homo sapiens

<400> 240

caggtgcagc tggtggagtc tgggggaggc gtggtccagc ctgggaggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 241
<211> 90
<212> ADN
<213> Homo sapiens

<400> 241

caggtgcagc tggtggagtc tgggggaggc gtggtccagc ctgggaggtc cctgagactc 60
tcctgtgcag cgtctggatt caccttcagt 90

<210> 242
<211> 90
<212> ADN
<213> Homo sapiens

<400> 242

gaagtgcagc tggtggagtc tgggggagtc gtggtacagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccTttgat 90

<210> 243
<211> 90
<212> ADN
<213> Homo sapiens

<400> 243

gaggtgcagc tggtggagtc tgggggaggc ttggtacagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 244
<211> 90
<212> ADN
<213> Homo sapiens

<400> 244

gaggtgcagc tggtggagtc tgggggaggc ttggtacagc cagggcggtc cctgagactc 60
tcctgtacag cttctggatt caccTttggt 90

<210> 245
<211> 90
<212> ADN
<213> Homo sapiens

<400> 245

gaggtgcagc tggtggagac tggaggaggc ttgatccagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctgggtt caccgtcagt 90

<210> 246
<211> 90
<212> ADN
<213> Homo sapiens

<400> 246

gaggtgcagc tggtggagtc tgggggaggc ttggtccagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 247
<211> 90
<212> ADN
<213> Homo sapiens

<400> 247

gaggtgcagc tggtggagtc tgggggaggc ttggtccagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccgtcagt 90

<210> 248
<211> 90
<212> ADN
<213> Homo sapiens

<400> 248

gaggtgcagc tggtggagtc tgggggaggc ttggtccagc ctggagggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 249
<211> 90
<212> ADN
<213> Homo sapiens

<400> 249

gaggtgcagc tggtggagtc tgggggaggc ttggtccagc ctgggggGtc cctgaaactc 60
tcctgtgcag cctctgggtt caccttcagt 90

<210> 250
<211> 90
<212> ADN
<213> Homo sapiens

<400> 250

gaggtgcagc tggtggagtc cgggggaggc ttagttcagc ctgggggGtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt 90

<210> 251
<211> 90
<212> ADN
<213> Homo sapiens

<400> 251

gaggtgcagc tggtggagtc tcggggagtc ttggtacagc ctgggggggtc cctgagactc 60
tcctgtgcag cctctggatt caccgtcagt 90

<210> 252
<211> 90
<212> ADN
<213> Homo sapiens

<400> 252
caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggggac cctgtccctc 60
acctgcgctg tctctggtgg ctccatcagc 90

<210> 253
<211> 90
<212> ADN
<213> Homo sapiens

<400> 253

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggacac cctgtccctc 60
acctgcgctg tctctggtta ctccatcagc 90

<210> 254
<211> 90
<212> ADN
<213> Homo sapiens

<400> 254

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcacagac cctgtccctc 60
acctgcactg tctctggtgg ctccatcagc 90

<210> 255
<211> 90
<212> ADN
<213> Homo sapiens

<400> 255

cagctgcagc tgcaggagtc cggctcagga ctggtgaagc cttcacagac cctgtccctc 60
acctgcgctg tctctggtgg ctccatcagc 90

<210> 256
<211> 90
<212> ADN
<213> Homo sapiens

<400> 256

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcacagac cctgtccctc 60
acctgcactg tctctggtgg ctccatcagc 90

<210> 257
<211> 90
<212> ADN
<213> Homo sapiens

<400> 257

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcacagac cctgtccctc 60
acctgcactg tctctggtgg ctccatcagc 90

<210> 258
<211> 90
<212> ADN
<213> Homo sapiens

<400> 258

caggtgcagc tacagcagtg gggcgcagga ctgttgaagc cttcggagac cctgtccctc 60
acctgcgctg tctatggtgg gtccttcagt 90

<210> 259
<211> 90
<212> ADN
<213> Homo sapiens

<400> 259

cagctgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggagac cctgtccctc 60
acctgcactg tctctggtgg ctccatcagc 90

<210> 260
<211> 90
<212> ADN
<213> Homo sapiens

<400> 260

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggagac cctgtccctc 60
acctgcactg tctctggtgg ctccatcagt 90

<210> 261
<211> 90
<212> ADN
<213> Homo sapiens

<400> 261

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggagac cctgtccctc 60
acctgcactg tctctggtgg ctccgtcagc 90

<210> 262
<211> 90
<212> ADN
<213> Homo sapiens

<400> 262

caggtgcagc tgcaggagtc gggcccagga ctggtgaagc cttcggagac cctgtccctc 60
acctgcgctg tctctggtta ctccatcagc 90

<210> 263
<211> 90
<212> ADN
<213> Homo sapiens

<400> 263

gaggtgcagc tggtgcagtc tggagcagag gtgaaaaagc ccggggagtc tctgaagatc 60
tcctgtaagg gttctggata cagctttacc 90

<210> 264
<211> 90
<212> ADN
<213> Homo sapiens

<400> 264

gaagtgcagc tggtgcagtc tggagcagag gtgaaaaagc ccggggagtc tctgaggatc 60
tcctgtaagg gttctggata cagctttacc 90

<210> 265
<211> 90
<212> ADN
<213> Homo sapiens

<400> 265

caggtacagc tgcagcagtc aggtccagga ctggtgaagc cctcgcagac cctctcactc 60
acctgtgcca tctccgggga cagtgtctct 90

<210> 266
<211> 90
<212> ADN
<213> Homo sapiens

<400> 266

caggtgcagc tggtgcaatc tgggtctgag ttgaagaagc ctggggcctc agtgaaggtt 60
tcctgcaagg cttctggata caccttcact 90

<210> 267
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 267
ccgtgtatta ctgtgcgaga ga 22

<210> 268
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 268
ctgtgtatta ctgtgcgaga ga 22

<210> 269
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 269
ccgtgtatta ctgtgcgaga gg 22

<210> 270
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 270
ccgtatatta ctgtgcgaaa ga 22

<210> 271
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 271
ctgtgtatta ctgtgcgaaa ga 22

<210> 272
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 272
ctgtgtatta ctgtgcgaga ca 22

<210> 273
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 273
ccatgtatta ctgtgcgaga ca 22

<210> 274
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 274
ccatgtatta ctgtgcgaga aa 22

<210> 275
<211> 69
<212> ADN
<213> Homo sapiens

<400> 275

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 276
<211> 69
<212> ADN
<213> Homo sapiens

<400> 276

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 277
<211> 69
<212> ADN
<213> Homo sapiens

<400> 277

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 278
<211> 69
<212> ADN
<213> Homo sapiens

<400> 278

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 279
<211> 69
<212> ADN
<213> Homo sapiens

<400> 279

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 280
<211> 69
<212> ADN
<213> Homo sapiens

<400> 280

gacatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 281
<211> 69
<212> ADN
<213> Homo sapiens

<400> 281

aacatccaga tgacccagtc tccatctgcc atgtctgcat ctgtaggaga cagagtcacc 60
atcacttgt 69

<210> 282
<211> 69
<212> ADN
<213> Homo sapiens

<400> 282

gacatccaga tgacccagtc tccatcctca ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgt 69

<210> 283
<211> 69
<212> ADN
<213> Homo sapiens

<400> 283

gacatccaga tgacccagtc tccatcctca ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgt 69

<210> 284
<211> 69
<212> ADN
<213> Homo sapiens

<400> 284

gccatccagt tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 285
<211> 69
<212> ADN
<213> Homo sapiens

<400> 285

gccatccagt tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 286
<211> 69
<212> ADN
<213> Homo sapiens

<400> 286

gacatccaga tgacccagtc tccatcttcc gtgtctgcat ctgtaggaga cagagtcacc 60
atcacttgt 69

<210> 287
<211> 69
<212> ADN
<213> Homo sapiens

<400> 287

gacatccaga tgacccagtc tccatcttct gtgtctgcat ctgtaggaga cagagtcacc 60
atcacttgt 69

<210> 288
<211> 69
<212> ADN
<213> Homo sapiens

<400> 288

gacatccagt tgacccagtc tccatccttc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 289
<211> 69
<212> ADN
<213> Homo sapiens

<400> 289

gccatccgga tgacccagtc tccattctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 290
<211> 69
<212> ADN
<213> Homo sapiens

<400> 290

gccatccgga tgacccagtc tccatcctca ttctctgcat ctacaggaga cagagtcacc 60
atcacttgt 69

<210> 291
<211> 69
<212> ADN
<213> Homo sapiens

<400> 291

gtcatctgga tgacccagtc tccatcctta ctctctgcat ctacaggaga cagagtcacc 60
atcagttgt 69

<210> 292
<211> 69
<212> ADN
<213> Homo sapiens

<400> 292

gccatccaga tgacccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 293
<211> 69
<212> ADN
<213> Homo sapiens

<400> 293

gacatccaga tgacccagtc tccttccacc ctgtctgcat ctgtaggaga cagagtcacc 60
atcacttgc 69

<210> 294
<211> 69
<212> ADN
<213> Homo sapiens

<400> 294

gatattgtga tgacccagac tccactctcc ctgcccgtca cccctggaga gccggcctcc 60
atctcctgc 69

<210> 295
<211> 69
<212> ADN
<213> Homo sapiens

<400> 295

gatattgtga tgacccagac tccactctcc ctgcccgtca cccctggaga gccggcctcc 60
atctcctgc 69

<210> 296
<211> 69
<212> ADN
<213> Homo sapiens

<400> 296

gatgttgtga tgactcagtc tccactctcc ctgcccgtca cccttggaca gccggcctcc 60
atctcctgc 69

<210> 297
<211> 69
<212> ADN
<213> Homo sapiens

<400> 297

gatgttgtga tgactcagtc tccactctcc ctgcccgtca cccttggaca gccggcctcc 60
atctcctgc 69

<210> 298
<211> 69
<212> ADN
<213> Homo sapiens

<400> 298

gatattgtga tgacccagac tccactctct ctgtccgtca cccctggaca gccggcctcc 60
atctcctgc 69

<210> 299
<211> 69
<212> ADN
<213> Homo sapiens

<400> 299

gatattgtga tgacccagac tccactctct ctgtccgtca cccctggaca gccggcctcc 60
atctcctgc 69

<210> 300
<211> 69
<212> ADN
<213> Homo sapiens

<400> 300

gatattgtga tgactcagtc tccactctcc ctgcccgtca cccctggaga gccggcctcc 60
atctcctgc 69

<210> 301
<211> 69
<212> ADN
<213> Homo sapiens

<400> 301

gatattgtga tgactcagtc tccactctcc ctgcccgtca cccctggaga gccggcctcc 60
atctcctgc 69

<210> 302
<211> 69
<212> ADN
<213> Homo sapiens

<400> 302

gatattgtga tgacccagac tccactctcc tcacctgtca cccttggaca gccggcctcc 60
atctcctgc 69

<210> 303
<211> 69
<212> ADN
<213> Homo sapiens

<400> 303

gaaattgtgt tgacgcagtc tccaggcacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 304
<211> 69
<212> ADN
<213> Homo sapiens

<400> 304

gaaattgtgt tgacgcagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 305
<211> 69
<212> ADN
<213> Homo sapiens

<400> 305

gaaatagtga tgacgcagtc tccagccacc ctgtctgtgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 306
<211> 69
<212> ADN
<213> Homo sapiens

<400> 306

gaaatagtga tgacgcagtc tccagccacc ctgtctgtgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 307
<211> 69
<212> ADN
<213> Homo sapiens

<400> 307

gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 308
<211> 69
<212> ADN
<213> Homo sapiens

<400> 308

gaaattgtgt tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 309
<211> 69
<212> ADN
<213> Homo sapiens

<400> 309

gaaattgtaa tgacacagtc tccagccacc ctgtctttgt ctccagggga aagagccacc 60
ctctcctgc 69

<210> 310
<211> 69
<212> ADN
<213> Homo sapiens

<400> 310

gacatcgtga tgacccagtc tccagactcc ctggctgtgt ctctgggcga gagggccacc 60
atcaactgc 69

<210> 311
<211> 69
<212> ADN
<213> Homo sapiens

<400> 311

gaaacgacac tcacgcagtc tccagcattc atgtcagcga ctccaggaga caaagtcaac 60
atctcctgc 69

<210> 312
<211> 69
<212> ADN
<213> Homo sapiens

<400> 312

gaaattgtgc tgactcagtc tccagacttt cagtctgtga ctccaaagga gaaagtcacc 60
atcacctgc 69

<210> 313
<211> 69
<212> ADN
<213> Homo sapiens

<400> 313

gaaattgtgc tgactcagtc tccagacttt cagtctgtga ctccaaagga gaaagtcacc 60
atcacctgc 69

<210> 314
<211> 69
<212> ADN
<213> Homo sapiens

<400> 314

gatgttgtga tgacacagtc tccagctttc ctctctgtga ctccagggga gaaagtcacc 60
atcacctgc 69

<210> 315
<211> 66
<212> ADN
<213> Homo sapiens

<400> 315

cagtctgtgc tgactcagcc accctcggtg tctgaagccc ccaggcagag ggtcaccatc 60
tcctgt 66

<210> 316
<211> 66
<212> ADN
<213> Homo sapiens

<400> 316

cagtctgtgc tgacgcagcc gccctcagtg tctggggccc cagggcagag ggtcaccatc 60
tcctgc 66

<210> 317
<211> 66
<212> ADN
<213> Homo sapiens

<400> 317

cagtctgtgc tgactcagcc accctcagcg tctgggaccc ccgggcagag ggtcaccatc 60
tcttgt 66

<210> 318
<211> 66
<212> ADN
<213> Homo sapiens

<400> 318

cagtctgtgc tgactcagcc accctcagcg tctgggaccc ccgggcagag ggtcaccatc 60
tcttgt 66

<210> 319
<211> 66
<212> ADN
<213> Homo sapiens

<400> 319

cagtctgtgt tgacgcagcc gccctcagtg tctgcggccc caggacagaa ggtcaccatc 60
tcctgc 66

<210> 320
<211> 66
<212> ADN
<213> Homo sapiens

<400> 320

cagtctgccc tgactcagcc tccctccgcg tccgggtctc ctggacagtc agtcaccatc 60
tcctgc 66

<210> 321
<211> 66
<212> ADN
<213> Homo sapiens

<400> 321

cagtctgccc tgactcagcc tcgctcagtg tccgggtctc ctggacagtc agtcaccatc 60
tcctgc 66

<210> 322
<211> 66
<212> ADN
<213> Homo sapiens

<400> 322

cagtctgccc tgactcagcc tgcctccgtg tctgggtctc ctggacagtc gatcaccatc 60
tcctgc 66

<210> 323
<211> 66
<212> ADN
<213> Homo sapiens

<400> 323

cagtctgccc tgactcagcc tccctccgtg tccgggtctc ctggacagtc agtcaccatc 60
tcctgc 66

<210> 324
<211> 66
<212> ADN
<213> Homo sapiens

<400> 324

cagtctgccc tgactcagcc tgcctccgtg tctgggtctc ctggacagtc gatcaccatc 60
tcctgc 66

<210> 325
<211> 66
<212> ADN
<213> Homo sapiens

<400> 325

tcctatgagc tgactcagcc accctcagtg tccgtgtccc caggacagac agccagcatc 60
acctgc 66

<210> 326
<211> 66
<212> ADN
<213> Homo sapiens

<400> 326

tcctatgagc tgactcagcc actctcagtg tcagtggccc tgggacagac ggccaggatt 60
acctgt 66

<210> 327
<211> 66
<212> ADN
<213> Homo sapiens

<400> 327

tcctatgagc tgacacagcc accctcggtg tcagtgtccc caggacaaac ggccaggatc 60
acctgc 66

<210> 328
<211> 66
<212> ADN
<213> Homo sapiens

<400> 328

tcctatgagc tgacacagcc accctcggtg tcagtgtccc taggacagat ggccaggatc 60
acctgc 66

<210> 329
<211> 66
<212> ADN
<213> Homo sapiens

<400> 329

tcttctgagc tgactcagga ccctgctgtg tctgtggcct tgggacagac agtcaggatc 60
acatgc 66

<210> 330
<211> 66
<212> ADN
<213> Homo sapiens

<400> 330

tcctatgtgc tgactcagcc accctcagtg tcagtggccc caggaaagac ggccaggatt 60
acctgt 66

<210> 331
<211> 66
<212> ADN
<213> Homo sapiens

<400> 331

tcctatgagc tgacacagct accctcggtg tcagtgtccc caggacagac agccaggatc 60
acctgc 66

<210> 332
<211> 66
<212> ADN
<213> Homo sapiens

<400> 332

tcctatgagc tgatgcagcc accctcggtg tcagtgtccc caggacagac ggccaggatc 60
acctgc 66

<210> 333
<211> 66
<212> ADN
<213> Homo sapiens

<400> 333

tcctatgagc tgacacagcc atcctcagtg tcagtgtctc cgggacagac agccaggatc 60
acctgc 66

<210> 334
<211> 66
<212> ADN
<213> Homo sapiens

<400> 334

ctgcctgtgc tgactcagcc cccgtctgca tctgccttgc tgggagcctc gatcaagctc 60
acctgc 66

<210> 335
<211> 66
<212> ADN
<213> Homo sapiens

<400> 335

cagcctgtgc tgactcaatc atcctctgcc tctgcttccc tgggatcctc ggtcaagctc 60
acctgc 66

<210> 336
<211> 66
<212> ADN
<213> Homo sapiens

<400> 336

cagcttgtgc tgactcaatc gccctctgcc tctgcctccc tgggagcctc ggtcaagctc 60
acctgc 66

<210> 337
<211> 66
<212> ADN
<213> Homo sapiens

<400> 337

cagcctgtgc tgactcagcc accttcctcc tccgcatctc ctggagaatc cgccagactc 60
acctgc 66

<210> 338
<211> 66
<212> ADN
<213> Homo sapiens

<400> 338

caggctgtgc tgactcagcc ggcttccctc tctgcatctc ctggagcatc agccagtctc 60
acctgc 66

<210> 339
<211> 66
<212> ADN
<213> Homo sapiens

<400> 339

cagcctgtgc tgactcagcc atcttcccat tctgcatctt ctggagcatc agtcagactc 60
acctgc 66

<210> 340
<211> 66
<212> ADN
<213> Homo sapiens

<400> 340

aattttatgc tgactcagcc ccactctgtg tcggagtctc cggggaagac ggtaaccatc 60
tcctgc 66

<210> 341
<211> 66
<212> ADN
<213> Homo sapiens

<400> 341

cagactgtgg tgactcagga gccctcactg actgtgtccc caggagggac agtcactctc 60
acctgt 66

<210> 342
<211> 66
<212> ADN
<213> Homo sapiens

<400> 342

caggctgtgg tgactcagga gccctcactg actgtgtccc caggagggac agtcactctc 60
acctgt 66

<210> 343
<211> 66
<212> ADN
<213> Homo sapiens

<400> 343

cagactgtgg tgacccagga gccatcgttc tcagtgtccc ctggagggac agtcacactc 60
acttgt 66

<210> 344
<211> 66
<212> ADN
<213> Homo sapiens

<400> 344

cagcctgtgc tgactcagcc accttctgca tcagcctccc tgggagcctc ggtcacactc 60
acctgc 66

<210> 345
<211> 66
<212> ADN
<213> Homo sapiens

<400> 345

caggcagggc tgactcagcc accctcggtg tccaagggct tgagacagac cgccacactc 60
acctgc 66

<210> 346
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 346
nnnnnngact c 11

<210> 347
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 347
gagtcnnnnn n 11

<210> 348
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (3)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 348
gcnnnnnnng c 11

<210> 349
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 349
acctgcnnnn n 11

<210> 350
<211> 25
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 350
cacatccgtg ttgttcacgg atgtg 25

<210> 351
<211> 88
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 351
aatagtagac tgcagtgtcc tcagccctta agctgttcat ctgcaagtag agagtattct 60
tagagttgtc tctagactta gtgaagcg 88

<210> 352
<211> 88
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 352
cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agggctgagg acactgcagt ctactatt 88

<210> 353
<211> 95
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 353
cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agggctgagg acactgcagt ctactattgt gcgag 95

<210> 354
<211> 95
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 354
cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agggctgagg acactgcagt ctactattgt acgag 95

<210> 355
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 355
cgcttcacta agtctagaga caac 24

<210> 356
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (8)..(15)
<223> A, T, C, G, outro ou desconhecido

<400> 356
cacctgcnnn nnnnn 15

<210> 357
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 357
cagctcnnnn nnnnnnn 17

<210> 358
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 358
gaagacnnnn nnnnnnn 17

<210> 359
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 359
gcagcnnnnn nnnnnnn 17

<210> 360
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 360
gaagacnnnn nn 12

<210> 361
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(22)
<223> A, T, C, G, outro ou desconhecido

<400> 361
cttgagnnnn nnnnnnnnnn nn 22

<210> 362
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(19)
<223> A, T, C, G, outro ou desconhecido

<400> 362
acggcnnnnn nnnnnnnnn 19

<210> 363
<211> 18
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(18)
<223> A, T, C, G, outro ou desconhecido

<400> 363
acggcnnnnn nnnnnnnn 18

<210> 364
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 364
gtatccnnnn nn 12

<210> 365
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 365
actgggnnnn n 11

<210> 366
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 366
ggatcnnnnn 10

<210> 367
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 367
gcatcnnnnn n 11

<210> 368
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(16)
<223> A, T, C, G, outro ou desconhecido

<400> 368
gaggagnnnn nnnnnn 16

<210> 369
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(19)
<223> A, T, C, G, outro ou desconhecido

<400> 369
gggacnnnnn nnnnnnnnn 19

<210> 370
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(14)
<223> A, T, C, G, outro ou desconhecido

<400> 370
acctgcnnnn nnnn 14

<210> 371
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 371
ggcggannnn nnnnnnn 17

<210> 372
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(22)
<223> A, T, C, G, outro ou desconhecido

<400> 372
ctgaagnnnn nnnnnnnnnn nn 22

<210> 373
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 373
cccgcnnnnn n 11

<210> 374
<211> 18
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(18)
<223> A, T, C, G, outro ou desconhecido

<400> 374
ggatgnnnnn nnnnnnnn 18

<210> 375
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(22)
<223> A, T, C, G, outro ou desconhecido

<400> 375
ctggagnnnn nnnnnnnnnn nn 22

<210> 376
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(15)
<223> A, T, C, G, outro ou desconhecido

<400> 376
gacgcnnnnn nnnnn 15

<210> 377
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(13)
<223> A, T, C, G, outro ou desconhecido

<400> 377
ggtgannnnn nnn 13

<210> 378
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(13)
<223> A, T, C, G, outro ou desconhecido

<400> 378
gaagannnnn nnn 13

<210> 379
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 379
gagtcnnnnn 10

<210> 380
<211> 26
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(26)
<223> A, T, C, G, outro ou desconhecido

<400> 380
tccracnnnn nnnnnnnnnn nnnnnn 26

<210> 381
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (5)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 381
cctcnnnnnn n 11

<210> 382
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 382
gagtcnnnnn 10

<210> 383
<211> 18
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(18)
<223> A, T, C, G, outro ou desconhecido

<400> 383
cccacannnn nnnnnnnn 18

<210> 384
<211> 14
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(14)
<223> A, T, C, G, outro ou desconhecido

<400> 384
gcatcnnnnn nnnn 14

<210> 385
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(13)
<223> A, T, C, G, outro ou desconhecido

<400> 385
ggtgannnnn nnn 13

<210> 386
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (5)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 386
cccgnnnnnn nn 12

<210> 387
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (6)..(19)
<223> A, T, C, G, outro ou desconhecido

<400> 387
ggatgnnnnn nnnnnnnnn 19

<210> 388
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 388
gaccgannnn nnnnnnn 17

<210> 389
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 389
cacccannnn nnnnnnn 17

<210> 390
<211> 17
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(17)
<223> A, T, C, G, outro ou desconhecido

<400> 390
caarcannnn nnnnnnn 17

<210> 391
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 391
gctgtgtatt actgtgcgag 20

<210> 392
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 392
gccgtgtatt actgtgcgag 20

<210> 393
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 393
gccgtatatt actgtgcgag 20

<210> 394
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 394
gccgtgtatt actgtacgag 20

<210> 395
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sonda sintética

<400> 395
gccatgtatt actgtgcgag 20

<210> 396
<211> 25
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 396
cacatccgtg ttgttcacgg atgtg 25

<210> 397
<211> 88
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 397

aatagtagac tgcagtgtcc tcagccctta agctgttcat ctgcaagtag agagtattct 60
tagagttgtc tctagactta gtgaagcg 88

<210> 398
<211> 95
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 398

cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agggctgagg acactgcagt ctactattgt gcgag 95

<210> 399
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 399
cgcttcacta agtctagaga caac 24

<210> 400
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 400
cacatccgtg ttgttcacgg atgtgggagg atggagactg ggtc 44

<210> 401
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 401
cacatccgtg ttgttcacgg atgtgggaga gtggagactg agtc 44

<210> 402
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 402
cacatccgtg ttgttcacgg atgtgggtgc ctggagactg cgtc 44

<210> 403
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 403
cacatccgtg ttgttcacgg atgtgggtgg ctggagactg cgtc 44

<210> 404
<211> 34
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 404
cctctactct tgtcacagtg cacaagacat ccag 34

<210> 405
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 405
cctctactct tgtcacagtg 20

<210> 406
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 406
ggaggatgga ctggatgtct tgtgcactgt gacaagagta gagg 44

<210> 407
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 407
ggagagtgga ctggatgtct tgtgcactgt gacaagagta gagg 44

<210> 408
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 408
ggtgcctgga ctggatgtct tgtgcactgt gacaagagta gagg 44

<210> 409
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 409
ggtggctgga ctggatgtct tgtgcactgt gacaagagta gagg 44

<210> 410
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 410
cacatccgtg ttgttcacgg atgtggatcg actgtccagg agac 44

<210> 411
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 411
cacatccgtg ttgttcacgg atgtggactg tctgtcccaa ggcc 44

<210> 412
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 412
cacatccgtg ttgttcacgg atgtggactg actgtccagg agac 44

<210> 413
<211> 44
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 413
cacatccgtg ttgttcacgg atgtggaccc tctgccctgg ggcc 44

<210> 414
<211> 59
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 414
cctctgactg agtgcacaga gtgctttaac ccaaccggct agtgttagcg gttccccgg 59

<210> 415
<211> 69
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 415

cctctgactg agtgcacaga gtgctttaac ccaaccggct agtgttagcg gttccccggg 60
acagtcgat 69

<210> 416
<211> 69
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 416

cctctgactg agtgcacaga gtgctttaac ccaaccggct agtgttagcg gttccccggg 60
acagacagt 69

<210> 417
<211> 69
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 417

cctctgactg agtgcacaga gtgctttaac ccaaccggct agtgttagcg gttccccggg 60
acagtcagt 69

<210> 418
<211> 70
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 418

cctctgactg agtgcacaga gtgctttaac ccaaccggct agtgttagcg gtstccccgg 60
ggcagagggt 70

<210> 419
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<400> 419
cctctgactg agtgcacaga gtgc 24

<210> 420
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (5)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 420
ggccnnnnng gcc 13

<210> 421
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 421
ccannnnnnn nntgg 15

<210> 422
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 422
cgannnnnnt gc 12

<210> 423
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 423
gccnnnnngg c 11

<210> 424
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 424
gatnnnnatc 10

<210> 425
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 425
gacnnnnngt c 11

<210> 426
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 426
gcannnnntg c 11

<210> 427
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 427
gtatccnnnn nn 12

<210> 428
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 428
gacnnnnnng tc 12

<210> 429
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 429
ccannnnntg g 11

<210> 430
<211> 12
<212> ADN

<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 430
nnnnnngaga cg 12

<210> 431
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 431
ccannnnnnt gg 12

<210> 432
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 432
gaannnnttc 10

<210> 433
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 433
ggtctcnnnn n 11

<210> 434
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 434
nnnnnnnnnn ctcctc 16

<210> 435
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 435
nnnnnnnnnt ccgcc 15

<210> 436
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (5)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 436
ggccnnnnng gcc 13

<210> 437
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 437
ccannnnnnt gg 12

<210> 438
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 438
gacnnnnnng tc 12

<210> 439
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 439
cgannnnnnt gc 12

<210> 440
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 440
gcannnnntg c 11

<210> 441
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 441
ccannnnntg g 11

<210> 442
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 442
gaannnnttc 10

<210> 443
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 443
nnnnnngaga cg 12

<210> 444
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 444
gtatccnnnn nn 12

<210> 445
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, C, G, outro ou desconhecido

<400> 445
gacnnnnngt c 11

<210> 446
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 446
ggtctcnnnn n 11

<210> 447
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 447
gccnnnnngg c 11

<210> 448
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (4)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 448
ccannnnnnn nntgg 15

<210> 449
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 449
nnnnnnnnnn ctcctc 16

<210> 450
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido Sintético

<220>
<221> base_modificada
<222> (1)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 450
nnnnnnnnnt ccgcc 15

<210> 451
<211> 9532
<212> ADN
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência nucleotídica MALIA3

<220>
<221> CDS
<222> (1579)..(1638)

<220>
<221> CDS
<222> (2343)..(3443)

<220>
<221> CDS
<222> (3945)..(4400)

<220>
<221> CDS
<222> (4406)..(4450)

<220>
<221> CDS
<222> (4746)..(5789)

<400> 451
aatgctacta ctattagtag aattgatgcc acctttcag ctcgcgcccc aaatgaaaat 60

atagctaaac aggttattga ccatttgcga aatgtatcta atggtcaaac taaatctact 120

cgttcgcaga attgggaatc aactgttaca tggaatgaaa cttccagaca ccgtacttta 180

gttgcatatt taaaacatgt tgagctacag caccagattc agcaattaag ctctaagcca 240

tccgcaaaaa tgacctctta tcaaaaggag caattaaagg tactctctaa tcctgacctg 300

ttggagtttg cttccggtct ggttcgcttt gaagctcgaa ttaaaacgcg atatttgaag 360

tctttcgggc ttcctcttaa tctttttgat gcaatccgct ttgcttctga ctataatagt 420

cagggtaaag acctgatttt tgatttatgg tcattctcgt tttctgaact gtttaaagca 480

tttgaggggg attcaatgaa tatttatgac gattccgcag tattggacgc tatccagtct 540

aaacatttta ctattacccc ctctggcaaa acttcttttg caaaagcctc tcgctatttt 600

```
ggtttttatc gtcgtctggt aaacgagggt tatgatagtg ttgctcttac tatgcctcgt 660

aattcctttt ggcgttatgt atctgcatta gttgaatgtg gtattcctaa atctcaactg 720

atgaatcttt ctacctgtaa taatgttgtt ccgttagttc gttttattaa cgtagatttt 780

tcttcccaac gtcctgactg gtataatgag ccagttctta aaatcgcata aggtaattca 840

caatgattaa agttgaaatt aaaccatctc aagcccaatt tactactcgt tctggtgttt 900

ctcgtcaggg caagccttat tcactgaatg agcagctttg ttacgttgat ttgggtaatg 960

aatatccggt tcttgtcaag attactcttg atgaaggtca gccagcctat gcgcctggtc 1020

tgtacaccgt tcatctgtcc tctttcaaag ttggtcagtt cggttccctt atgattgacc 1080

gtctgcgcct cgttccggct aagtaacatg gagcaggtcg cggatttcga cacaatttat 1140

caggcgatga tacaaatctc cgttgtactt tgtttcgcgc ttggtataat cgctgggggt 1200

caaagatgag tgttttagtg tattctttcg cctctttcgt tttaggttgg tgccttcgta 1260

gtggcattac gtattttacc cgtttaatgg aaacttcctc atgaaaaagt ctttagtcct 1320

caaagcctct gtagccgttg ctaccctcgt tccgatgctg tctttcgctg ctgagggtga 1380

cgatcccgca aaagcggcct ttaactccct gcaagcctca gcgaccgaat atatcggtta 1440

tgcgtgggcg atggttgttg tcattgtcgg cgcaactatc ggtatcaagc tgtttaagaa 1500

attcacctcg aaagcaagct gataaaccga tacaattaaa ggctcctttt ggagcctttt 1560

tttttggaga ttttcaac gtg aaa aaa tta tta ttc gca att cct tta gtt   1611
                  Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val
                    1               5                  10

gtt cct ttc tat tct cac agt gca cag tctgtcgtga cgcagccgcc         1658
Val Pro Phe Tyr Ser His Ser Ala Gln
             15                  20

ctcagtgtct ggggccccag ggcagagggt caccatctcc tgcactggga gcagctccaa 1718

catcggggca ggttatgatg tacactggta ccagcagctt ccaggaacag cccccaaact 1778

cctcatctat ggtaacagca atcggccctc aggggtccct gaccgattct ctggctccaa 1838

gtctggcacc tcagcctccc tggccatcac tgggctccag gctgaggatg aggctgatta 1898

ttactgccag tcctatgaca gcagcctgag tggcctttat gtcttcggaa ctgggaccaa 1958

ggtcaccgtc ctaggtcagc ccaaggccaa ccccactgtc actctgttcc cgccctcctc 2018

tgaggagctc caagccaaca aggccacact agtgtgtctg atcagtgact tctacccggg 2078

agctgtgaca gtggcctgga aggcagatag cagccccgtc aaggcgggag tggagaccac 2138

cacaccctcc aaacaaagca acaacaagta cgcggccagc agctatctga gcctgacgcc 2198

tgagcagtgg aagtcccaca gaagctacag ctgccaggtc acgcatgaag ggagcaccgt 2258
```

```
ggagaagaca gtggccccta cagaatgttc ataataaacc gcctccaccg ggcgcgccaa 2318

ttctatttca aggagacagt cata atg aaa tac cta ttg cct acg gca gcc    2369
                           Met Lys Tyr Leu Leu Pro Thr Ala Ala
                                                25

gct gga ttg tta tta ctc gcg gcc cag ccg gcc atg gcc gaa gtt caa   2417
Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro Ala Met Ala Glu Val Gln
 30                  35                  40                  45

ttg tta gag tct ggt ggc ggt ctt gtt cag cct ggt ggt tct tta cgt   2465
Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg
                 50                  55                  60

ctt tct tgc gct gct tcc gga ttc act ttc tct tcg tac gct atg tct   2513
Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr Ala Met Ser
             65                  70                  75

tgg gtt cgc caa gct cct ggt aaa ggt ttg gag tgg gtt tct gct atc   2561
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Ala Ile
         80                  85                  90

tct ggt tct ggt ggc agt act tac tat gct gac tcc gtt aaa ggt cgc   2609
Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys Gly Arg
     95                 100                 105

ttc act atc tct aga gac aac tct aag aat act ctc tac ttg cag atg   2657
Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met
110                 115                 120                 125

aac agc tta agg gct gag gac act gca gtc tac tat tgc gct aaa gac   2705
Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp
                130                 135                 140

tat gaa ggt act ggt tat gct ttc gac ata tgg ggt caa ggt act atg   2753
Tyr Glu Gly Thr Gly Tyr Ala Phe Asp Ile Trp Gly Gln Gly Thr Met
            145                 150                 155

gtc acc gtc tct agt gcc tcc acc aag ggc cca tcg gtc ttc ccc ctg   2801
Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
        160                 165                 170

gca ccc tcc tcc aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc   2849
Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys
    175                 180                 185

ctg gtc aag gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg aac tca   2897
Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser
190                 195                 200                 205

ggc gcc ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta cag tct   2945
Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser
                210                 215                 220

agc gga ctc tac tcc ctc agc agc gta gtg acc gtg ccc tct tct agc   2993
Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser
            225                 230                 235

ttg ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc agc aac   3041
Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn
        240                 245                 250
```

```
acc aag gtg gac aag aaa gtt gag ccc aaa tct tgt gcg gcc gct cat   3089
Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Ala Ala Ala His
    255                 260                 265

cac cac cat cat cac tct gct gaa caa aaa ctc atc tca gaa gag gat   3137
His His His His His Ser Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp
270                 275                 280                 285

ctg aat ggt gcc gca gat atc aac gat gat cgt atg gct ggc gcc gct   3185
Leu Asn Gly Ala Ala Asp Ile Asn Asp Asp Arg Met Ala Gly Ala Ala
                290                 295                 300

gaa act gtt gaa agt tgt tta gca aaa ccc cat aca gaa aat tca ttt   3233
Glu Thr Val Glu Ser Cys Leu Ala Lys Pro His Thr Glu Asn Ser Phe
            305                 310                 315

act aac gtc tgg aaa gac gac aaa act tta gat cgt tac gct aac tat   3281
Thr Asn Val Trp Lys Asp Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr
        320                 325                 330

gag ggt tgt ctg tgg aat gct aca ggc gtt gta gtt tgt act ggt gac   3329
Glu Gly Cys Leu Trp Asn Ala Thr Gly Val Val Val Cys Thr Gly Asp
    335                 340                 345

gaa act cag tgt tac ggt aca tgg gtt cct att ggg ctt gct atc cct   3377
Glu Thr Gln Cys Tyr Gly Thr Trp Val Pro Ile Gly Leu Ala Ile Pro
350                 355                 360                 365

gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct gag ggt ggc ggt   3425
Glu Asn Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
                370                 375                 380

tct gag ggt ggc ggt act aaacctcctg agtacggtga tacacctatt          3473
Ser Glu Gly Gly Gly Thr
            385

ccgggctata cttatatcaa ccctctcgac ggcacttatc cgcctggtac tgagcaaaac 3533

cccgctaatc ctaatccttc tcttgaggag tctcagcctc ttaatacttt catgtttcag 3593

aataataggt tccgaaatag gcaggggggca ttaactgttt atacgggcac tgttactcaa 3653

ggcactgacc ccgttaaaac ttattaccag tacactcctg tatcatcaaa agccatgtat 3713

gacgcttact ggaacggtaa attcagagac tgcgctttcc attctggctt taatgaagat 3773

ccattcgttt gtgaatatca aggccaatcg tctgacctgc ctcaacctcc tgtcaatgct 3833

ggcggcggct ctggtggtgg ttctggtggc ggctctgagg gtggtggctc tgagggtggc 3893

ggttctgagg gtggcggctc tgagggaggc ggttccggtg gtggctctgg t tcc ggt  3950
                                                          Ser Gly

gat ttt gat tat gaa aag atg gca aac gct aat aag ggg gct atg acc   3998
Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr
390                 395                 400                 405

gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc aaa ctt   4046
Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu
                410                 415                 420
```

```
gat tct gtc gct act gat tac ggt gct gct atc gat ggt ttc att ggt   4094
Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly
            425                 430                 435

gac gtt tcc ggc ctt gct aat ggt aat ggt gct act ggt gat ttt gct   4142
Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala
        440                 445                 450

ggc tct aat tcc caa atg gct caa gtc ggt gac ggt gat aat tca cct   4190
Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro
    455                 460                 465

tta atg aat aat ttc cgt caa tat tta cct tcc ctc cct caa tcg gtt   4238
Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val
470                 475                 480                 485

gaa tgt cgc cct ttt gtc ttt agc gct ggt aaa cca tat gaa ttt tct   4286
Glu Cys Arg Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser
                490                 495                 500

att gat tgt gac aaa ata aac tta ttc cgt ggt gtc ttt gcg ttt ctt   4334
Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu
            505                 510                 515

tta tat gtt gcc acc ttt atg tat gta ttt tct acg ttt gct aac ata   4382
Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile
        520                 525                 530

ctg cgt aat aag gag tct taatc atg cca gtt ctt ttg ggt att ccg tta 4432
Leu Arg Asn Lys Glu Ser       Met Pro Val Leu Leu Gly Ile Pro Leu
    535                       540                 545

tta ttg cgt ttc ctc ggt ttccttctgg taactttgtt cggctatctg          4480
Leu Leu Arg Phe Leu Gly
    550

cttacttttc ttaaaaaggg cttcggtaag atagctattg ctatttcatt gtttcttgct 4540

cttattattg ggcttaactc aattcttgtg ggttatctct ctgatattag cgctcaatta 4600

ccctctgact ttgttcaggg tgttcagtta attctcccgt ctaatgcgct tccctgtttt 4660

tatgttattc tctctgtaaa ggctgctatt ttcatttttg acgttaaaca aaaaatcgtt 4720

tcttatttgg attgggataa ataat atg gct gtt tat ttt gta act ggc aaa   4772
                          Met Ala Val Tyr Phe Val Thr Gly Lys
                          555                 560

tta ggc tct gga aag acg ctc gtt agc gtt ggt aag att cag gat aaa   4820
Leu Gly Ser Gly Lys Thr Leu Val Ser Val Gly Lys Ile Gln Asp Lys
    565                 570                 575

att gta gct ggg tgc aaa ata gca act aat ctt gat tta agg ctt caa   4868
Ile Val Ala Gly Cys Lys Ile Ala Thr Asn Leu Asp Leu Arg Leu Gln
580                 585                 590                 595

aac ctc ccg caa gtc ggg agg ttc gct aaa acg cct cgc gtt ctt aga   4916
Asn Leu Pro Gln Val Gly Arg Phe Ala Lys Thr Pro Arg Val Leu Arg
                600                 605                 610
```

```
ata ccg gat aag cct tct ata tct gat ttg ctt gct att ggg cgc ggt   4964
Ile Pro Asp Lys Pro Ser Ile Ser Asp Leu Leu Ala Ile Gly Arg Gly
            615                 620                 625

aat gat tcc tac gat gaa aat aaa aac ggc ttg ctt gtt ctc gat gag   5012
Asn Asp Ser Tyr Asp Glu Asn Lys Asn Gly Leu Leu Val Leu Asp Glu
        630                 635                 640

tgc ggt act tgg ttt aat acc cgt tct tgg aat gat aag gaa aga cag   5060
Cys Gly Thr Trp Phe Asn Thr Arg Ser Trp Asn Asp Lys Glu Arg Gln
    645                 650                 655

ccg att att gat tgg ttt cta cat gct cgt aaa tta gga tgg gat att   5108
Pro Ile Ile Asp Trp Phe Leu His Ala Arg Lys Leu Gly Trp Asp Ile
660                 665                 670                 675

att ttt ctt gtt cag gac tta tct att gtt gat aaa cag gcg cgt tct   5156
Ile Phe Leu Val Gln Asp Leu Ser Ile Val Asp Lys Gln Ala Arg Ser
                680                 685                 690

gca tta gct gaa cat gtt gtt tat tgt cgt cgt ctg gac aga att act   5204
Ala Leu Ala Glu His Val Val Tyr Cys Arg Arg Leu Asp Arg Ile Thr
            695                 700                 705

tta cct ttt gtc ggt act tta tat tct ctt att act ggc tcg aaa atg   5252
Leu Pro Phe Val Gly Thr Leu Tyr Ser Leu Ile Thr Gly Ser Lys Met
        710                 715                 720

cct ctg cct aaa tta cat gtt ggc gtt gtt aaa tat ggc gat tct caa   5300
Pro Leu Pro Lys Leu His Val Gly Val Val Lys Tyr Gly Asp Ser Gln
    725                 730                 735

tta agc cct act gtt gag cgt tgg ctt tat act ggt aag aat ttg tat   5348
Leu Ser Pro Thr Val Glu Arg Trp Leu Tyr Thr Gly Lys Asn Leu Tyr
740                 745                 750                 755

aac gca tat gat act aaa cag gct ttt tct agt aat tat gat tcc ggt   5396
Asn Ala Tyr Asp Thr Lys Gln Ala Phe Ser Ser Asn Tyr Asp Ser Gly
                760                 765                 770

gtt tat tct tat tta acg cct tat tta tca cac ggt cgg tat ttc aaa   5444
Val Tyr Ser Tyr Leu Thr Pro Tyr Leu Ser His Gly Arg Tyr Phe Lys
            775                 780                 785

cca tta aat tta ggt cag aag atg aaa tta act aaa ata tat ttg aaa   5492
Pro Leu Asn Leu Gly Gln Lys Met Lys Leu Thr Lys Ile Tyr Leu Lys
        790                 795                 800

aag ttt tct cgc gtt ctt tgt ctt gcg att gga ttt gca tca gca ttt   5540
Lys Phe Ser Arg Val Leu Cys Leu Ala Ile Gly Phe Ala Ser Ala Phe
    805                 810                 815

aca tat agt tat ata acc caa cct aag ccg gag gtt aaa aag gta gtc   5588
Thr Tyr Ser Tyr Ile Thr Gln Pro Lys Pro Glu Val Lys Lys Val Val
820                 825                 830                 835

tct cag acc tat gat ttt gat aaa ttc act att gac tct tct cag cgt   5636
Ser Gln Thr Tyr Asp Phe Asp Lys Phe Thr Ile Asp Ser Ser Gln Arg
                840                 845                 850

ctt aat cta agc tat cgc tat gtt ttc aag gat tct aag gga aaa tta   5684
```

```
Leu Asn Leu Ser Tyr Arg Tyr Val Phe Lys Asp Ser Lys Gly Lys Leu
            855                 860                 865

att aat agc gac gat tta cag aag caa ggt tat tca ctc aca tat att   5732
Ile Asn Ser Asp Asp Leu Gln Lys Gln Gly Tyr Ser Leu Thr Tyr Ile
        870                 875                 880

gat tta tgt act gtt tcc att aaa aaa ggt aat tca aat gaa att gtt   5780
Asp Leu Cys Thr Val Ser Ile Lys Lys Gly Asn Ser Asn Glu Ile Val
    885                 890                 895

aaa tgt aat taattttgtt ttcttgatgt ttgtttcatc atcttctttt           5829
Lys Cys Asn
900

gctcaggtaa ttgaaatgaa taattcgcct ctgcgcgatt ttgtaacttg gtattcaaag 5889

caatcaggcg aatccgttat tgtttctccc gatgtaaaag gtactgttac tgtatattca 5949

tctgacgtta aacctgaaaa tctacgcaat ttctttattt ctgttttacg tgctaataat 6009

tttgatatgg ttggttcaat tccttccata attcagaagt ataatccaaa caatcaggat 6069

tatattgatg aattgccatc atctgataat caggaatatg atgataattc cgctccttct 6129

ggtggtttct ttgttccgca aaatgataat gttactcaaa cttttaaaat taataacgtt 6189

cgggcaaagg atttaatacg agttgtcgaa ttgtttgtaa agtctaatac ttctaaatcc 6249

tcaaatgtat tatctattga cggctctaat ctattagttg tttctgcacc taaagatatt 6309

ttagataacc ttcctcaatt cctttctact gttgatttgc caactgacca gatattgatt 6369

gagggtttga tatttgaggt tcagcaaggt gatgctttag atttttcatt tgctgctggc 6429

tctcagcgtg gcactgttgc aggcggtgtt aatactgacc gcctcacctc tgttttatct 6489

tctgctggtg gttcgttcgg tatttttaat ggcgatgttt tagggctatc agttcgcgca 6549

ttaaagacta atagccattc aaaaatattg tctgtgccac gtattcttac gctttcaggt 6609

cagaagggtt ctatctctgt tggccagaat gtcccttta ttactggtcg tgtgactggt 6669

gaatctgcca atgtaaataa tccatttcag acgattgagc gtcaaaatgt aggtatttcc 6729

atgagcgttt ttcctgttgc aatggctggc ggtaatattg ttctggatat taccagcaag 6789

gccgatagtt tgagttcttc tactcaggca agtgatgtta ttactaatca aagaagtatt 6849

gctacaacgg ttaatttgcg tgatggacag actcttttac tcggtggcct cactgattat 6909

aaaaacactt ctcaagattc tggcgtaccg ttcctgtcta aaatcccttt aatcggcctc 6969

ctgtttagct cccgctctga ttccaacgag gaaagcacgt tatacgtgct cgtcaaagca 7029

accatagtac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag 7089

cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt 7149

tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttagggtt 7209
```

ccgatttagt gctttacggc acctcgaccc caaaaaactt gatttgggtg atggttcacg 7269

tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt 7329

taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg gctattcttt 7389

tgatttataa gggattttgc cgatttcgga accaccatca aacaggattt tcgcctgctg 7449

gggcaaacca gcgtggaccg cttgctgcaa ctctctcagg gccaggcggt gaagggcaat 7509

cagctgttgc ccgtctcact ggtgaaaaga aaaaccaccc tggatccaag cttgcaggtg 7569

gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa 7629

atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 7689

agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg gcattttgcc 7749

ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg 7809

gcgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc 7869

gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt catacactat 7929

tatcccgtat tgacgccggg caagagcaac tcggtcgccg ggcgcggtat tctcagaatg 7989

acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag 8049

aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa 8109

cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc 8169

gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca 8229

cgatgcctgt agcaatgcca acaacgttgc gcaaactatt aactggcgaa ctacttactc 8289

tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc 8349

tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg 8409

ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta 8469

tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag 8529

gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga 8589

ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc 8649

tcatgaccaa aatcccttaa cgtgagtttt cgttccactg tacgtaagac ccccaagctt 8709

gtcgactgaa tggcgaatgg cgctttgcct ggtttccggc accagaagcg gtgccggaaa 8769

gctggctgga gtgcgatctt cctgaggccg atactgtcgt cgtcccctca aactggcaga 8829

tgcacggtta cgatgcgccc atctacacca acgtaaccta tcccattacg gtcaatccgc 8889

cgtttgttcc cacggagaat ccgacgggtt gttactcgct cacatttaat gttgatgaaa 8949

gctggctaca ggaaggccag acgcgaatta tttttgatgg cgttcctatt ggttaaaaaa 9009

tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgt ttacaattta 9069

aatatttgct tatacaatct tcctgttttt ggggcttttc tgattatcaa ccggggtaca 9129

tatgattgac atgctagttt tacgattacc gttcatcgat tctcttgttt gctccagact 9189

ctcaggcaat gacctgatag cctttgtaga tctctcaaaa atagctaccc tctccggcat 9249

gaatttatca gctagaacgg ttgaatatca tattgatggt gatttgactg tctccggcct 9309

ttctcaccct tttgaatctt tacctacaca ttactcaggc attgcattta aaatatatga 9369

gggttctaaa aatttttatc cttgcgttga ataaaggct tctcccgcaa aagtattaca 9429

gggtcataat gtttttggta caaccgattt agctttatgc tctgaggctt tattgcttaa 9489

ttttgctaat tctttgcctt gcctgtatga tttattggat gtt 9532

<210> 452
<211> 20
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência peptídica MALIA3

<400> 452
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
1 5 10 15

His Ser Ala Gln
20

<210> 453
<211> 367
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência proteica

<400> 453
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly
20 25 30

Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
35 40 45

Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln Ala Pro Gly
50 55 60

Lys Gly Leu Glu Trp Val Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr
65 70 75 80

Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
85 90 95

```
Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
            100                 105                 110

Thr Ala Val Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr Ala
        115                 120                 125

Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser Ala Ser
    130                 135                 140

Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr
145                 150                 155                 160

Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro
                165                 170                 175

Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val
            180                 185                 190

His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser
        195                 200                 205

Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile
    210                 215                 220

Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val
225                 230                 235                 240

Glu Pro Lys Ser Cys Ala Ala Ala His His His His His His Ser Ala
                245                 250                 255

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala Ala Asp Ile
            260                 265                 270

Asn Asp Asp Arg Met Ala Gly Ala Ala Glu Thr Val Glu Ser Cys Leu
        275                 280                 285

Ala Lys Pro His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys Asp Asp
    290                 295                 300

Lys Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala
305                 310                 315                 320

Thr Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr
                325                 330                 335

Trp Val Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser
            340                 345                 350

Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr
        355                 360                 365


<210> 454
<211> 152
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência proteica MALIA3

<400> 454
```

```
Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala
  1               5                  10                  15

Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly
             20                  25                  30

Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe
         35                  40                  45

Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp
     50                  55                  60

Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn
 65                  70                  75                  80

Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln
                85                  90                  95

Ser Val Glu Cys Arg Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu
            100                 105                 110

Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala
        115                 120                 125

Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala
    130                 135                 140

Asn Ile Leu Arg Asn Lys Glu Ser
145                 150


<210> 455
<211> 15
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência peptídica MALIA3

<400> 455
Met Pro Val Leu Leu Gly Ile Pro Leu Leu Leu Arg Phe Leu Gly
  1               5                  10                  15


<210> 456
<211> 34B
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: sequência proteica MALIA3

<400> 456
Met Ala Val Tyr Phe Val Thr Gly Lys Leu Gly Ser Gly Lys Thr Leu
  1               5                  10                  15

Val Ser Val Gly Lys Ile Gln Asp Lys Ile Val Ala Gly Cys Lys Ile
             20                  25                  30

Ala Thr Asn Leu Asp Leu Arg Leu Gln Asn Leu Pro Gln Val Gly Arg
         35                  40                  45
```

```
Phe Ala Lys Thr Pro Arg Val Leu Arg Ile Pro Asp Lys Pro Ser Ile
     50                  55                  60

Ser Asp Leu Leu Ala Ile Gly Arg Gly Asn Asp Ser Tyr Asp Glu Asn
 65                  70                  75                  80

Lys Asn Gly Leu Leu Val Leu Asp Glu Cys Gly Thr Trp Phe Asn Thr
                 85                  90                  95

Arg Ser Trp Asn Asp Lys Glu Arg Gln Pro Ile Ile Asp Trp Phe Leu
            100                 105                 110

His Ala Arg Lys Leu Gly Trp Asp Ile Ile Phe Leu Val Gln Asp Leu
        115                 120                 125

Ser Ile Val Asp Lys Gln Ala Arg Ser Ala Leu Ala Glu His Val Val
    130                 135                 140

Tyr Cys Arg Arg Leu Asp Arg Ile Thr Leu Pro Phe Val Gly Thr Leu
145                 150                 155                 160

Tyr Ser Leu Ile Thr Gly Ser Lys Met Pro Leu Pro Lys Leu His Val
                165                 170                 175

Gly Val Val Lys Tyr Gly Asp Ser Gln Leu Ser Pro Thr Val Glu Arg
            180                 185                 190

Trp Leu Tyr Thr Gly Lys Asn Leu Tyr Asn Ala Tyr Asp Thr Lys Gln
        195                 200                 205

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

Tyr Leu Ser His Gly Arg Tyr Phe Lys Pro Leu Asn Leu Gly Gln Lys
225                 230                 235                 240

Met Lys Leu Thr Lys Ile Tyr Leu Lys Lys Phe Ser Arg Val Leu Cys
                245                 250                 255

Leu Ala Ile Gly Phe Ala Ser Ala Phe Thr Tyr Ser Tyr Ile Thr Gln
            260                 265                 270

Pro Lys Pro Glu Val Lys Lys Val Val Ser Gln Thr Tyr Asp Phe Asp
        275                 280                 285

Lys Phe Thr Ile Asp Ser Ser Gln Arg Leu Asn Leu Ser Tyr Arg Tyr
    290                 295                 300

Val Phe Lys Asp Ser Lys Gly Lys Leu Ile Asn Ser Asp Asp Leu Gln
305                 310                 315                 320

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

Lys Lys Gly Asn Ser Asn Glu Ile Val Lys Cys Asn
            340                 345
```

<210> 457
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 457
tggaagaggc acgttctttt cttt 24

<210> 458
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 458
cttttctttg ttgccgttgg ggtg 24

<210> 459
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 459
acactctccc ctgttgaagc tctt 24

<210> 460
<211> 51
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 460
accgcctcca ccgggcgcgc cttattaaca ctctcccctg ttgaagctct t 51

<210> 461
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 461
tgaacattct gtaggggcca ctg 23

<210> 462
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 462
agagcattct gcaggggcca ctg 23

<210> 463
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 463
accgcctcca ccgggcgcgc cttattatga acattctgta ggggccactg 50

<210> 464
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 464
accgcctcca ccgggcgcgc cttattaaga gcattctgca ggggccactg 50

<210> 465
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 465
cgactggagc acgaggacac tga 23

<210> 466
<211> 26
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 466
ggacactgac atggactgaa ggagta 26

<210> 467
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 467
gggaggatgg agactgggtc 20

<210> 468
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 468
gggaagatgg agactgggtc 20

<210> 469
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 469
gggagagtgg agactgagtc 20

<210> 470
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 470
gggtgcctgg agactgcgtc 20

<210> 471
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 471
gggtggctgg agactgcgtc 20

<210> 472
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 472
gggaggatgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 473
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 473
gggaagatgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 474
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 474
gggagagtgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 475
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 475
gggtgcctgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 476
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 476
gggtggctgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 477
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 477
gggagtctgg agactgggtc atctggatgt cttgtgcact gtgacagagg 50

<210> 478
<211> 42
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 478
cctctgtcac agtgcacaag acatccagat gacccagtct cc 42

<210> 479
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 479
cctctgtcac agtgcacaag ac 22

<210> 480
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 480
acactctccc ctgttgaagc tctt 24

<210> 481
<211> 668
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(668)

<400> 481
```
agt gca caa gac atc cag atg acc cag tct cca gcc acc ctg tct gtg   48
Ser Ala Gln Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser Val
  1               5                  10                  15

tct cca ggg gaa agg gcc acc ctc tcc tgc agg gcc agt cag agt gtt   96
Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val
             20                  25                  30

agt aac aac tta gcc tgg tac cag cag aaa cct ggc cag gtt ccc agg   144
Ser Asn Asn Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Val Pro Arg
         35                  40                  45

ctc ctc atc tat ggt gca tcc acc agg gcc act gat atc cca gcc agg   192
Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr Asp Ile Pro Ala Arg
     50                  55                  60

ttc agt ggc agt ggg tct ggg aca gac ttc act ctc acc atc agc aga   240
Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg
 65                  70                  75                  80

ctg gag cct gaa gat ttt gca gtg tat tac tgt cag cgg tat ggt agc   288
Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Arg Tyr Gly Ser
                 85                  90                  95

tca ccg ggg tgg acg ttc ggc caa ggg acc aag gtg gaa atc aaa cga   336
Ser Pro Gly Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
            100                 105                 110

act gtg gct gca cca tct gtc ttc atc ttc ccg cca tct gat gag cag   384
Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
        115                 120                 125

ttg aaa tct gga act gcc tct gtt gtg tgc ctg ctg aat aac ttc tat   432
Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
    130                 135                 140

ccc aga gag gcc aaa gta cag tgg aag gtg gat aac gcc ctc caa tcg   480
Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
145                 150                 155                 160

ggt aac tcc cag gag agt gtc aca gag cag gac agc aag gac agc acc   528
Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
                165                 170                 175

tac agc ctc agc agc acc ctg acg ctg agc aaa gca gac tac gag aaa   576
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
            180                 185                 190

cac aaa gtc tac gcc tgc gaa gtc acc cat cag ggc ctg agc tcg cct   624
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
        195                 200                 205
```

```
gtc aca aag agc ttc aac aaa gga gag tgt aag ggc gaa ttc gc        668
Val Thr Lys Ser Phe Asn Lys Gly Glu Cys Lys Gly Glu Phe Ala
    210                 215                 220


<210> 482
<211> 223
<212> PRT
<213> Homo sapiens

<400> 482
Ser Ala Gln Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser Val
  1               5                  10                  15

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val
             20                  25                  30

Ser Asn Asn Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Val Pro Arg
         35                  40                  45

Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr Asp Ile Pro Ala Arg
     50                  55                  60

Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg
 65                  70                  75                  80

Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Arg Tyr Gly Ser
                 85                  90                  95

Ser Pro Gly Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
            100                 105                 110

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
        115                 120                 125

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
    130                 135                 140

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
145                 150                 155                 160

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
                165                 170                 175

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
            180                 185                 190

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
        195                 200                 205

Val Thr Lys Ser Phe Asn Lys Gly Glu Cys Lys Gly Glu Phe Ala
    210                 215                 220


<210> 483
<211> 13
<212> ADN
<213> Sequência Artificial
```

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 483
agccaccctg tct 13

<210> 484
<211> 700
<212> ADN
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(699)

<400> 484
agt gca caa gac atc cag atg acc cag tct cct gcc acc ctg tct gtg 48
Ser Ala Gln Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser Val
1 5 10 15

tct cca ggt gaa aga gcc acc ctc tcc tgc agg gcc agt cag gtg tct 96
Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Val Ser
20 25 30

cca ggg gaa aga gcc acc ctc tcc tgc aat ctt ctc agc aac tta gcc 144
Pro Gly Glu Arg Ala Thr Leu Ser Cys Asn Leu Leu Ser Asn Leu Ala
35 40 45

tgg tac cag cag aaa cct ggc cag gct ccc agg ctc ctc atc tat ggt 192
Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly
50 55 60

gct tcc acc ggg gcc att ggt atc cca gcc agg ttc agt ggc agt ggg 240
Ala Ser Thr Gly Ala Ile Gly Ile Pro Ala Arg Phe Ser Gly Ser Gly
65 70 75 80

tct ggg aca gag ttc act ctc acc atc agc agc ctg cag tct gaa gat 288
Ser Gly Thr Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Ser Glu Asp
85 90 95

ttt gca gtg tat ttc tgt cag cag tat ggt acc tca ccg ccc act ttc 336
Phe Ala Val Tyr Phe Cys Gln Gln Tyr Gly Thr Ser Pro Pro Thr Phe
100 105 110

ggc gga ggg acc aag gtg gag atc aaa cga act gtg gct gca cca tct 384
Gly Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser
115 120 125

gtc ttc atc ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc 432
Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala
130 135 140

tct gtt gtg tgc ccg ctg aat aac ttc tat ccc aga gag gcc aaa gta 480
Ser Val Val Cys Pro Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val
145 150 155 160

cag tgg aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt 528
Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser
165 170 175

```
gtc aca gag cag gac aac aag gac agc acc tac agc ctc agc agc acc   576
Val Thr Glu Gln Asp Asn Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr
            180                 185                 190

ctg acg ctg agc aaa gta gac tac gag aaa cac gaa gtc tac gcc tgc   624
Leu Thr Leu Ser Lys Val Asp Tyr Glu Lys His Glu Val Tyr Ala Cys
        195                 200                 205

gaa gtc acc cat cag ggc ctt agc tcg ccc gtc acg aag agc ttc aac   672
Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn
    210                 215                 220

agg gga gag tgt aag aaa gaa ttc gtt t                             700
Arg Gly Glu Cys Lys Lys Glu Phe Val
225                 230
```

<210> 485
<211> 233
<212> PRT
<213> Homo sapiens

<400> 485

```
Ser Ala Gln Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser Val
  1               5                  10                  15

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Val Ser
             20                  25                  30

Pro Gly Glu Arg Ala Thr Leu Ser Cys Asn Leu Leu Ser Asn Leu Ala
         35                  40                  45

Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly
     50                  55                  60

Ala Ser Thr Gly Ala Ile Gly Ile Pro Ala Arg Phe Ser Gly Ser Gly
 65                  70                  75                  80

Ser Gly Thr Glu Phe Thr Leu Thr Ile Ser Ser Leu Gln Ser Glu Asp
                 85                  90                  95

Phe Ala Val Tyr Phe Cys Gln Gln Tyr Gly Thr Ser Pro Pro Thr Phe
            100                 105                 110

Gly Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser
        115                 120                 125

Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala
    130                 135                 140

Ser Val Val Cys Pro Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val
145                 150                 155                 160

Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser
                165                 170                 175

Val Thr Glu Gln Asp Asn Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr
            180                 185                 190

Leu Thr Leu Ser Lys Val Asp Tyr Glu Lys His Glu Val Tyr Ala Cys
        195                 200                 205
```

```
Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn
    210                 215                 220

Arg Gly Glu Cys Lys Lys Glu Phe Val
225                 230
```

<210> 486
<211> 419
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sintética 3-23
VH Sequência nucleotídica

<220>
<221> CDS
<222> (12)..(419)

<400> 486

```
ctgtctgaac g gcc cag ccg gcc atg gcc gaa gtt caa ttg tta gag tct  50
             Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser
               1               5                  10

ggt ggc ggt ctt gtt cag cct ggt ggt tct tta cgt ctt tct tgc gct   98
Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala
     15                  20                  25

gct tcc gga ttc act ttc tct tcg tac gct atg tct tgg gtt cgc caa   146
Ala Ser Gly Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln
 30                  35                  40                  45

gct cct ggt aaa ggt ttg gag tgg gtt tct gct atc tct ggt tct ggt   194
Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Ala Ile Ser Gly Ser Gly
                 50                  55                  60

ggc agt act tac tat gct gac tcc gtt aaa ggt cgc ttc act atc tct   242
Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
             65                  70                  75

aga gac aac tct aag aat act ctc tac ttg cag atg aac agc tta agg   290
Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
         80                  85                  90

gct gag gac act gca gtc tac tat tgc gct aaa gac tat gaa ggt act   338
Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr
     95                  100                 105

ggt tat gct ttc gac ata tgg ggt caa ggt act atg gtc acc gtc tct   386
Gly Tyr Ala Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser
110                 115                 120                 125

agt gcc tcc acc aag ggc cca tcg gtc ttc ccc                      419
Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro
                130                 135
```

<210> 487
<211> 136
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sintética 3-23
VH Sequência proteica

<400> 487
Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly
1 5 10 15

Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
20 25 30

Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln Ala Pro Gly
35 40 45

Lys Gly Leu Glu Trp Val Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr
50 55 60

Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
65 70 75 80

Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
85 90 95

Thr Ala Val Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr Ala
100 105 110

Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser Ala Ser
115 120 125

Thr Lys Gly Pro Ser Val Phe Pro
130 135

<210> 488
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 488
ctgtctgaac ggcccagccg 20

<210> 489
<211> 83
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 489
ctgtctgaac ggcccagccg gccatggccg aagttcaatt gttagagtct ggtggcggtc 60
ttgttcagcc tggtggttct tta 83

<210> 490
<211> 54
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 490
gaaagtgaat ccggaagcag cgcaagaaag acgtaaagaa ccaccaggct gaac 54

<210> 491
<211> 42
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 491
agaaacccac tccaaacctt taccaggagc ttggcgaacc ca 42

<210> 492
<211> 94
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 492
agtgtcctca gcccttaagc tgttcatctg caagtagaga gtattcttag agttgtctct 60
agagatagtg aagcgacctt taacggagtc agca 94

<210> 493
<211> 81
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 493
gcttaagggc tgaggacact gcagtctact attgcgctaa agactatgaa ggtactggtt 60
atgctttcga catatggggt c 81

<210> 494
<211> 72
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 494
ggggaagacc gatgggccct tggtggaggc actagagacg gtgaccatag taccttgacc 60
tatgtcgaaa gc 72

<210> 495
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 495
ggggaagacc gatgggccct tgg 23

<210> 496
<211> 56
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (22)..(24)
<223> A, T, C, G, outro ou desconhecido

<220>
<221> base_modificada
<222> (28)..(30)
<223> A, T, C, G, outro ou desconhecido

<220>
<221> base_modificada
<222> (34)..(36)
<223> A, T, C, G, outro ou desconhecido

<220>
<223> nnn código para qualquer aminoácido mas Cys

<400> 496
gcttccggat tcactttctc tnnntacnnn atgnnntggg ttcgccaagc tcctgg 56

<210> 497
<211> 68
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (19)..(21)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (25)..(30)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (40)..(42)
<223> A, T, C or G

<220>
<221> base_modificada
<222> (46)..(48)
<223> A, T, C or G

<400> 497
ggtttggagt gggtttctnn natcnnnnnn tctggtggcn nnactnnnta tgctgactcc 60
gttaaagg 68

<210> 498
<211> 912
<212> ADN
<213> Escherichia coli

<400> 498
tccggagctt cagatctgtt tgccttttttg tggggtggtg cagatcgcgt tacggagatc 60
gaccgactgc ttgagcaaaa gccacgctta actgctgatc aggcatggga tgttattcgc 120
caaaccagtc gtcaggatct taacctgagg ctttttttac ctactctgca agcagcgaca 180
tctggtttga cacagagcga tccgcgtcgt cagttggtag aaacattaac acgttgggat 240
ggcatcaatt tgcttaatga tgatggtaaa acctggcagc agccaggctc tgccatcctg 300
aacgtttggc tgaccagtat gttgaagcgt accgtagtgg ctgccgtacc tatgccattt 360
gataagtggt acagcgccag tggctacgaa acaacccagg acggcccaac tggttcgctg 420
aatataagtg ttggagcaaa aattttgtat gaggcggtgc agggagacaa atcaccaatc 480
ccacaggcgg ttgatctgtt tgctgggaaa ccacagcagg aggttgtgtt ggctgcgctg 540
gaagatacct gggagactct ttccaaacgc tatggcaata atgtgagtaa ctggaaaaca 600
cctgcaatgg ccttaacgtt ccgggcaaat aatttctttg gtgtaccgca ggccgcagcg 660
gaagaaacgc gtcatcaggc ggagtatcaa aaccgtggaa cagaaaacga tatgattgtt 720
ttctcaccaa cgacaagcga tcgtcctgtg cttgcctggg atgtggtcgc acccggtcag 780
agtgggttta ttgctcccga tggaacagtt gataagcact atgaagatca gctgaaaatg 840
tacgaaaatt ttggccgtaa gtcgctctgg ttaacgaagc aggatgtgga ggcgcataag 900
gagtcgtcta ga 912

<210> 499
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 499
gatnnnnatc 10

<210> 500
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(15)
<223> A, T, C, G, outro ou desconhecido

<400> 500
nnnnnnnnnn nnnnngtccc 20

<210> 501
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 501
gcannnnntg c 11

<210> 502
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (9)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 502
gacnnnngtc 10

<210> 503
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 503
nnnnnnngcg gg 12

<210> 504
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 504
gtatccnnnn nn 12

<210> 505
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 505
gcannnnnnt cg 12

<210> 506
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 506
gccnnnnngg c 11

<210> 507
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 507
ggtctcnnnn n 11

<210> 508
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 508
gacnnnnngt c 11

<210> 509
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 509
gacnnnnngt c 11

<210> 510
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 510
gacnnnnnng tc 12

<210> 511
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 511
ccannnnntg g 11

<210> 512
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 512
nnnnnnnnng caggt 15

<210> 513
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 513
acctgcnnnn n 11

<210> 514
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (5)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 514
ggccnnnnng gcc 13

<210> 515
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 515
ccannnnnnn nntgg 15

<210> 516
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 516
cgtctcnnnn n 11

<210> 517
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(6)
<223> A, T, C, G, outro ou desconhecido

<400> 517
nnnnnngaga cg 12

<210> 518
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 518
nnnnnnnnnn ctcctc 16

<210> 519
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(16)
<223> A, T, C, G, outro ou desconhecido

<400> 519
gaggagnnnn nnnnnn 16

<210> 520
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 520
cctnnnnnag g 11

<210> 521
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 521
ccannnnnnt gg 12

<210> 522
<211> 6680
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência nucleotídica do Vector pCES5

<220>
<221> CDS
<222> (201)..(1058)

<220>
<221> CDS
<222> (2269)..(2682)

<220>
<221> CDS
<222> (2723)..(2866)

<220>
<221> CDS
<222> (3767)..(3850)

<220>
<221> CDS
<222> (4198)..(5799)

<400> 522
gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt 60

cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt 120

tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 180

aatattgaaa aaggaagagt atg agt att caa cat ttc cgt gtc gcc ctt att 233
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile
1 5 10

ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca gaa acg 281
Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr
15 20 25

ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gcc cga gtg ggt 329
Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly
30 35 40

tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt ttt cgc 377
Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg
45 50 55

ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt 425
Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys
60 65 70 75

ggc gcg gta tta tcc cgt att gac gcc ggg caa gag caa ctc ggt cgc 473
Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg
80 85 90

cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca gtc aca 521
Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr
95 100 105

gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc agt gct 569
Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala
110 115 120

gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg aca acg 617
Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr
125 130 135

```
atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg ggg gat   665
Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp
140                 145                 150                 155

cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc ata   713
His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile
                160                 165                 170

cca aac gac gag cgt gac acc acg atg cct gta gca atg gca aca acg   761
Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr
            175                 180                 185

ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa   809
Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln
        190                 195                 200

caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg   857
Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu
    205                 210                 215

cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct gga gcc   905
Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala
220                 225                 230                 235

ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca gat ggt   953
Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly
                240                 245                 250

aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag gca act   1001
Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr
            255                 260                 265

atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca ctg att   1049
Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile
        270                 275                 280

aag cat tgg taactgtcag accaagttta ctcatatata ctttagattg           1098
Lys His Trp
    285

atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca 1158

tgaccaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga 1218

tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa 1278

aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga 1338

aggtaactgg cttcagcaga gcgcagatac caaatactgt ccttctagtg tagccgtagt 1398

taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt 1458

taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat 1518

agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcata cagcccagct 1578

tggagcgaac gacctacacc gaactgagat acctacagcg tgagcattga gaaagcgcca 1638

cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag 1698

agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc 1758
```

```
gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcgg agcctatgga 1818

aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 1878

tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag 1938

ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg 1998

aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatgcagct 2058

ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt 2118

agctcactca ttaggcaccc caggctttac actttatgct tccggctcgt atgttgtgtg 2178

gaattgtgag cggataacaa tttcacacag gaaacagcta tgaccatgat tacgccaagc 2238

tttggagcct ttttttgga gattttcaac gtg aaa aaa tta tta ttc gca att  2292
                              Met Lys Lys Leu Leu Phe Ala Ile
                                              290

cct tta gtt gtt cct ttc tat tct cac agt gca cag gtc caa ctg cag   2340
Pro Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln Val Gln Leu Gln
295                 300                 305                 310

gtc gac ctc gag atc aaa cgt gga act gtg gct gca cca tct gtc ttc   2388
Val Asp Leu Glu Ile Lys Arg Gly Thr Val Ala Ala Pro Ser Val Phe
                315                 320                 325

atc ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt   2436
Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val
            330                 335                 340

gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag tgg   2484
Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp
        345                 350                 355

aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc aca   2532
Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr
    360                 365                 370

gag cag gac agc aag gac agc acc tac agc ctc agc agc acc ctg acg   2580
Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr
375                 380                 385                 390

ctg agc aaa gca gac tac gag aaa cac aaa gtc tac gcc tgc gaa gtc   2628
Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val
                395                 400                 405

acc cat cag ggc ctg agt tca ccg gtg aca aag agc ttc aac agg gga   2676
Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly
            410                 415                 420

gag tgt taataaggcg cgccaattct atttcaagga gacagtcata atg aaa tac   2731
Glu Cys                                             Met Lys Tyr
                                                    425

cta ttg cct acg gca gcc gct gga ttg tta tta ctc gcg gcc cag ccg   2779
Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala Ala Gln Pro
        430                 435                 440
```

```
gcc atg gcc gaa gtt caa ttg tta gag tct ggt ggc ggt ctt gtt cag   2827
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
    445                 450                 455

cct ggt ggt tct tta cgt ctt tct tgc gct gct tcc gga gcttcagatc    2876
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
460                 465                 470

tgtttgcctt tttgtggggt ggtgcagatc gcgttacgga gatcgaccga ctgcttgagc 2936

aaaagccacg cttaactgct gatcaggcat gggatgttat tcgccaaacc agtcgtcagg 2996

atcttaacct gaggcttttt ttacctactc tgcaagcagc gacatctggt ttgacacaga 3056

gcgatccgcg tcgtcagttg gtagaaacat taacacgttg ggatggcatc aatttgctta 3116

atgatgatgg taaaacctgg cagcagccag gctctgccat cctgaacgtt tggctgacca 3176

gtatgttgaa gcgtaccgta gtggctgccg tacctatgcc atttgataag tggtacagcg 3236

ccagtggcta cgaaacaacc caggacggcc caactggttc gctgaatata agtgttggag 3296

caaaaatttt gtatgaggcg gtgcagggag acaaatcacc aatcccacag gcggttgatc 3356

tgtttgctgg gaaaccacag caggaggttg tgttggctgc gctggaagat acctgggaga 3416

ctctttccaa acgctatggc aataatgtga gtaactggaa aacacctgca atggccttaa 3476

cgttccgggc aaataatttc tttggtgtac cgcaggccgc agcggaagaa acgcgtcatc 3536

aggcggagta tcaaaaccgt ggaacagaaa acgatatgat tgttttctca ccaacgacaa 3596

gcgatcgtcc tgtgcttgcc tgggatgtgg tcgcacccgg tcagagtggg tttattgctc 3656

ccgatggaac agttgataag cactatgaag atcagctgaa aatgtacgaa aattttggcc 3716

gtaagtcgct ctggttaacg aagcaggatg tggaggcgca taaggagtcg tct aga     3772
                                                      Ser Arg

gac aac tct aag aat act ctc tac ttg cag atg aac agc tta agt ctg    3820
Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Ser Leu
475                 480                 485                 490

agc att cgg tcc ggg caa cat tct cca aac tgaccagacg acacaaacgg      3870
Ser Ile Arg Ser Gly Gln His Ser Pro Asn
                495                 500

cttacgctaa atcccgcgca tgggatggta aagaggtggc gtctttgctg gcctggactc 3930

atcagatgaa ggccaaaaat tggcaggagt ggacacagca ggcagcgaaa caagcactga 3990

ccatcaactg gtactatgct gatgtaaacg gcaatattgg ttatgttcat actggtgctt 4050

atccagatcg tcaatcaggc catgatccgc gattacccgt tcctggtacg ggaaaatggg 4110

actggaaagg gctattgcct tttgaaatga accctaaggt gtataacccc cagaagctag 4170

cctgcggctt cggtcaccgt ctcaagc gcc tcc acc aag ggc cca tcg gtc ttc 4224
                            Ala Ser Thr Lys Gly Pro Ser Val Phe
                                                505
```

```
ccc ctg gca ccc tcc tcc aag agc acc tct ggg ggc aca gcg gcc ctg   4272
Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu
510                 515                 520                 525

ggc tgc ctg gtc aag gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg   4320
Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
                530                 535                 540

aac tca ggc gcc ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta   4368
Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu
            545                 550                 555

cag tcc tca gga ctc tac tcc ctc agc agc gta gtg acc gtg ccc tcc   4416
Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser
        560                 565                 570

agc agc ttg ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc   4464
Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro
    575                 580                 585

agc aac acc aag gtg gac aag aaa gtt gag ccc aaa tct tgt gcg gcc   4512
Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Ala Ala
590                 595                 600                 605

gca cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc atc tca   4560
Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu Ile Ser
                610                 615                 620

gaa gag gat ctg aat ggg gcc gca tag act gtt gaa agt tgt tta gca   4608
Glu Glu Asp Leu Asn Gly Ala Ala     Thr Val Glu Ser Cys Leu Ala
            625                 630                 635

aaa cct cat aca gaa aat tca ttt act aac gtc tgg aaa gac gac aaa   4656
Lys Pro His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys Asp Asp Lys
            640                 645                 650

act tta gat cgt tac gct aac tat gag ggc tgt ctg tgg aat gct aca   4704
Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala Thr
        655                 660                 665

ggc gtt gtg gtt tgt act ggt gac gaa act cag tgt tac ggt aca tgg   4752
Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr Trp
    670                 675                 680

gtt cct att ggg ctt gct atc cct gaa aat gag ggt ggt ggc tct gag   4800
Val Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser Glu
685                 690                 695                 700

ggt ggc ggt tct gag ggt ggc ggt tct gag ggt ggc ggt act aaa cct   4848
Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr Lys Pro
                705                 710                 715

cct gag tac ggt gat aca cct att ccg ggc tat act tat atc aac cct   4896
Pro Glu Tyr Gly Asp Thr Pro Ile Pro Gly Tyr Thr Tyr Ile Asn Pro
            720                 725                 730

ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct aat cct   4944
Leu Asp Gly Thr Tyr Pro Pro Gly Thr Glu Gln Asn Pro Ala Asn Pro
        735                 740                 745

aat cct tct ctt gag gag tct cag cct ctt aat act ttc atg ttt cag   4992
```

```
Asn Pro Ser Leu Glu Glu Ser Gln Pro Leu Asn Thr Phe Met Phe Gln
    750                 755                 760

aat aat agg ttc cga aat agg cag ggt gca tta act gtt tat acg ggc   5040
Asn Asn Arg Phe Arg Asn Arg Gln Gly Ala Leu Thr Val Tyr Thr Gly
765                 770                 775                 780

act gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act   5088
Thr Val Thr Gln Gly Thr Asp Pro Val Lys Thr Tyr Tyr Gln Tyr Thr
                785                 790                 795

cct gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttc   5136
Pro Val Ser Ser Lys Ala Met Tyr Asp Ala Tyr Trp Asn Gly Lys Phe
            800                 805                 810

aga gac tgc gct ttc cat tct ggc ttt aat gag gat cca ttc gtt tgt   5184
Arg Asp Cys Ala Phe His Ser Gly Phe Asn Glu Asp Pro Phe Val Cys
        815                 820                 825

gaa tat caa ggc caa tcg tct gac ctg cct caa cct cct gtc aat gct   5232
Glu Tyr Gln Gly Gln Ser Ser Asp Leu Pro Gln Pro Pro Val Asn Ala
    830                 835                 840

ggc ggc ggc tct ggt ggt ggt tct ggt ggc ggc tct gag ggt ggc ggc   5280
Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Glu Gly Gly Gly
845                 850                 855                 860

tct gag ggt ggc ggt tct gag ggt ggc ggc tct gag ggt ggc ggt tcc   5328
Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser
                865                 870                 875

ggt ggc ggc tcc ggt tcc ggt gat ttt gat tat gaa aaa atg gca aac   5376
Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn
            880                 885                 890

gct aat aag ggg gct atg acc gaa aat gcc gat gaa aac gcg cta cag   5424
Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln
        895                 900                 905

tct gac gct aaa ggc aaa ctt gat tct gtc gct act gat tac ggt gct   5472
Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala
    910                 915                 920

gct atc gat ggt ttc att ggt gac gtt tcc ggc ctt gct aat ggt aat   5520
Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn
925                 930                 935                 940

ggt gct act ggt gat ttt gct ggc tct aat tcc caa atg gct caa gtc   5568
Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val
                945                 950                 955

ggt gac ggt gat aat tca cct tta atg aat aat ttc cgt caa tat tta   5616
Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu
            960                 965                 970

cct tct ttg cct cag tcg gtt gaa tgt cgc cct tat gtc ttt ggc gct   5664
Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro Tyr Val Phe Gly Ala
        975                 980                 985
```

```
ggt aaa cca tat gaa ttt tct att gat tgt gac aaa ata aac tta ttc   5712
Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe
    990                 995                1000

cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat gta   5760
Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val
1005              1010                1015                1020

ttt tcg acg ttt gct aac ata ctg cgt aat aag gag tct taataagaat    5809
Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
               1025                1030

tcactggccg tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat 5869

cgccttgcag cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat 5929

cgcccttccc aacagttgcg cagcctgaat ggcgaatggc gcctgatgcg gtattttctc 5989

cttacgcatc tgtgcggtat ttcacaccgc atataaattg taaacgttaa tattttgtta 6049

aaattcgcgt taaatttttg ttaaatcagc tcatttttta accaataggc cgaaatcggc 6109

aaaatccctt ataaatcaaa agaatagccc gagatagggt tgagtgttgt tccagtttgg 6169

aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat 6229

cagggcgatg gcccactacg tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc 6289

cgtaaagcac taaatcggaa ccctaaaggg agcccccgat ttagagcttg acggggaaag 6349

ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg 6409

gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta 6469

cagggcgcgt actatggttg ctttgacggg tgcagtctca gtacaatctg ctctgatgcc 6529

gcatagttaa gccagccccg acacccgcca acacccgctg acgcgccctg acgggcttgt 6589

ctgctcccgg catccgctta cagacaagct gtgaccgtct ccgggagctg catgtgtcag 6649

aggttttcac cgtcatcacc gaaacgcgcg a                                  6680


<210> 523
<211> 286
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica do Vector pCES5


<400> 523
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
  1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
             20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
         35                  40                  45
```

```
Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
     50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
 65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275                 280                 285


<210> 524
<211> 138
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica do Vector pCES5


<400> 524
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
  1               5                  10                  15
```

```
His Ser Ala Gln Val Gln Leu Gln Val Asp Leu Glu Ile Lys Arg Gly
             20                  25                  30

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
         35                  40                  45

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
     50                  55                  60

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
 65                  70                  75                  80

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
                 85                  90                  95

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
            100                 105                 110

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
        115                 120                 125

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
    130                 135
```

<210> 525
<211> 48
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica do Vector pCES5

```
<400> 525
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
  1               5                  10                  15

Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly
             20                  25                  30

Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
         35                  40                  45
```

<210> 526
<211> 28
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica do Vector pCES5

```
<400> 526
Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu
  1               5                  10                  15

Ser Leu Ser Ile Arg Ser Gly Gln His Ser Pro Asn
             20                  25
```

<210> 527
<211> 533
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica do Vector pCES5

<400> 527
Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val Glu Pro Lys Ser Cys Ala Ala Ala His His His His His His
100 105 110

Gly Ala Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala
115 120 125

Ala Thr Val Glu Ser Cys Leu Ala Lys Pro His Thr Glu Asn Ser Phe
130 135 140

Thr Asn Val Trp Lys Asp Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr
145 150 155 160

Glu Gly Cys Leu Trp Asn Ala Thr Gly Val Val Val Cys Thr Gly Asp
165 170 175

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

Glu Asn Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
195 200 205

Ser Glu Gly Gly Gly Thr Lys Pro Pro Glu Tyr Gly Asp Thr Pro Ile
210 215 220

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

Thr Glu Gln Asn Pro Ala Asn Pro Asn Pro Ser Leu Glu Glu Ser Gln
245 250 255

```
Pro Leu Asn Thr Phe Met Phe Gln Asn Asn Arg Phe Arg Asn Arg Gln
            260                 265                 270

Gly Ala Leu Thr Val Tyr Thr Gly Thr Val Thr Gln Gly Thr Asp Pro
        275                 280                 285

Val Lys Thr Tyr Tyr Gln Tyr Thr Pro Val Ser Ser Lys Ala Met Tyr
    290                 295                 300

Asp Ala Tyr Trp Asn Gly Lys Phe Arg Asp Cys Ala Phe His Ser Gly
305                 310                 315                 320

Phe Asn Glu Asp Pro Phe Val Cys Glu Tyr Gln Gly Gln Ser Ser Asp
                325                 330                 335

Leu Pro Gln Pro Pro Val Asn Ala Gly Gly Gly Ser Gly Gly Gly Ser
            340                 345                 350

Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly
        355                 360                 365

Gly Gly Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp
    370                 375                 380

Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu
385                 390                 395                 400

Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp
                405                 410                 415

Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp
            420                 425                 430

Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly
        435                 440                 445

Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu
    450                 455                 460

Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu
465                 470                 475                 480

Cys Arg Pro Tyr Val Phe Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile
                485                 490                 495

Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu
            500                 505                 510

Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu
        515                 520                 525

Arg Asn Lys Glu Ser
    530
```

<210> 528
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 528
acctcactgg cttccggatt cactttctct 30

<210> 529
<211> 42
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 529
agaaacccac tccaaacctt taccaggagc ttggcgaacc ca 42

<210> 530
<211> 51
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 530
ggaaggcagt gatctagaga tagtgaagcg acctttaacg gagtcagcat a 51

<210> 531
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 531
ggaaggcagt gatctagaga tag 23

<210> 532
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 532
gtgctgactc agccaccctc 20

<210> 533
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 533
gccctgactc agcctgcctc 20

<210> 534
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 534
gagctgactc aggaccctgc 20

<210> 535
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 535
gagctgactc agccaccctc 20

<210> 536
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 536
cctcgacagc gaagtgcaca gagcgtcttg actcagcc 38

<210> 537
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 537
cctcgacagc gaagtgcaca gagcgtcttg 30

<210> 538
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 538
cctcgacagc gaagtgcaca gagcgctttg actcagcc 38

<210> 539
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 539
cctcgacagc gaagtgcaca gagcgctttg 30

<210> 540
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 540
cctcgacagc taagtgcaca gagcgctttg actcagcc 38

<210> 541
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 541
cctcgacagc gaagtgcaca gagcgctttg 30

<210> 542
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 542
cctcgacagc gaagtgcaca gagcgaattg actcagcc 38

<210> 543
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 543
cctcgacagc gaagtgcaca gagcgaattg 30

<210> 544
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 544
cctcgacagc gaagtgcaca gtacgaattg actcagcc 38

<210> 545
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 545
cctcgacagc gaagtgcaca gtacgaattg 30

<210> 546
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 546
cctcgacagc gaagtgcaca g 21

<210> 547
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 547
ccgtgtatta ctgtgcgaga g 21

<210> 548
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 548
ctgtgtatta ctgtgcgaga g 21

<210> 549
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 549
ccgtatatta ctgtgcgaaa g 21

<210> 550
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 550
ctgtgtatta ctgtgcgaaa g 21

<210> 551
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 551
ctgtgtatta ctgtgcgaga c 21

<210> 552
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 552
ccatgtatta ctgtgcgaga c 21

<210> 553
<211> 94
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 553
ggtgtagtga tctagtgaca actctaagaa tactctctac ttgcagatga acagctttag 60
ggctgaggac actgcagtct actattgtgc gaga 94

<210> 554
<211> 94
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 554
ggtgtagtga tctagtgaca actctaagaa tactctctac ttgcagatga acagctttag 60
ggctgaggac actgcagtct actattgtgc gaaa 94

<210> 555
<211> 85
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 555
atagtagact gcagtgtcct cagcccttaa gctgttcatc tgcaagtaga gagtattctt 60
agagttgtct ctagatcact acacc 85

<210> 556
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 556
gactgggtgt agtgatctag 20

<210> 557
<211> 24
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 557
cttttctttg ttgccgttgg ggtg 24

<210> 558
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 558
nnnnnnnnng caggt 15

<210> 559
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 559
acctgcnnnn n 11

<210> 560
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 560
gatnnnnatc 10

<210> 561
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(16)
<223> A, T, C, G, outro ou desconhecido

<400> 561
gaggagnnnn nnnnnn 16

<210> 562
<211> 16
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 562
nnnnnnnnnn ctcctc 16

<210> 563
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 563
ctcttcnnnn 10

<210> 564
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(5)
<223> A, T, C, G, outro ou desconhecido

<400> 564
nnnnngaaga g 11

<210> 565
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (1)..(15)
<223> A, T, C, G, outro ou desconhecido

<400> 565
nnnnnnnnnn nnnnngtccc 20

<210> 566
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 566
gacnnnnnng tc 12

<210> 567
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 567
cgtctcnnnn n 11

<210> 568
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 568
gtatccnnnn nn 12

<210> 569
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 569
gcannnnnnt cg 12

<210> 570
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 570
gccnnnnngg c 11

<210> 571
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(11)
<223> A, T, C, G, outro ou desconhecido

<400> 571
ggtctcnnnn n 11

<210> 572
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 572
gacnnnnngt c 11

<210> 573
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 573
gacnnnnngt c 11

<210> 574
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 574
ccannnnntg g 11

<210> 575
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 575
ccannnnnnn nntgg 15

<210> 576
<211> 13
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (5)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 576
ggccnnnnng gcc 13

<210> 577
<211> 12
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(9)
<223> A, T, C, G, outro ou desconhecido

<400> 577
ccannnnnnt gg 12

<210> 578
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 578
cctnnnnnag g 11

<210> 579
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 579
gacnnnngtc 10

<210> 580
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(12)
<223> A, T, C, G, outro ou desconhecido

<400> 580
ccannnnnnn nntgg 15

<210> 581
<211> 11
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(8)
<223> A, T, C, G, outro ou desconhecido

<400> 581
gcannnnntg c 11

<210> 582
<211> 10251
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência nucleotídica CJRA05

<220>
<221> CDS
<222> (1578)..(1916)

<220>
<221> CDS
<222> (2388)..(2843)

<220>
<221> CDS
<222> (2849)..(2893)

<220>
<221> CDS
<222> (3189)..(4232)

<220>
<221> CDS
<222> (7418)..(8119)

<220>
<221> CDS
<222> (8160)..(9452)

<400> 582
aatgctacta ctattagtag aattgatgcc acctttttcag ctcgcgcccc aaatgaaaat 60

atagctaaac aggttattga ccatttgcga aatgtatcta atggtcaaac taaatctact 120

cgttcgcaga attgggaatc aactgttata tggaatgaaa cttccagaca ccgtacttta 180

gttgcatatt taaaacatgt tgagctacag cattatattc agcaattaag ctctaagcca 240

tccgcaaaaa tgacctctta tcaaaaggag caattaaagg tactctctaa tcctgacctg 300

ttggagtttg cttccggtct ggttcgcttt gaagctcgaa ttaaaacgcg atatttgaag 360

tctttcgggc ttcctcttaa tctttttgat gcaatccgct ttgcttctga ctataatagt 420

cagggtaaag acctgatttt tgatttatgg tcattctcgt tttctgaact gtttaaagca 480

tttgaggggg attcaatgaa tatttatgac gattccgcag tattggacgc tatccagtct 540

aaacatttta ctattacccc ctctggcaaa acttcttttg caaaagcctc tcgctatttt 600

ggtttttatc gtcgtctggt aaacgagggt tatgatagtg ttgctcttac tatgcctcgt 660

aattccttttt ggcgttatgt atctgcatta gttgaatgtg gtattcctaa atctcaactg 720

atgaatcttt ctacctgtaa taatgttgtt ccgttagttc gttttattaa cgtagatttt 780

tcttcccaac gtcctgactg gtataatgag ccagttctta aaatcgcata aggtaattca 840

caatgattaa agttgaaatt aaaccatctc aagcccaatt tactactcgt tctggtgttt 900

ctcgtcaggg caagccttat tcactgaatg agcagctttg ttacgttgat ttgggtaatg 960

aatatccggt tcttgtcaag attactcttg atgaaggtca gccagcctat gcgcctggtc 1020

tgtacaccgt tcatctgtcc tctttcaaag ttggtcagtt cggttccctt atgattgacc 1080

gtctgcgcct cgttccggct aagtaacatg gagcaggtcg cggatttcga cacaatttat 1140

caggcgatga tacaaatctc cgttgtactt tgtttcgcgc ttggtataat cgctgggggt 1200

caaagatgag tgttttagtg tattcttttg cctctttcgt tttaggttgg tgccttcgta 1260

gtggcattac gtattttacc cgtttaatgg aaacttcctc atgaaaaagt ctttagtcct 1320

caaagcctct gtagccgttg ctaccctcgt tccgatgctg tctttcgctg ctgagggtga 1380

cgatcccgca aaagcggcct ttaactccct gcaagcctca gcgaccgaat atatcggtta 1440

tgcgtgggcg atggttgttg tcattgtcgg cgcaactatc ggtatcaagc tgtttaagaa 1500

attcacctcg aaagcaagct gataaaccga tacaattaaa ggctcctttt ggagcctttt 1560

```
ttttggagat tttcaac gtg aaa aaa tta tta ttc gca att cct tta gtt    1610
                  Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val
                    1               5                  10

gtt cct ttc tat tct ggc gcg gcc gaa tca cat cta gac ggc gcc gct   1658
Val Pro Phe Tyr Ser Gly Ala Ala Glu Ser His Leu Asp Gly Ala Ala
             15                  20                  25

gaa act gtt gaa agt tgt tta gca aaa tcc cat aca gaa aat tca ttt   1706
Glu Thr Val Glu Ser Cys Leu Ala Lys Ser His Thr Glu Asn Ser Phe
         30                  35                  40

act aac gtc tgg aaa gac gac aaa act tta gat cgt tac gct aac tat   1754
Thr Asn Val Trp Lys Asp Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr
     45                  50                  55

gag ggc tgt ctg tgg aat gct aca ggc gtt gta gtt tgt act ggt gac   1802
Glu Gly Cys Leu Trp Asn Ala Thr Gly Val Val Val Cys Thr Gly Asp
 60                  65                  70                  75

gaa act cag tgt tac ggt aca tgg gtt cct att ggg ctt gct atc cct   1850
Glu Thr Gln Cys Tyr Gly Thr Trp Val Pro Ile Gly Leu Ala Ile Pro
                 80                  85                  90

gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct gag ggt ggc ggt   1898
Glu Asn Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
             95                 100                 105

tct gag ggt ggc ggt act aaacctcctg agtacggtga tacacctatt          1946
Ser Glu Gly Gly Gly Thr
        110

ccgggctata cttatatcaa ccctctcgac ggcacttatc cgcctggtac tgagcaaaac 2006

cccgctaatc ctaatccttc tcttgaggag tctcagcctc ttaatacttt catgtttcag 2066

aataataggt tccgaaatag gcaggggggca ttaactgttt atacgggcac tgttactcaa 2126

ggcactgacc ccgttaaaac ttattaccag tacactcctg tatcatcaaa agccatgtat 2186

gacgcttact ggaacggtaa attcagagac tgcgctttcc attctggctt taatgaggat 2246

ttatttgttt gtgaatatca aggccaatcg tctgacctgc ctcaacctcc tgtcaatgct 2306

ggcggcggct ctggtggtgg ttctggtggc ggctctgagg gtggtggctc tgagggaggc 2366

ggttccggtg gtggctctgg t tcc ggt gat ttt gat tat gaa aag atg gca   2417
                       Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala
                           115                 120

aac gct aat aag ggg gct atg acc gaa aat gcc gat gaa aac gcg cta   2465
Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu
    125                 130                 135

cag tct gac gct aaa ggc aaa ctt gat tct gtc gct act gat tac ggt   2513
Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly
140                 145                 150                 155

gct gct atc gat ggt ttc att ggt gac gtt tcc ggc ctt gct aat ggt   2561
Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly
                160                 165                 170
```

```
aat ggt gct act ggt gat ttt gct ggc tct aat tcc caa atg gct caa   2609
Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln
            175                 180                 185

gtc ggt gac ggt gat aat tca cct tta atg aat aat ttc cgt caa tat   2657
Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr
        190                 195                 200

tta cct tcc ctc cct caa tcg gtt gaa tgt cgc cct ttt gtc ttt ggc   2705
Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe Gly
    205                 210                 215

gct ggt aaa cca tat gaa ttt tct att gat tgt gac aaa ata aac tta   2753
Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu
220                 225                 230                 235

ttc cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg tat   2801
Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr
                240                 245                 250

gta ttt tct acg ttt gct aac ata ctg cgt aat aag gag tct taatc atg 2851
Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser       Met
            255                 260                 265

cca gtt ctt ttg ggt att ccg tta tta ttg cgt ttc ctc ggt           2893
Pro Val Leu Leu Gly Ile Pro Leu Leu Leu Arg Phe Leu Gly
            270                 275                 280

ttccttctgg taactttgtt cggctatctg cttacttttc ttaaaaaggg cttcggtaag 2953

atagctattg ctatttcatt gtttcttgct cttattattg ggcttaactc aattcttgtg 3013

ggttatctct ctgatattag cgctcaatta ccctctgact ttgttcaggg tgttcagtta 3073

attctcccgt ctaatgcgct tccctgtttt tatgttattc tctctgtaaa ggctgctatt 3133

ttcatttttg acgttaaaca aaaaatcgtt tcttatttgg attgggataa ataat atg  3191
                                                              Met

gct gtt tat ttt gta act ggc aaa tta ggc tct gga aag acg ctc gtt   3239
Ala Val Tyr Phe Val Thr Gly Lys Leu Gly Ser Gly Lys Thr Leu Val
            285                 290                 295

agc gtt ggt aag att cag gat aaa att gta gct ggg tgc aaa ata gca   3287
Ser Val Gly Lys Ile Gln Asp Lys Ile Val Ala Gly Cys Lys Ile Ala
        300                 305                 310

act aat ctt gat tta agg ctt caa aac ctc ccg caa gtc ggg agg ttc   3335
Thr Asn Leu Asp Leu Arg Leu Gln Asn Leu Pro Gln Val Gly Arg Phe
    315                 320                 325

gct aaa acg cct cgc gtt ctt aga ata ccg gat aag cct tct ata tct   3383
Ala Lys Thr Pro Arg Val Leu Arg Ile Pro Asp Lys Pro Ser Ile Ser
330                 335                 340                 345

gat ttg ctt gct att ggg cgc ggt aat gat tcc tac gat gaa aat aaa   3431
Asp Leu Leu Ala Ile Gly Arg Gly Asn Asp Ser Tyr Asp Glu Asn Lys
                350                 355                 360
```

```
aac ggc ttg ctt gtt ctc gat gag tgc ggt act tgg ttt aat acc cgt   3479
Asn Gly Leu Leu Val Leu Asp Glu Cys Gly Thr Trp Phe Asn Thr Arg
            365                 370                 375

tct tgg aat gat aag gaa aga cag ccg att att gat tgg ttt cta cat   3527
Ser Trp Asn Asp Lys Glu Arg Gln Pro Ile Ile Asp Trp Phe Leu His
        380                 385                 390

gct cgt aaa tta gga tgg gat att att ttt ctt gtt cag gac tta tct   3575
Ala Arg Lys Leu Gly Trp Asp Ile Ile Phe Leu Val Gln Asp Leu Ser
    395                 400                 405

att gtt gat aaa cag gcg cgt tct gca tta gct gaa cat gtt gtt tat   3623
Ile Val Asp Lys Gln Ala Arg Ser Ala Leu Ala Glu His Val Val Tyr
410                 415                 420                 425

tgt cgt cgt ctg gac aga att act tta cct ttt gtc ggt act tta tat   3671
Cys Arg Arg Leu Asp Arg Ile Thr Leu Pro Phe Val Gly Thr Leu Tyr
                430                 435                 440

tct ctt att act ggc tcg aaa atg cct ctg cct aaa tta cat gtt ggc   3719
Ser Leu Ile Thr Gly Ser Lys Met Pro Leu Pro Lys Leu His Val Gly
            445                 450                 455

gtt gtt aaa tat ggc gat tct caa tta agc cct act gtt gag cgt tgg   3767
Val Val Lys Tyr Gly Asp Ser Gln Leu Ser Pro Thr Val Glu Arg Trp
        460                 465                 470

ctt tat act ggt aag aat ttg tat aac gca tat gat act aaa cag gct   3815
Leu Tyr Thr Gly Lys Asn Leu Tyr Asn Ala Tyr Asp Thr Lys Gln Ala
    475                 480                 485

ttt tct agt aat tat gat tcc ggt gtt tat tct tat tta acg cct tat   3863
Phe Ser Ser Asn Tyr Asp Ser Gly Val Tyr Ser Tyr Leu Thr Pro Tyr
490                 495                 500                 505

tta tca cac ggt cgg tat ttc aaa cca tta aat tta ggt cag aag atg   3911
Leu Ser His Gly Arg Tyr Phe Lys Pro Leu Asn Leu Gly Gln Lys Met
                510                 515                 520

aaa tta act aaa ata tat ttg aaa aag ttt tct cgc gtt ctt tgt ctt   3959
Lys Leu Thr Lys Ile Tyr Leu Lys Lys Phe Ser Arg Val Leu Cys Leu
            525                 530                 535

gcg att gga ttt gca tca gca ttt aca tat agt tat ata acc caa cct   4007
Ala Ile Gly Phe Ala Ser Ala Phe Thr Tyr Ser Tyr Ile Thr Gln Pro
        540                 545                 550

aag ccg gag gtt aaa aag gta gtc tct cag acc tat gat ttt gat aaa   4055
Lys Pro Glu Val Lys Lys Val Val Ser Gln Thr Tyr Asp Phe Asp Lys
    555                 560                 565

ttc act att gac tct tct cag cgt ctt aat cta agc tat cgc tat gtt   4103
Phe Thr Ile Asp Ser Ser Gln Arg Leu Asn Leu Ser Tyr Arg Tyr Val
570                 575                 580                 585

ttc aag gat tct aag gga aaa tta att aat agc gac gat tta cag aag   4151
Phe Lys Asp Ser Lys Gly Lys Leu Ile Asn Ser Asp Asp Leu Gln Lys
                590                 595                 600
```

```
caa ggt tat tca ctc aca tat att gat tta tgt act gtt tcc att aaa   4199
Gln Gly Tyr Ser Leu Thr Tyr Ile Asp Leu Cys Thr Val Ser Ile Lys
            605                 610                 615

aaa ggt aat tca aat gaa att gtt aaa tgt aat taattttgtt ttcttgatgt 4252
Lys Gly Asn Ser Asn Glu Ile Val Lys Cys Asn
        620                 625

ttgtttcatc atcttctttt gctcaggtaa ttgaaatgaa taattcgcct ctgcgcgatt 4312

ttgtaacttg gtattcaaag caatcaggcg aatccgttat tgtttctccc gatgtaaaag 4372

gtactgttac tgtatattca tctgacgtta aacctgaaaa tctacgcaat ttctttattt 4432

ctgttttacg tgcaaataat tttgatatgg taggttctaa cccttccatt attcagaagt 4492

ataatccaaa caatcaggat tatattgatg aattgccatc atctgataat caggaatatg 4552

atgataattc cgctccttct ggtggtttct ttgttccgca aaatgataat gttactcaaa 4612

cttttaaaat taataacgtt cgggcaaagg atttaatacg agttgtcgaa ttgtttgtaa 4672

agtctaatac ttctaaatcc tcaaatgtat tatctattga cggctctaat ctattagttg 4732

ttagtgctcc taaagatatt ttagataacc ttcctcaatt cctttcaact gttgatttgc 4792

caactgacca gatattgatt gagggtttga tatttgaggt tcagcaaggt gatgctttag 4852

atttttcatt tgctgctggc tctcagcgtg gcactgttgc aggcggtgtt aatactgacc 4912

gcctcacctc tgttttatct tctgctggtg gttcgttcgg tatttttaat ggcgatgttt 4972

tagggctatc agttcgcgca ttaaagacta atagccattc aaaaatattg tctgtgccac 5032

gtattcttac gctttcaggt cagaagggtt ctatctctgt tggccagaat gtccctttta 5092

ttactggtcg tgtgactggt gaatctgcca atgtaaataa tccatttcag acgattgagc 5152

gtcaaaatgt aggtatttcc atgagcgttt ttcctgttgc aatggctggc ggtaatattg 5212

ttctggatat taccagcaag gccgatagtt tgagttcttc tactcaggca agtgatgtta 5272

ttactaatca aagaagtatt gctacaacgg ttaatttgcg tgatggacag actcttttac 5332

tcggtggcct cactgattat aaaaacactt ctcaggattc tggcgtaccg ttcctgtcta 5392

aaatcccttt aatcggcctc ctgtttagct cccgctctga ttctaacgag gaaagcacgt 5452

tatacgtgct cgtcaaagca accatagtac gcgccctgta gcggcgcatt aagcgcggcg 5512

ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct 5572

ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat 5632

cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt 5692

gatttgggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg 5752

acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac 5812

cctatctcgg gctattcttt tgatttataa gggattttgc cgatttcgga accaccatca 5872
```

```
aacaggattt tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa ctctctcagg 5932

gccaggcggt gaagggcaat cagctgttgc ccgtctcact ggtgaaaaga aaaaccaccc 5992

tggatccaag cttgcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt 6052

atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct 6112

tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc 6172

ctttttttgcg gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa 6232

agatgctgaa gatcagttgg gcgcactagt gggttacatc gaactggatc tcaacagcgg 6292

taagatcctt gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt 6352

tctgctatgt ggcgcggtat tatcccgtat tgacgccggg caagagcaac tcggtcgccg 6412

catacactat tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac 6472

ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc 6532

ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgcacaa 6592

catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc 6652

aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt 6712

aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga 6772

taaagttgca ggaccacttc tgcgctcggc ccttccggct ggctggttta ttgctgataa 6832

atctggagcc ggtgagcgtg ggtctcgcgg tatcattgca gcactggggc cagatggtaa 6892

gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa 6952

tagacagatc gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt 7012

ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt 7072

gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg 7132

tacgtaagac ccccaagctt gtcgactgaa tggcgaatgg cgctttgcct ggtttccggc 7192

accagaagcg gtgccggaaa gctggctgga gtgcgatctt cctgacgctc gagcgcaacg 7252

caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 7312

ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 7372

atgattacgc caagctttgg agccttttt ttggagattt tcaac gtg aaa aaa tta 7429
                                                Met Lys Lys Leu
                                                    630

tta ttc gca att cct tta gtt gtt cct ttc tat tct cac agt gca caa   7477
Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln
        635                 640                 645
```

```
gac atc cag atg acc cag tct cca gcc acc ctg tct ttg tct cca ggg   7525
Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
    650                 655                 660

gaa aga gcc acc ctc tcc tgc agg gcc agt cag ggt gtt agc agc tac   7573
Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Gly Val Ser Ser Tyr
665                 670                 675                 680

tta gcc tgg tac cag cag aaa cct ggc cag gct ccc agg ctc ctc atc   7621
Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
                685                 690                 695

tat gat gca tcc aac agg gcc act ggc atc cca gcc agg ttc agt ggc   7669
Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
            700                 705                 710

agt ggg cct ggg aca gac ttc act ctc acc atc agc agc cta gag cct   7717
Ser Gly Pro Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
        715                 720                 725

gaa gat ttt gca gtt tat tac tgt cag cag cgt aac tgg cat ccg tgg   7765
Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asn Trp His Pro Trp
    730                 735                 740

acg ttc ggc caa ggg acc aag gtg gaa atc aaa cga act gtg gct gca   7813
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
745                 750                 755                 760

cca tct gtc ttc atc ttc ccg cca tct gat gag cag ttg aaa tct gga   7861
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
                765                 770                 775

act gcc tct gtt gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc   7909
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
            780                 785                 790

aaa gta cag tgg aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag   7957
Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
        795                 800                 805

gag agt gtc aca gag cgg gac agc aag gac agc acc tac agc ctc agc   8005
Glu Ser Val Thr Glu Arg Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
    810                 815                 820

agc acc ctg acg ctg agc aaa gca gac tac gag aaa cac aaa gtc tac   8053
Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
825                 830                 835                 840

gcc tgc gaa gtc acc cat cag ggc ctg agc tcg ccc gtc aca aag agc   8101
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
                845                 850                 855

ttc aac agg gga gag tgt taataaggcg cgccaattct atttcaagga          8149
Phe Asn Arg Gly Glu Cys
            860

gacagtcata atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta    8198
          Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu
                  865                 870                 875
```

```
tta ctc gcg gcc cag ccg gcc atg gcc gaa gtt caa ttg tta gag tct   8246
Leu Leu Ala Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser
                880                 885                 890

ggt ggc ggt ctt gtt cag cct ggt ggt tct tta cgt ctt tct tgc gct   8294
Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala
            895                 900                 905

gct tcc gga ttc act ttc tct act tac gag atg cgt tgg gtt cgc caa   8342
Ala Ser Gly Phe Thr Phe Ser Thr Tyr Glu Met Arg Trp Val Arg Gln
        910                 915                 920

gct cct ggt aaa ggt ttg gag tgg gtt tct tat atc gct cct tct ggt   8390
Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Tyr Ile Ala Pro Ser Gly
    925                 930                 935

ggc gat act gct tat gct gac tcc gtt aaa ggt cgc ttc act atc tct   8438
Gly Asp Thr Ala Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
940                 945                 950                 955

aga gac aac tct aag aat act ctc tac ttg cag atg aac agc tta agg   8486
Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
                960                 965                 970

gct gag gac act gca gtc tac tat tgt gcg agg agg ctc gat ggc tat   8534
Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Arg Leu Asp Gly Tyr
            975                 980                 985

att tcc tac tac tac ggt atg gac gtc tgg ggc caa ggg acc acg gtc   8582
Ile Ser Tyr Tyr Tyr Gly Met Asp Val Trp Gly Gln Gly Thr Thr Val
        990                 995                1000

acc gtc tca agc gcc tcc acc aag ggc cca tcg gtc ttc ccc ctg gca   8630
Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala
   1005                1010                1015

ccc tcc tcc aag agc acc tct ggg ggc aca gcg gcc ctg ggc tgc ctg   8678
Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu
1020                1025                1030                1035

gtc aag gac tac ttc ccc gaa ccg gtg acg gtg tcg tgg aac tca ggc   8726
Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly
               1040                1045                1050

gcc ctg acc agc ggc gtc cac acc ttc ccg gct gtc cta cag tcc tca   8774
Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser
           1055                1060                1065

gga ctc tac tcc ctc agc agc gta gtg acc gtg ccc tcc agc agc ttg   8822
Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu
       1070                1075                1080

ggc acc cag acc tac atc tgc aac gtg aat cac aag ccc agc aac acc   8870
Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr
   1085                1090                1095

aag gtg gac aag aaa gtt gag ccc aaa tct tgt gcg gcc gca cat cat   8918
Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Ala Ala Ala His His
1100                1105                1110                1115
```

```
cat cac cat cac ggg gcc gca gaa caa aaa ctc atc tca gaa gag gat   8966
His His His His Gly Ala Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp
               1120                1125                1130

ctg aat ggg gcc gca tag gct agc tct gct wsy ggy gay tty gay tay   9014
Leu Asn Gly Ala Ala Gln Ala Ser Ser Ala Ser Gly Asp Phe Asp Tyr
           1135                1140                1145

gar aar atg gct aaw gcy aay aar ggs gcy atg acy gar aay gcy gay   9062
Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp
       1150                1155                1160

gar aay gck ytr car wsy gay gcy aar ggy aar ytw gay wsy gtc gck   9110
Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala
   1165                1170                1175

acy gay tay ggy gcy gcc atc gay ggy tty aty ggy gay gtc wsy ggy   9158
Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly
1180                1185                1190                1195

ytk gcy aay ggy aay ggy gcy acy ggw gay tty gcw ggy tck aat tcy   9206
Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser
               1200                1205                1210

car atg gcy car gty ggw gay ggk gay aay wsw cck ytw atg aay aay   9254
Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn
           1215                1220                1225

tty mgw car tay ytw cck tcy cty cck car wsk gty gar tgy cgy ccw   9302
Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro
       1230                1235                1240

tty gty tty wsy gcy ggy aar ccw tay gar tty wsy aty gay tgy gay   9350
Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp
   1245                1250                1255

aar atm aay ytw tty cgy ggy gty tty gck tty ytk yta tay gty gcy   9398
Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala
1260                1265                1270                1275

acy tty atg tay gtw tty wsy ack tty gcy aay atw ytr cgy aay aar   9446
Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys
               1280                1285                1290

gar wsy tagtgatctc ctaggaagcc cgcctaatga gcgggctttt tttttctggt    9502
Glu Ser

atgcatcctg aggccgatac tgtcgtcgtc ccctcaaact ggcagatgca cggttacgat 9562

gcgcccatct acaccaacgt gacctatccc attacggtca atccgccgtt tgttcccacg 9622

gagaatccga cgggttgtta ctcgctcaca tttaatgttg atgaaagctg gctacaggaa 9682

ggccagacgc gaattatttt tgatggcgtt cctattggtt aaaaaatgag ctgatttaac 9742

aaaaatttaa tgcgaatttt aacaaaatat taacgtttac aatttaaata tttgcttata 9802

caatcttcct gtttttgggg cttttctgat tatcaaccgg ggtacatatg attgacatgc 9862

tagttttacg attaccgttc atcgattctc ttgtttgctc cagactctca ggcaatgacc 9922
```

tgatagcctt tgtagatctc tcaaaaatag ctaccctctc cggcattaat ttatcagcta 9982

gaacggttga atatcatatt gatggtgatt tgactgtctc cggcctttct cacccttttg 10042

aatctttacc tacacattac tcaggcattg catttaaaat atatgagggt tctaaaaatt 10102

tttatccttg cgttgaaata aaggcttctc ccgcaaaagt attacagggt cataatgttt 10162

ttggtacaac cgatttagct ttatgctctg aggctttatt gcttaatttt gctaattctt 10222

tgccttgcct gtatgattta ttggatgtt 10251

<210> 583
<211> 113
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica CJRA05

<400> 583
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
1 5 10 15

Gly Ala Ala Glu Ser His Leu Asp Gly Ala Ala Glu Thr Val Glu Ser
20 25 30

Cys Leu Ala Lys Ser His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys
35 40 45

Asp Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp
50 55 60

Asn Ala Thr Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr
65 70 75 80

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

Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
100 105 110

Thr

<210> 584
<211> 152
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica CJRA05

<400> 584
Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala
1 5 10 15

```
Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly
            20                  25                  30

Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe
        35                  40                  45

Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp
    50                  55                  60

Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn
65                  70                  75                  80

Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln
                85                  90                  95

Ser Val Glu Cys Arg Pro Phe Val Phe Gly Ala Gly Lys Pro Tyr Glu
            100                 105                 110

Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala
        115                 120                 125

Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala
    130                 135                 140

Asn Ile Leu Arg Asn Lys Glu Ser
145                 150
```

<210> 585
<211> 15
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência peptídica CJRA05

<400> 585
```
Met Pro Val Leu Leu Gly Ile Pro Leu Leu Leu Arg Phe Leu Gly
  1               5                  10                  15
```

<210> 586
<211> 348
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica CJRA05

<400> 586
```
Met Ala Val Tyr Phe Val Thr Gly Lys Leu Gly Ser Gly Lys Thr Leu
  1               5                  10                  15

Val Ser Val Gly Lys Ile Gln Asp Lys Ile Val Ala Gly Cys Lys Ile
            20                  25                  30

Ala Thr Asn Leu Asp Leu Arg Leu Gln Asn Leu Pro Gln Val Gly Arg
        35                  40                  45
```

```
Phe Ala Lys Thr Pro Arg Val Leu Arg Ile Pro Asp Lys Pro Ser Ile
     50                  55                  60

Ser Asp Leu Leu Ala Ile Gly Arg Gly Asn Asp Ser Tyr Asp Glu Asn
 65                  70                  75                  80

Lys Asn Gly Leu Leu Val Leu Asp Glu Cys Gly Thr Trp Phe Asn Thr
                 85                  90                  95

Arg Ser Trp Asn Asp Lys Glu Arg Gln Pro Ile Ile Asp Trp Phe Leu
            100                 105                 110

His Ala Arg Lys Leu Gly Trp Asp Ile Ile Phe Leu Val Gln Asp Leu
        115                 120                 125

Ser Ile Val Asp Lys Gln Ala Arg Ser Ala Leu Ala Glu His Val Val
    130                 135                 140

Tyr Cys Arg Arg Leu Asp Arg Ile Thr Leu Pro Phe Val Gly Thr Leu
145                 150                 155                 160

Tyr Ser Leu Ile Thr Gly Ser Lys Met Pro Leu Pro Lys Leu His Val
                165                 170                 175

Gly Val Val Lys Tyr Gly Asp Ser Gln Leu Ser Pro Thr Val Glu Arg
            180                 185                 190

Trp Leu Tyr Thr Gly Lys Asn Leu Tyr Asn Ala Tyr Asp Thr Lys Gln
        195                 200                 205

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

Tyr Leu Ser His Gly Arg Tyr Phe Lys Pro Leu Asn Leu Gly Gln Lys
225                 230                 235                 240

Met Lys Leu Thr Lys Ile Tyr Leu Lys Lys Phe Ser Arg Val Leu Cys
                245                 250                 255

Leu Ala Ile Gly Phe Ala Ser Ala Phe Thr Tyr Ser Tyr Ile Thr Gln
            260                 265                 270

Pro Lys Pro Glu Val Lys Lys Val Val Ser Gln Thr Tyr Asp Phe Asp
        275                 280                 285

Lys Phe Thr Ile Asp Ser Ser Gln Arg Leu Asn Leu Ser Tyr Arg Tyr
    290                 295                 300

Val Phe Lys Asp Ser Lys Gly Lys Leu Ile Asn Ser Asp Asp Leu Gln
305                 310                 315                 320

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

Lys Lys Gly Asn Ser Asn Glu Ile Val Lys Cys Asn
            340                 345
```

<210> 587
<211> 234
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica CJRA05

<400> 587
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
1 5 10 15

His Ser Ala Gln Asp Ile Gln Met Thr Gln Ser Pro Ala Thr Leu Ser
20 25 30

Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Gly
35 40 45

Val Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro
50 55 60

Arg Leu Leu Ile Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala
65 70 75 80

Arg Phe Ser Gly Ser Gly Pro Gly Thr Asp Phe Thr Leu Thr Ile Ser
85 90 95

Ser Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Arg Asn
100 105 110

Trp His Pro Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
115 120 125

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
130 135 140

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
145 150 155 160

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
165 170 175

Gly Asn Ser Gln Glu Ser Val Thr Glu Arg Asp Ser Lys Asp Ser Thr
180 185 190

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
195 200 205

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
210 215 220

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225 230

<210> 588
<211> 431
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica CJRA05

<400> 588
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly
20 25 30

Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
35 40 45

Phe Thr Phe Ser Thr Tyr Glu Met Arg Trp Val Arg Gln Ala Pro Gly
50 55 60

Lys Gly Leu Glu Trp Val Ser Tyr Ile Ala Pro Ser Gly Gly Asp Thr
65 70 75 80

Ala Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
85 90 95

Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp
100 105 110

Thr Ala Val Tyr Tyr Cys Ala Arg Arg Leu Asp Gly Tyr Ile Ser Tyr
115 120 125

Tyr Tyr Gly Met Asp Val Trp Gly Gln Gly Thr Thr Val Thr Val Ser
130 135 140

Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser
145 150 155 160

Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp
165 170 175

Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr
180 185 190

Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr
195 200 205

Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln
210 215 220

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp
225 230 235 240

Lys Lys Val Glu Pro Lys Ser Cys Ala Ala Ala His His His His His
245 250 255

His Gly Ala Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly
260 265 270

Ala Ala Gln Ala Ser Ser Ala Ser Gly Asp Phe Asp Tyr Glu Lys Met
275 280 285

Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala
290 295 300

```
Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr
305                 310                 315                 320

Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn
                325                 330                 335

Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala
            340                 345                 350

Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln
        355                 360                 365

Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe
    370                 375                 380

Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn
385                 390                 395                 400

Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met
                405                 410                 415

Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
            420                 425                 430


<210> 589
<211> 5
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido ilustrativo


<400> 589
Glu Gly Gly Gly Ser
  1               5


<210> 590
<211> 1275
<212> ADN
<213> Organismo Desconhecido

<220>
<221> CDS
<222> (1)..(1272)

<220>
<223> Descrição do Organismo Desconhecido: Sequência nucleotídica M13


<400> 590
gtg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat tct   48
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
  1               5                  10                  15

cac tcc gct gaa act gtt gaa agt tgt tta gca aaa ccc cat aca gaa   96
His Ser Ala Glu Thr Val Glu Ser Cys Leu Ala Lys Pro His Thr Glu
             20                  25                  30
```

```
aat tca ttt act aac gtc tgg aaa gac gac aaa act tta gat cgt tac   144
Asn Ser Phe Thr Asn Val Trp Lys Asp Asp Lys Thr Leu Asp Arg Tyr
         35                  40                  45

gct aac tat gag ggt tgt ctg tgg aat gct aca ggc gtt gta gtt tgt   192
Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala Thr Gly Val Val Val Cys
     50                  55                  60

act ggt gac gaa act cag tgt tac ggt aca tgg gtt cct att ggg ctt   240
Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr Trp Val Pro Ile Gly Leu
 65                  70                  75                  80

gct atc cct gaa aat gag ggt ggt ggc tct gag ggt ggc ggt tct gag   288
Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu
                 85                  90                  95

ggt ggc ggt tct gag ggt ggc ggt act aaa cct cct gag tac ggt gat   336
Gly Gly Gly Ser Glu Gly Gly Gly Thr Lys Pro Pro Glu Tyr Gly Asp
            100                 105                 110

aca cct att ccg ggc tat act tat atc aac cct ctc gac ggc act tat   384
Thr Pro Ile Pro Gly Tyr Thr Tyr Ile Asn Pro Leu Asp Gly Thr Tyr
        115                 120                 125

ccg cct ggt act gag caa aac ccc gct aat cct aat cct tct ctt gag   432
Pro Pro Gly Thr Glu Gln Asn Pro Ala Asn Pro Asn Pro Ser Leu Glu
    130                 135                 140

gag tct cag cct ctt aat act ttc atg ttt cag aat aat agg ttc cga   480
Glu Ser Gln Pro Leu Asn Thr Phe Met Phe Gln Asn Asn Arg Phe Arg
145                 150                 155                 160

aat agg cag ggg gca tta act gtt tat acg ggc act gtt act caa ggc   528
Asn Arg Gln Gly Ala Leu Thr Val Tyr Thr Gly Thr Val Thr Gln Gly
                165                 170                 175

act gac ccc gtt aaa act tat tac cag tac act cct gta tca tca aaa   576
Thr Asp Pro Val Lys Thr Tyr Tyr Gln Tyr Thr Pro Val Ser Ser Lys
            180                 185                 190

gcc atg tat gac gct tac tgg aac ggt aaa ttc aga gac tgc gct ttc   624
Ala Met Tyr Asp Ala Tyr Trp Asn Gly Lys Phe Arg Asp Cys Ala Phe
        195                 200                 205

cat tct ggc ttt aat gag gat cca ttc gtt tgt gaa tat caa ggc caa   672
His Ser Gly Phe Asn Glu Asp Pro Phe Val Cys Glu Tyr Gln Gly Gln
    210                 215                 220

tcg tct gac ctg cct caa cct cct gtc aat gct ggc ggc ggc tct ggt   720
Ser Ser Asp Leu Pro Gln Pro Pro Val Asn Ala Gly Gly Gly Ser Gly
225                 230                 235                 240

ggt ggt tct ggt ggc ggc tct gag ggt ggt ggc tct gag ggt ggc ggt   768
Gly Gly Ser Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
                245                 250                 255

tct gag ggt ggc ggc tct gag gga ggc ggt tcc ggt ggt ggc tct ggt   816
Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly
            260                 265                 270
```

```
tcc ggt gat ttt gat tat gaa aag atg gca aac gct aat aag ggg gct   864
Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala
        275                 280                 285

atg acc gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc   912
Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly
    290                 295                 300

aaa ctt gat tct gtc gct act gat tac ggt gct gct atc gat ggt ttc   960
Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe
305                 310                 315                 320

att ggt gac gtt tcc ggc ctt gct aat ggt aat ggt gct act ggt gat   1008
Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp
                325                 330                 335

ttt gct ggc tct aat tcc caa atg gct caa gtc ggt gac ggt gat aat   1056
Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn
            340                 345                 350

tca cct tta atg aat aat ttc cgt caa tat tta cct tcc ctc cct caa   1104
Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln
        355                 360                 365

tcg gtt gaa tgt cgc cct ttt gtc ttt agc gct ggt aaa cca tat gaa   1152
Ser Val Glu Cys Arg Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu
    370                 375                 380

ttt tct att gat tgt gac aaa ata aac tta ttc cgt ggt gtc ttt gcg   1200
Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala
385                 390                 395                 400

ttt ctt tta tat gtt gcc acc ttt atg tat gta ttt tct acg ttt gct   1248
Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala
                405                 410                 415

aac ata ctg cgt aat aag gag tct taa                              1275
Asn Ile Leu Arg Asn Lys Glu Ser
            420
```

<210> 591
<211> 424
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: Sequência proteica M13

<400> 591

```
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
  1               5                  10                  15

His Ser Ala Glu Thr Val Glu Ser Cys Leu Ala Lys Pro His Thr Glu
             20                  25                  30

Asn Ser Phe Thr Asn Val Trp Lys Asp Asp Lys Thr Leu Asp Arg Tyr
         35                  40                  45
```

```
Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala Thr Gly Val Val Val Cys
     50                  55                  60

Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr Trp Val Pro Ile Gly Leu
 65                  70                  75                  80

Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu
                 85                  90                  95

Gly Gly Gly Ser Glu Gly Gly Gly Thr Lys Pro Pro Glu Tyr Gly Asp
            100                 105                 110

Thr Pro Ile Pro Gly Tyr Thr Tyr Ile Asn Pro Leu Asp Gly Thr Tyr
        115                 120                 125

Pro Pro Gly Thr Glu Gln Asn Pro Ala Asn Pro Asn Pro Ser Leu Glu
    130                 135                 140

Glu Ser Gln Pro Leu Asn Thr Phe Met Phe Gln Asn Asn Arg Phe Arg
145                 150                 155                 160

Asn Arg Gln Gly Ala Leu Thr Val Tyr Thr Gly Thr Val Thr Gln Gly
                165                 170                 175

Thr Asp Pro Val Lys Thr Tyr Tyr Gln Tyr Thr Pro Val Ser Ser Lys
            180                 185                 190

Ala Met Tyr Asp Ala Tyr Trp Asn Gly Lys Phe Arg Asp Cys Ala Phe
        195                 200                 205

His Ser Gly Phe Asn Glu Asp Pro Phe Val Cys Glu Tyr Gln Gly Gln
    210                 215                 220

Ser Ser Asp Leu Pro Gln Pro Pro Val Asn Ala Gly Gly Gly Ser Gly
225                 230                 235                 240

Gly Gly Ser Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
                245                 250                 255

Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly
            260                 265                 270

Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala
        275                 280                 285

Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly
    290                 295                 300

Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe
305                 310                 315                 320

Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp
                325                 330                 335

Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn
            340                 345                 350

Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln
        355                 360                 365
```

```
Ser Val Glu Cys Arg Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu
    370                 375                 380

Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala
385                 390                 395                 400

Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala
                405                 410                 415

Asn Ile Leu Arg Asn Lys Glu Ser
            420
```

<210> 592
<211> 35
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 592
caacgatgat cgtatggcgc atgctgccga gacag 35

<210> 593
<211> 1355
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência nucleotídica M13-III

<220>
<221> CDS
<222> (1)..(1305)

<400> 593

```
gcg gcc gca cat cat cat cac cat cac ggg gcc gca gaa caa aaa ctc   48
Ala Ala Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu
  1               5                  10                  15

atc tca gaa gag gat ctg aat ggg gcc gca tag gct agc gat atc aac   96
Ile Ser Glu Glu Asp Leu Asn Gly Ala Ala     Ala Ser Asp Ile Asn
            20                  25                  30

gat gat cgt atg gct tct act gcy gar acw gty gaa wsy tgy ytr gcm   144
Asp Asp Arg Met Ala Ser Thr Ala Glu Thr Val Glu Ser Cys Leu Ala
            35                  40                  45

aar ccy cay acw gar aat wsw tty acw aay gts tgg aar gay gay aar   192
Lys Pro His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys Asp Asp Lys
        50                  55                  60

acy ytw gat cgw tay gcy aay tay gar ggy tgy ytr tgg aat gcy acm   240
Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala Thr
    65                  70                  75
```

```
ggc gty gtw gty tgy ack ggy gay gar acw car tgy tay ggy acr tgg   288
Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr Trp
 80                  85                  90                  95

gtk cck atw ggs ytw gcy atm cck gar aay gar ggy ggy ggy wsy gar   336
Val Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser Glu
                100                 105                 110

ggy ggy ggy wsy gar ggy ggy ggw tcy gar ggw ggy ggw acy aar cck   384
Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr Lys Pro
            115                 120                 125

cck gar tay ggy gay acw cck atw cck ggy tay acy tay aty aay cck   432
Pro Glu Tyr Gly Asp Thr Pro Ile Pro Gly Tyr Thr Tyr Ile Asn Pro
        130                 135                 140

ytm gay ggm acy tay cck cck ggy acy gar car aay ccy gcy aay cck   480
Leu Asp Gly Thr Tyr Pro Pro Gly Thr Glu Gln Asn Pro Ala Asn Pro
    145                 150                 155

aay ccw wsy ytw gar gar wsy car cck ytw aay acy tty atg tty car   528
Asn Pro Ser Leu Glu Glu Ser Gln Pro Leu Asn Thr Phe Met Phe Gln
160                 165                 170                 175

aay aay mgk tty mgr aay mgk car ggk gcw ytw acy gtk tay ack ggm   576
Asn Asn Arg Phe Arg Asn Arg Gln Gly Ala Leu Thr Val Tyr Thr Gly
                180                 185                 190

acy gty acy car ggy acy gay ccy gty aar acy tay tay car tay acy   624
Thr Val Thr Gln Gly Thr Asp Pro Val Lys Thr Tyr Tyr Gln Tyr Thr
            195                 200                 205

cck gtm tcr wsw aar gcy atg tay gay gcy tay tgg aay ggy aar tty   672
Pro Val Ser Ser Lys Ala Met Tyr Asp Ala Tyr Trp Asn Gly Lys Phe
        210                 215                 220

mgw gay tgy gcy tty cay wsy ggy tty aay gar gay ccw tty gty tgy   720
Arg Asp Cys Ala Phe His Ser Gly Phe Asn Glu Asp Pro Phe Val Cys
    225                 230                 235

gar tay car ggy car wsk wsy gay ytr cck car ccw cck gty aay gck   768
Glu Tyr Gln Gly Gln Ser Ser Asp Leu Pro Gln Pro Pro Val Asn Ala
240                 245                 250                 255

ggy ggy ggy wsy ggy ggw ggy wsy ggy ggy ggy wsy gar ggy ggw ggy   816
Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Glu Gly Gly Gly
                260                 265                 270

wsy gar ggw ggy ggy wsy ggr ggy ggy wsy ggy wsy ggy gay tty gay   864
Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp
            275                 280                 285

tay gar aar atg gcw aay gcy aay aar ggs gcy atg acy gar aay gcy   912
Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala
        290                 295                 300

gay gar aay gcr ctr car wst gay gcy aar ggy aar ytw gay wsy gtc   960
Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val
    305                 310                 315

gcy acw gay tay ggt gct gcy atc gay ggy tty aty ggy gay gty wsy   1008
```

```
Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser
320                 325                 330                 335

ggy ctk gct aay ggy aay ggw gcy acy ggw gay tty gcw ggy tck aat   1056
Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn
                340                 345                 350

tcy car atg gcy car gty ggw gay ggk gay aay wsw cck ytw atg aay   1104
Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn
            355                 360                 365

aay tty mgw car tay ytw cck tcy cty cck car wsk gty gar tgy cgy   1152
Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg
        370                 375                 380

ccw tty gty tty wsy gcy ggy aar ccw tay gar tty wsy aty gay tgy   1200
Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys
    385                 390                 395

gay aar atm aay ytw ttc cgy ggy gty tty gck tty ytk yta tay gty   1248
Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val
400                 405                 410                 415

gcy acy tty atg tay gtw tty wsy ack tty gcy aay atw ytr cgy aay   1296
Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn
                420                 425                 430

aar gar wsy tagtgatctc ctaggaagcc cgcctaatga gcgggctttt           1345
Lys Glu Ser

ttttttctggt                                                       1355
```

<210> 594
<211> 434
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência proteica M13-III

<400> 594

```
Ala Ala Ala His His His His His His Gly Ala Ala Glu Gln Lys Leu
  1               5                  10                  15

Ile Ser Glu Glu Asp Leu Asn Gly Ala Ala Ala Ser Asp Ile Asn Asp
             20                  25                  30

Asp Arg Met Ala Ser Thr Ala Glu Thr Val Glu Ser Cys Leu Ala Lys
         35                  40                  45

Pro His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys Asp Asp Lys Thr
     50                  55                  60

Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn Ala Thr Gly
 65                  70                  75                  80

Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly Thr Trp Val
                 85                  90                  95
```

```
Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly Ser Glu Gly
            100                 105                 110

Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr Lys Pro Pro
        115                 120                 125

Glu Tyr Gly Asp Thr Pro Ile Pro Gly Tyr Thr Tyr Ile Asn Pro Leu
    130                 135                 140

Asp Gly Thr Tyr Pro Pro Gly Thr Glu Gln Asn Pro Ala Asn Pro Asn
145                 150                 155                 160

Pro Ser Leu Glu Glu Ser Gln Pro Leu Asn Thr Phe Met Phe Gln Asn
                165                 170                 175

Asn Arg Phe Arg Asn Arg Gln Gly Ala Leu Thr Val Tyr Thr Gly Thr
            180                 185                 190

Val Thr Gln Gly Thr Asp Pro Val Lys Thr Tyr Tyr Gln Tyr Thr Pro
        195                 200                 205

Val Ser Ser Lys Ala Met Tyr Asp Ala Tyr Trp Asn Gly Lys Phe Arg
    210                 215                 220

Asp Cys Ala Phe His Ser Gly Phe Asn Glu Asp Pro Phe Val Cys Glu
225                 230                 235                 240

Tyr Gln Gly Gln Ser Ser Asp Leu Pro Gln Pro Pro Val Asn Ala Gly
                245                 250                 255

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Glu Gly Gly Gly Ser
            260                 265                 270

Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp Tyr
        275                 280                 285

Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp
    290                 295                 300

Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala
305                 310                 315                 320

Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly
                325                 330                 335

Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser
            340                 345                 350

Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn Asn
        355                 360                 365

Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg Pro
    370                 375                 380

Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp
385                 390                 395                 400

Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala
                405                 410                 415
```

Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys
420 425 430

Glu Ser

<210> 595
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 595
cgttgatatc gctagcctat gc 22

<210> 596
<211> 30
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 596
gataggctta gctagcccgg agaacgaagg 30

<210> 597
<211> 37
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 597
ctttcacagc ggtttcgcta gcgacccttt tgtctgc 37

<210> 598
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 598
ctttcacagc ggtttcgcta gcgacccttt tgtcagcgag taccagggtc 50

<210> 599
<211> 37
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 599
gactgtctcg gcagcatgcg ccatacgatc atcgttg 37

<210> 600
<211> 37
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> CDS
<222> (2)..(25)

<400> 600

c aac gat gat cgt atg gcg cat gct gccgagacag tc 37
Asn Asp Asp Arg Met Ala His Ala
1 5

<210> 601
<211> 8
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Péptido Sintético

<400> 601

Asn Asp Asp Arg Met Ala His Ala
1 5

<210> 602
<211> 37
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 602
ctttcacagc ggtttgcatg cagacccttt tgtctgc 37

<210> 603
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 603
ctttcacagc ggtttgcatg cagacccttt tgtcagcgag taccagggtc 50

<210> 604
<211> 7
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido Ilustrativo

<400> 604
Tyr Ala Asp Ser Val Lys Gly
1 5

<210> 605
<211> 21
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 605
cctcgacagc gaagtgcaca g 21

<210> 606
<211> 38
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 606
ggctgagtca agacgctctg tgcacttcgc tgtcgagg 38

<210> 607
<211> 7
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido Ilustrativo

<400> 607
Gln Ser Ala Leu Thr Gln Pro
1 5

<210> 608
<211> 22
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 608
cctctgtcac agtgcacaag ac 22

<210> 609
<211> 42
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 609
cctctgtcac agtgcacaag acatccagat gacccagtct cc 42

<210> 610
<211> 50
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 610
gggaggatgg agactgggtc gtctggatgt cttgtgcact gtgacagagg 50

<210> 611
<211> 11
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido ilustrativo

<400> 611
Gln Asp Ile Gln Met Thr Gln Ser Pro Ser Ser
1 5 10

<210> 612
<211> 20
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Iniciador

<400> 612
gactgggtgt agtgatctag 20

<210> 613
<211> 28
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 613
ggtgtagtga tcttctagtg acaactct 28

<210> 614
<211> 6
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido sintético

<400> 614
Val Ser Ser Arg Asp Asn
1 5

<210> 615
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> CDS
<222> (1)..(15)

<400> 615
tac tat tgt gcg aaa 15
Tyr Tyr Cys Ala Lys
1 5

<210> 616
<211> 5
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido sintética

<400> 616
Tyr Tyr Cys Ala Lys
1 5

<210> 617
<211> 36
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 617
ggtgccgata ggcttgcatg caccggagaa cgaagg 36

<210> 618
<211> 95
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<400> 618
cgcttcacta agtctagaga caactctaag aatactctct acttgcagat gaacagctta 60
agggctgagg acactgcagt ctactattgt acgag 95

<210> 619
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Oligonucleótido sintético

<220>
<221> base_modificada
<222> (4)..(7)
<223> A, T, C, G, outro ou desconhecido

<400> 619
gatnnnnatc 10

<210> 620
<211> 10
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: péptido derivado de MALIA3

<400> 620
Met Lys Leu Leu Asn Val Ile Asn Phe Val
  1   5   10

<210> 621
<211> 29
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido derivado de CJRA05

<400> 621
Met Ser Val Leu Val Tyr Ser Phe Ala Ser Phe Val Leu Gly Trp Cys
  1   5   10   15

Leu Arg Ser Gly Ile Thr Tyr Phe Thr Arg Leu Met Glu
             20   25

<210> 622
<211> 15
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: Sequência nucleotídica ilustrativa

<400> 622
ttttttttttt ttttt   15

<210> 623
<211> 87
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: péptido derivado de MALIA3

<400> 623
Met Ile Lys Val Glu Ile Lys Pro Ser Gln Ala Gln Phe Thr Thr Arg
  1   5   10   15

Ser Gly Val Ser Arg Gln Gly Lys Pro Tyr Ser Leu Asn Glu Gln Leu
             20   25   30

Cys Tyr Val Asp Leu Gly Asn Glu Tyr Pro Val Leu Val Lys Ile Thr
         35   40   45

Leu Asp Glu Gly Gln Pro Ala Tyr Ala Pro Gly Leu Tyr Thr Val His
     50   55   60

Leu Ser Ser Phe Lys Val Gly Gln Phe Gly Ser Leu Met Ile Asp Arg
 65   70   75   80

Leu Arg Leu Val Pro Ala Lys
                 85

<210> 624
<211> 29
<212> PRT
<213> Organismo Desconhecido

<220>
<223> Descrição do Organismo Desconhecido: péptido derivado de MALIA3

<400> 624
Met Ser Val Leu Val Tyr Ser Phe Ala Ser Phe Val Leu Gly Trp Cys
  1   5   10   15

Leu Arg Ser Gly Ile Thr Tyr Phe Thr Arg Leu Met Glu
             20   25

<210> 625
<211> 10
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<220>
<221> base_modificada
<222> (7)..(10)
<223> A, T, C, G, outro ou desconhecido

<400> 625
ctcttcnnnn   10

<210> 626
<211> 87
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido derivado de CJRA05

<400> 626
Met Ile Lys Val Glu Ile Lys Pro Ser Gln Ala Gln Phe Thr Thr Arg
  1   5   10   15

Ser Gly Val Ser Arg Gln Gly Lys Pro Tyr Ser Leu Asn Glu Gln Leu
             20   25   30

Cys Tyr Val Asp Leu Gly Asn Glu Tyr Pro Val Leu Val Lys Ile Thr
         35   40   45

Leu Asp Glu Gly Gln Pro Ala Tyr Ala Pro Gly Leu Tyr Thr Val His
     50   55   60

Leu Ser Ser Phe Lys Val Gly Gln Phe Gly Ser Leu Met Ile Asp Arg
 65   70   75   80

Leu Arg Leu Val Pro Ala Lys
85

<210> 627
<211> 10
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: péptido derivado de CJRA05

<400> 627
Met Lys Leu Leu Asn Val Ile Asn Phe Val
1 5 10

<210> 628
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 628
gacccagtct ccatcctcc 19

<210> 629
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 629
gactcagtct ccactctcc 19

<210> 630
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 630
gacgcagtct ccaggcacc 19

<210> 631
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 631
gacgcagtct ccagccacc 19

<210> 632
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 632
gtctcctgga cagtcgatc 19

<210> 633
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 633
ggccttggga cagacagtc 19

<210> 634
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 634
gtctcctgga cagtcagtc 19

<210> 635
<211> 19
<212> ADN
<213> Sequência Artificial

<220>
<223> Descrição da Sequência Artificial: oligonucleótido sintético

<400> 635
ggccccaggg cagagggtc 19

Lisboa, 29 de Dezembro de 2011

REIVINDICAÇÕES

1. Método para produzir uma população ou biblioteca de genes de imunoglobina que compreende os passos de :

   (i) introdução de diversidade sintética em, pelo menos, uma das CDR1 VH ou CDR2 VH desses genes; e

   (ii) combinação da diversidade do passo (i) com a diversidade CDR3 VH capturada a partir de células B.

2. Método de acordo com a reivindicação 1, em que a diversidade sintética é introduzida tanto em CDR1 VH como em CDR2 VH.

3. Método de acordo com a reivindicação 2, em que a diversidade sintética compreende:

   (a) uma CDR1 VH tendo uma sequência de aminoácidos de acordo com a fórmula-X1-Y-X2-M-X3-, em que X1, X2 e X3 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; e

   (b) uma CDR2 VH, tendo uma sequência de aminoácidos de acordo com a fórmula X4-I-X5-X6-S-G-G-X7-T-X8-Y-A-D-S-V-K-G-, em que X4 e X5 são seleccionados independentemente do grupo consistindo em Y, R, W, V, G e S, X6 é seleccionado do grupo consistindo em P e S e X7 e X8 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y.

4. Biblioteca compreendendo uma colecção de pacotes genéticos que apresentam um membro de uma família diversa de péptidos, polipéptidos ou proteínas e que colectivamente apresentam, pelo menos, uma porção da família, sendo os péptidos, polipéptidos ou proteínas apresentados codificados por sequências de ADN compreendendo sequências codificando

(a) uma CDR1 VH tendo uma sequência de aminoácidos de acordo com a fórmula-X1-Y-X2-M-X3-, em que X1, X2 e X3 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y;

(b) uma CDR2 VH tendo uma sequência de aminoácidos de acordo com a fórmula X4-I-X5-X6-S-G-G-X7-T-X8-Y-A-D-S-V-K-G-, em que X4 e X5 são seleccionados independentemente do grupo consistindo em Y, R, W, V, G e S, X6 é seleccionado do grupo consistindo em P e S e X7 e X8 são seleccionados independentemente do grupo consistindo em A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y; e

(c) uma sequência codificando uma CDR3 VH, em que a referida CDR3 VH é uma CDR3 VH capturada da região CDR3 VH de um gene da imunoglobina de uma célula B.

5. Biblioteca de acordo com a reivindicação 4, em que as referidas sequências de ADN compreendem ainda uma sequência codificando uma cadeia leve de imunoglobina.

6. Biblioteca de acordo com a reivindicação 5, em que a referida sequência codificando uma cadeia leve de imunoglobina é capturada a partir de uma célula B.

7. Biblioteca de acordo com a reivindicação 6, em que a referida célula B é isolada de uma amostra sanguínea de um doente auto-imunitário.

8. Biblioteca de acordo com a reivindicação 7, em que o doente auto-imunitário é diagnosticado com um distúrbio seleccionado do grupo consistindo em lupus eritematoso sistémico, esclerose sistémica, artrite reumatóide, síndrome de antifosfolípido e vasculite.

9. Biblioteca de acordo com qualquer das reivindicações 4 ou 5 a 8, em que as referidas sequências de ADN compreendem ainda sequências codificando regiões estruturais VH 3-23.

10. Biblioteca de acordo com qualquer das reivindicações 4 ou 5 a 8, em que os referidos pacotes genéticos são fagos M13.

11. Biblioteca de acordo com a reivindicação 10, em que as referidas sequências de ADN estão num vector fágico.

12. Biblioteca de acordo com a reivindicação 10 ou 11, em que o referido fago compreende um gene de tipo selvagem iii e um gene truncado iii para apresentação de péptidos, polipéptidos ou proteínas.

13. Biblioteca de acordo com qualquer das reivindicações 4 ou 5 a 8, em que as referida sequências de ADN estão num vector fagemídico.

14. Biblioteca de acordo com qualquer das reivindicações 10 a 13, em que os referidos péptidos, polipéptidos ou proteínas

apresentados são apresentados através de um elemento de ligação curto na porção final do gene M13 III.

Lisboa, 29 de Dezembro de 2011

AMPIFICAÇÃO DE GENES VH
SEM UTILIZAR SEQUÊNCIAS VH

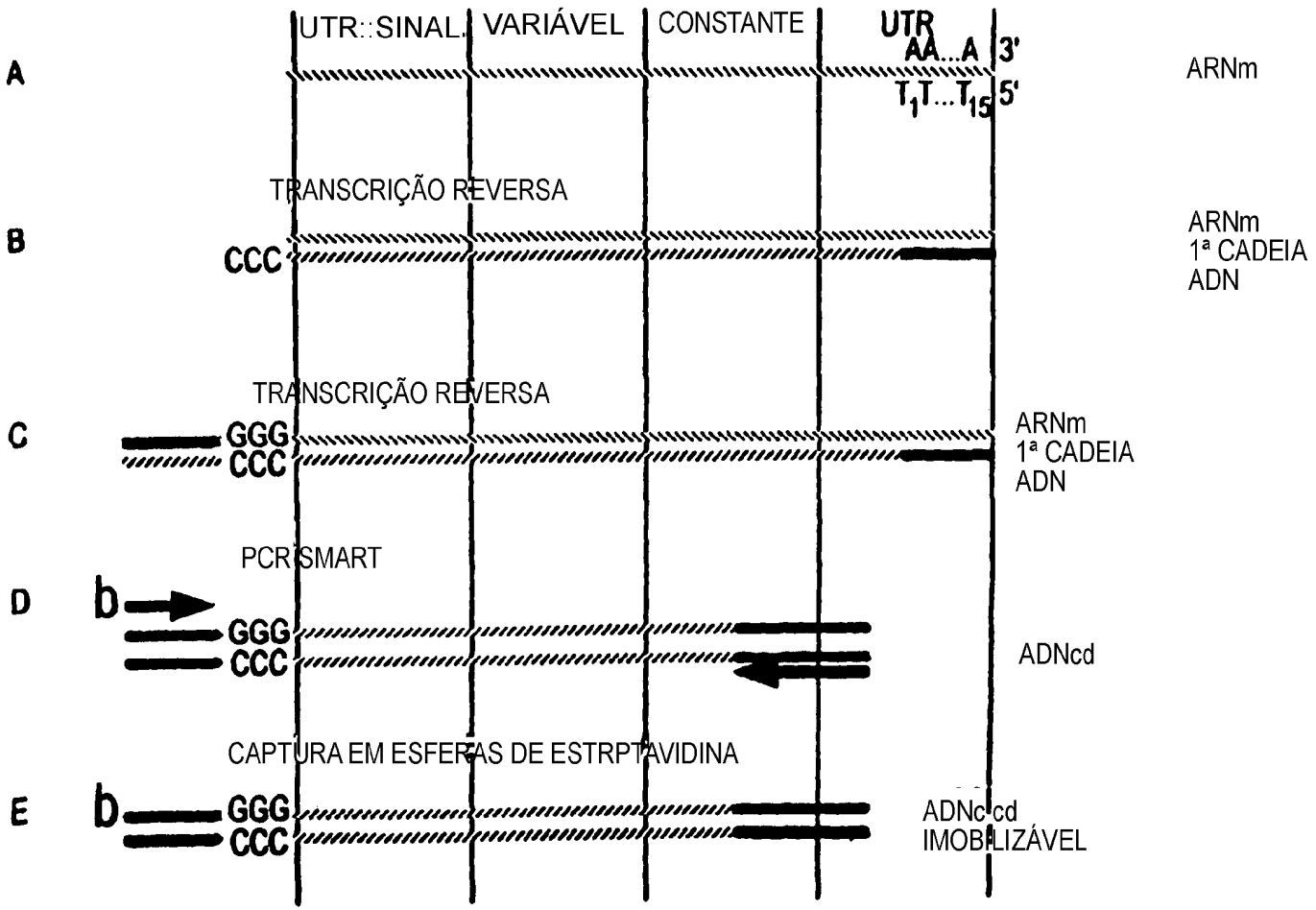


FIG. 1

AMPLIFICAÇÃO DE GENES VL SEM
UTILIZAR SEQUÊNCIAS VL

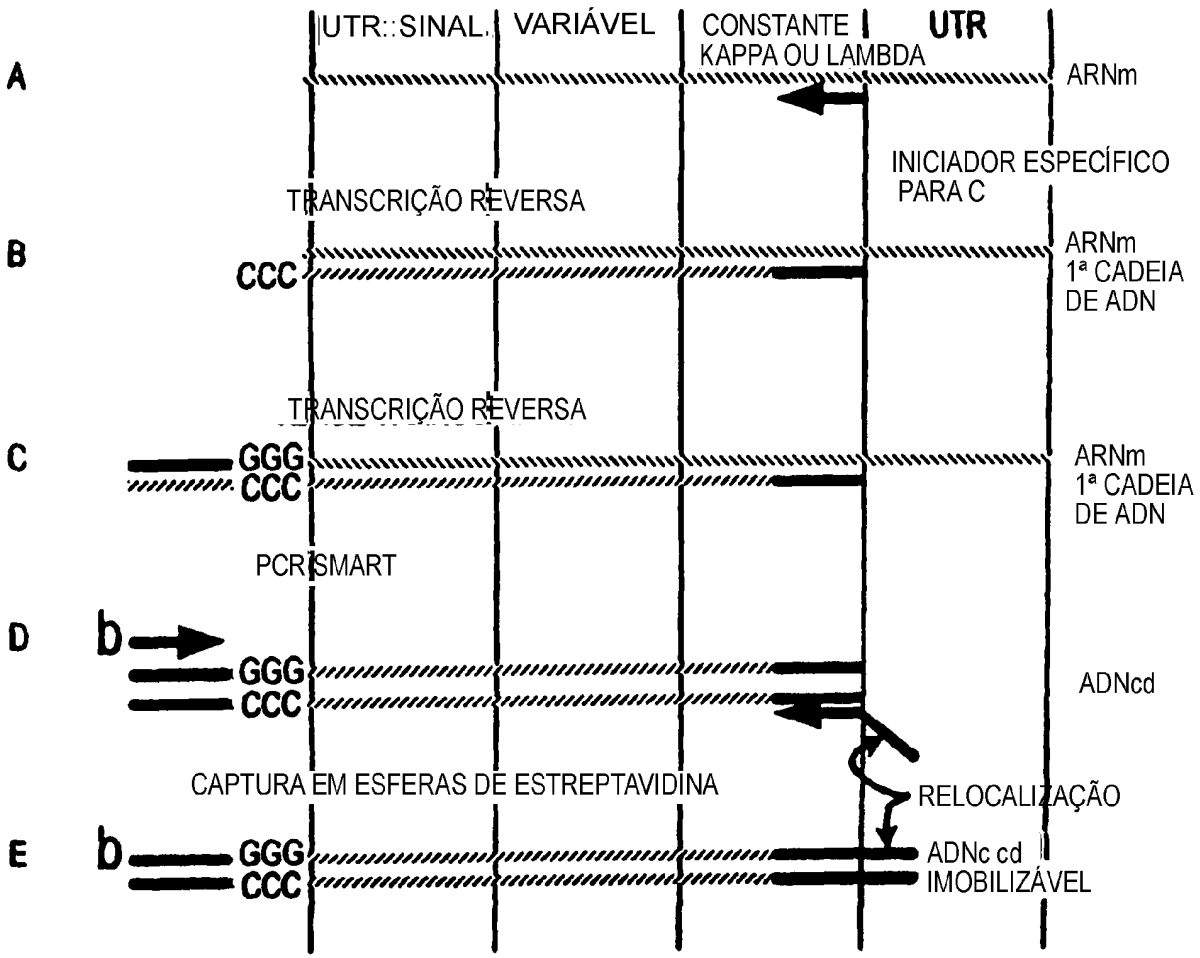


FIG. 2

Amplificação RACE não-enviesada do gene V do anticorpo

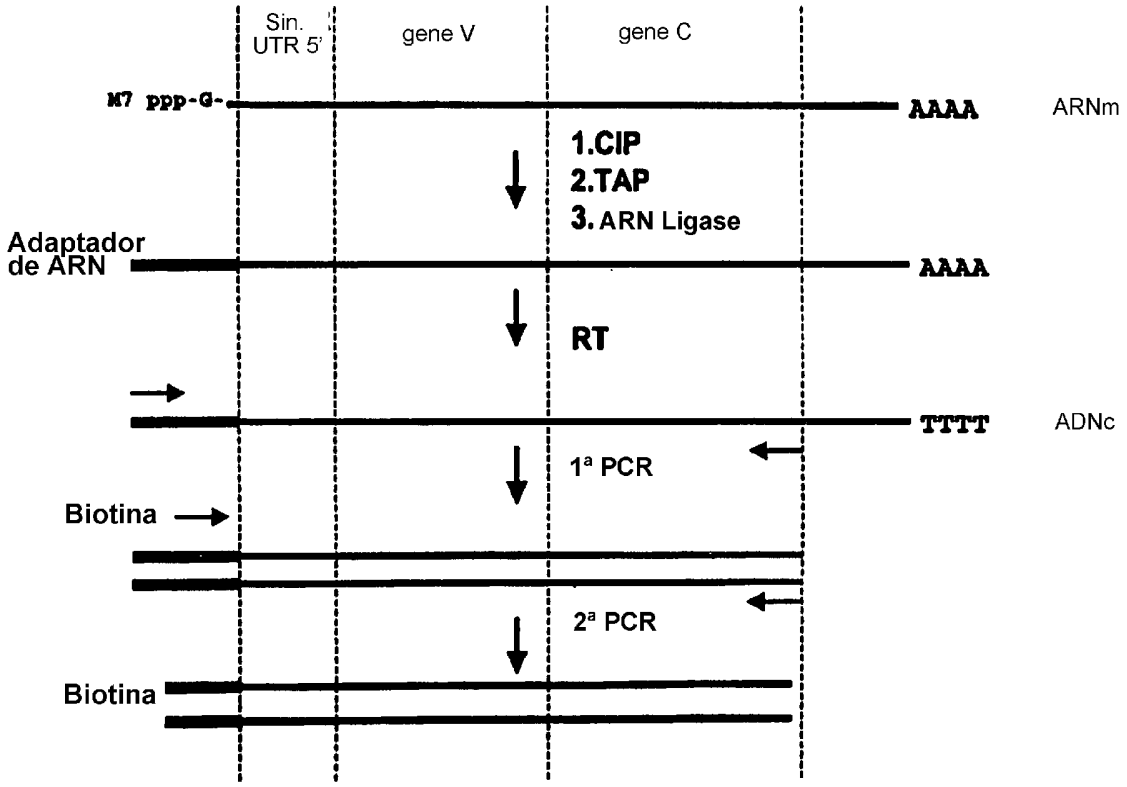
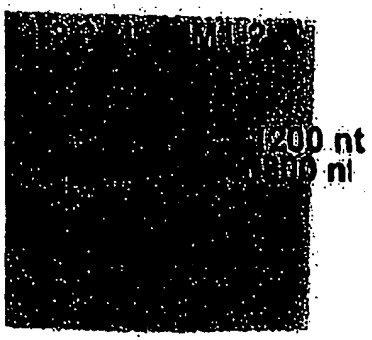


FIG. 3

1, 2, 3 e 4 são
amostras de doente

cadeias pesadas
da 1ª PCR



cadeias leves
da 1ª PCR

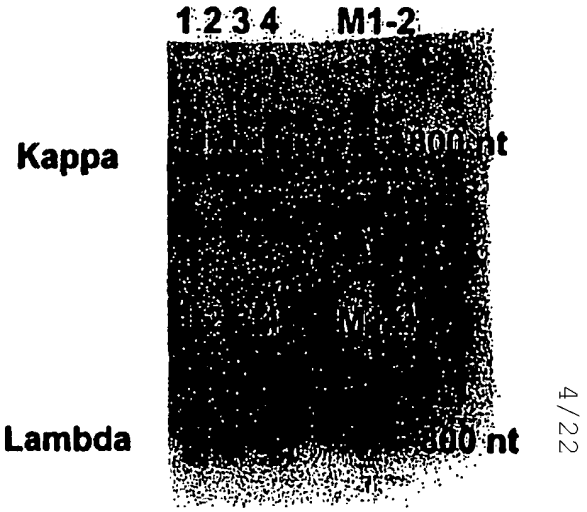


FIG. 4

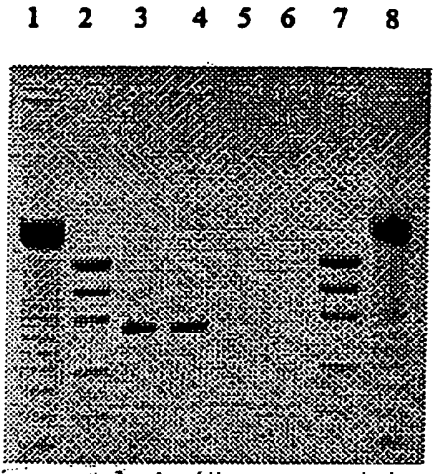
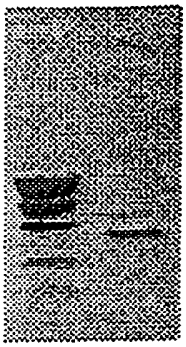


Figura 5: Análise em gel do produto de PCR de amplificação extensor-kappa Aprox. 75 ng/µL.→ 15 ng/µL

1- 100 pb
2- LDM
3- 50 ng de molde
4- 10 ng de molde
5- ADNcs não-ligado
6- Controlo negativo
7- LDM
8- 100 pb

FIG. 5

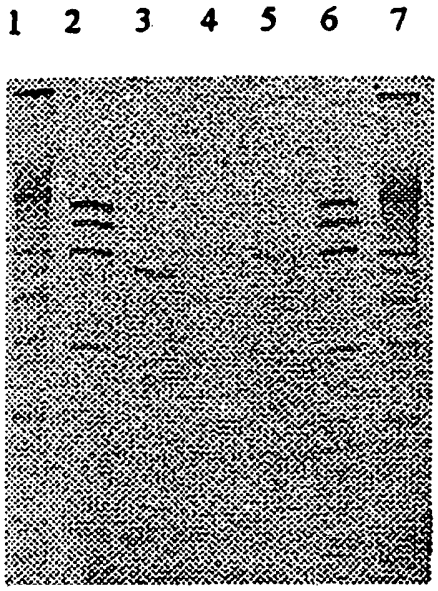
1 2



· Produto de PCR purificado em gel de amplificação extensor-kappa
Concentração: ±35 ng/µL

1- LDM
2- 1 µL purif.

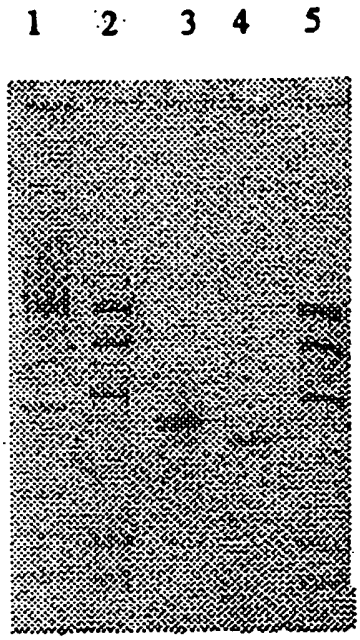
FIG. 6



Análise em gel de ADNcs-κ digerido
1 µL de ADNcs digerido ≈ 8 ng ADNcs
Volume total de 50 µL = 400 ng ADNcs
→ 400 ng ADNcs disponível para ligação das pontes-extensores

1- 100 pb
2- LDM
3- 1 µL de ADNcs puro
4 - 4µL de esferas após dig.
5- 8 µL de esferas após dig
6- Controlo negativo
7- LDM
8- 100 pb

FIG. 7

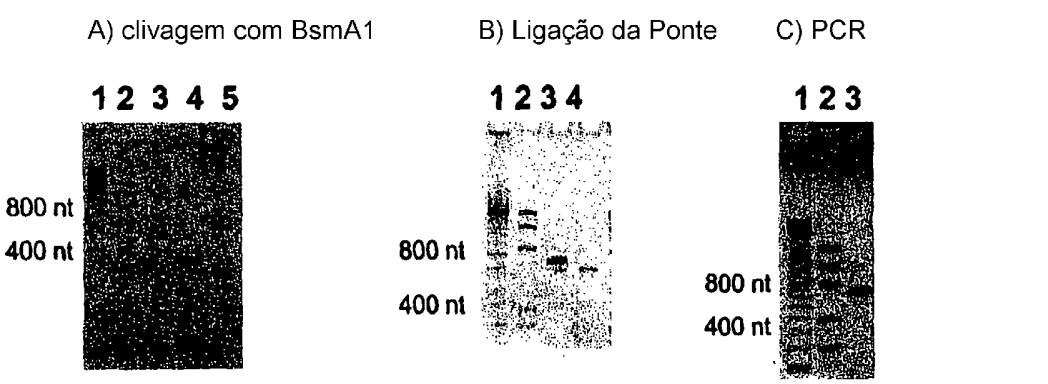


Análise em gel da ligação extensor-kappa quebrada
20 ng/5 µL de material eluído → 4 ng/µL

1- 100 pb
2- LDM
3- Mistura de ligação, 4 µL
4- ADNcs não ligado
5- LDM

FIG. 8

Clivagem e ligação de cadeias leves Kappa



1. Marcador de 100 pb
2. Marcador LDM
3. ADNcs sup. após dig.
4. esferas após dig (material não clivado)
5. ADN antes da clivagem

clivagem de 80%

1. Marcador de 100 pb
2. Marcador LDM
3. Mistura de ligação
4. ADNcs não ligado

Ligação de 90%

1. Marcador de 100 pb
2. Marcador LDM
3. 50 ng de molde (13 ciclos)

FIG. 9

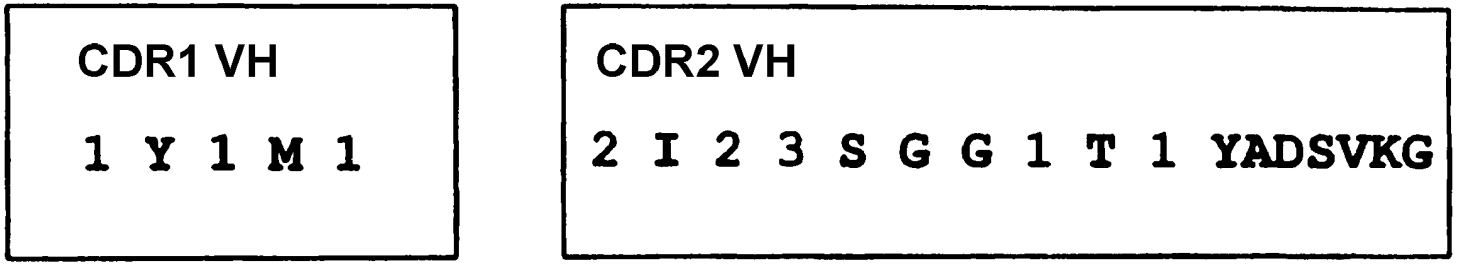


FIG. 10

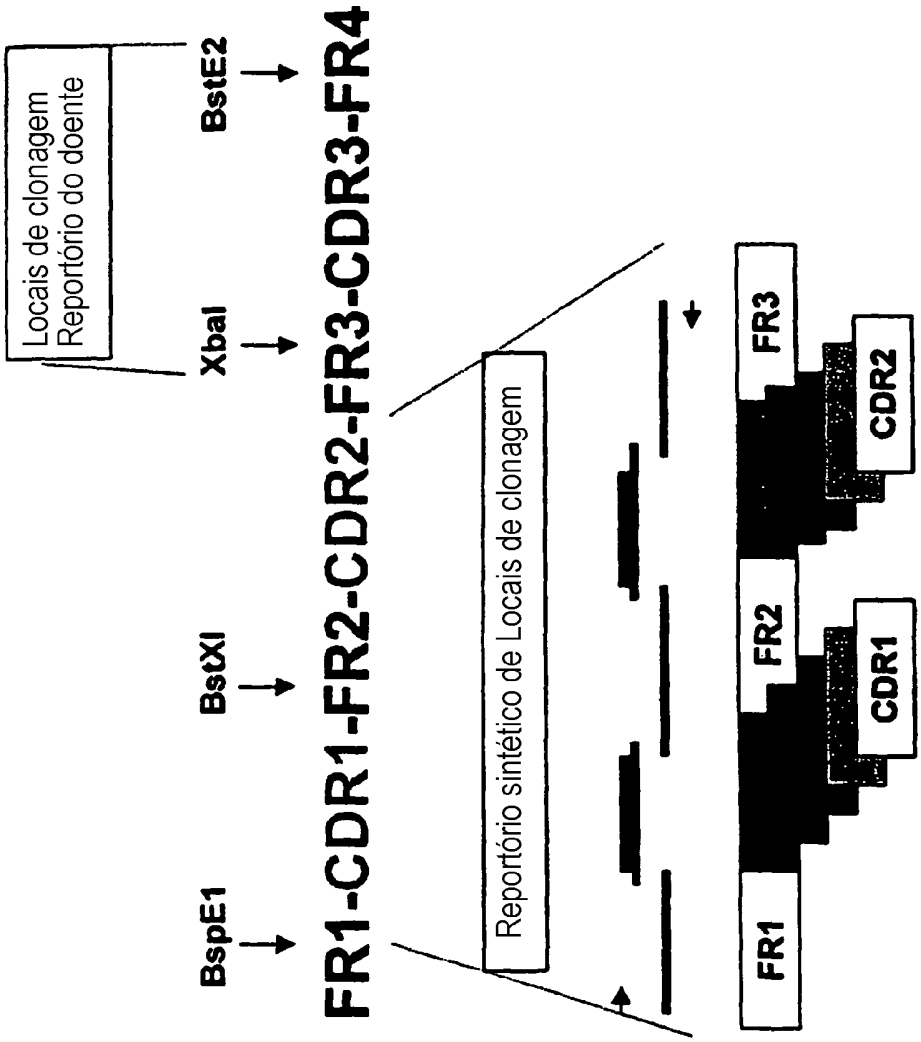

Locais de clonagem
Repertório do doente
BstE2
XbaI
BstXI
BspE1
FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4
Repertório sintético de Locais de clonagem
FR1
CDR1
FR2
CDR2
FR3

FIG. 11

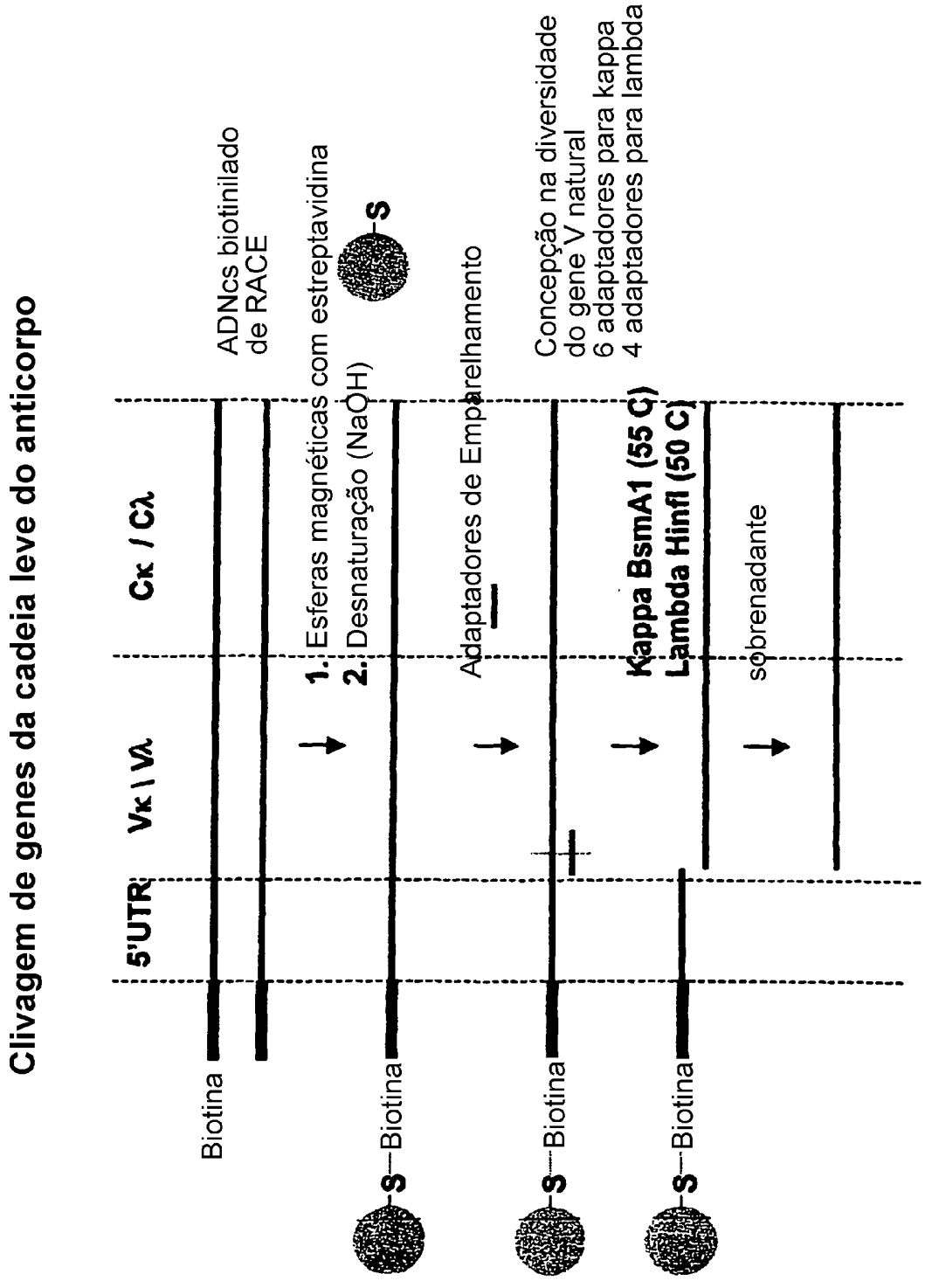

Clivagem de genes da cadeia leve do anticorpo
5'UTR
Vκ \ Vλ
Cκ / Cλ
ADNcs biotinilado de RACE
1. Esferas magnéticas com estreptavidina
2. Desnaturação (NaOH)
Adaptadores de Emparelhamento
Concepção na diversidade do gene V natural
6 adaptadores para kappa
4 adaptadores para lambda
Kappa BsmA1 (55 C)
Lambda HinfI (50 C)
sobrenadante
Biotina
S
Biotina
S
Biotina
S
Biotina
S

FIG. 12A

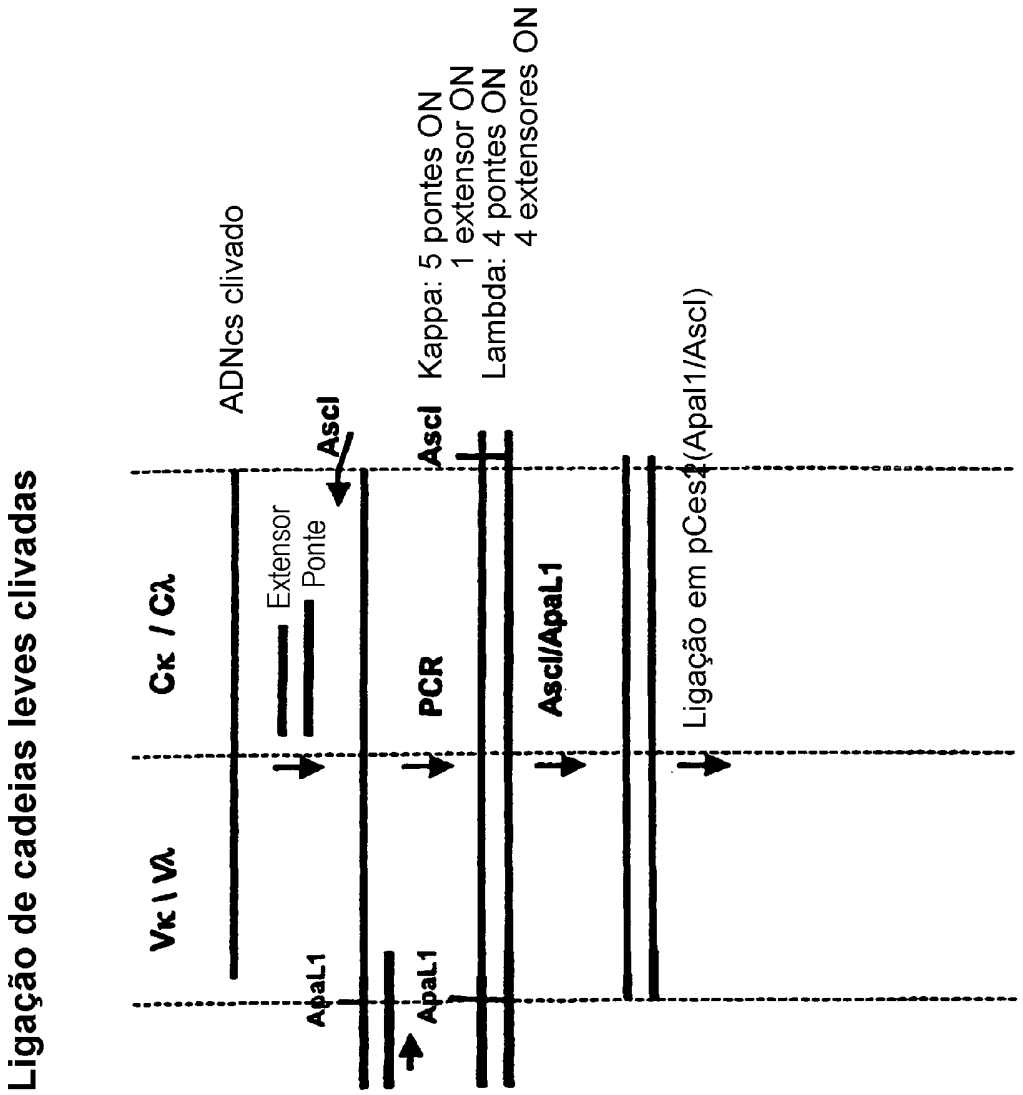
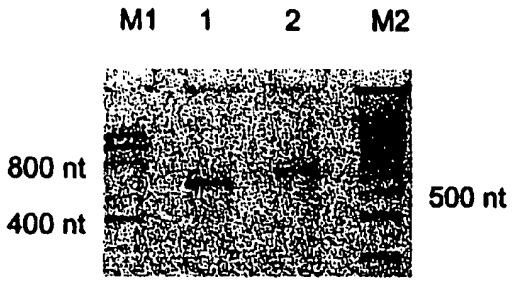

Ligação de cadeias leves clivadas
Vκ I Vλ
Cκ / Cλ
ADNcs clivado
Extensor
Ponte
AscI
ApaL1
ApaL1
PCR
AscI
Kappa: 5 pontes ON
1 extensor ON
Lambda: 4 pontes ON
4 extensores ON
AscI/ApaL1
Ligação em pCes1 (ApaI1/AscI)

FIG. 12B

Figura 3: Clivagem e ligação de cadeias leves lambda

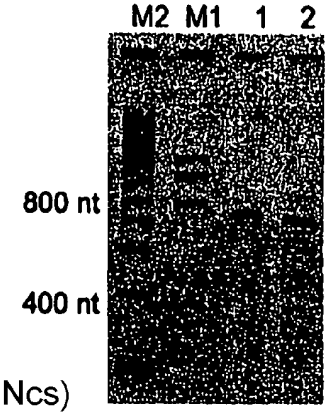
A) Clivagem com HinfI



M1 = Marcador PM LDM
1 = sup. após clivagem com Hinf (10 ng de ADNcs)
2 = esferas após clivagem com Hinf
(10 ng ADNcs(não clivado)
M2 = Marcador de 100 pb, de baixo para cima:
300 nt, 400 nt, 500 nt, 600 nt, 800 nt

Clivagem de 70%

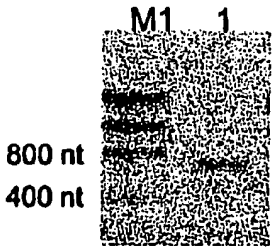
B) Ligação da ponte



1= ADN ligado
2= ADN não-ligado

Ligação de 95%

C) PCR



1 = 1 uL de produto
de PCR purificado/50 uL
(15 ciclos)

FIG. 13

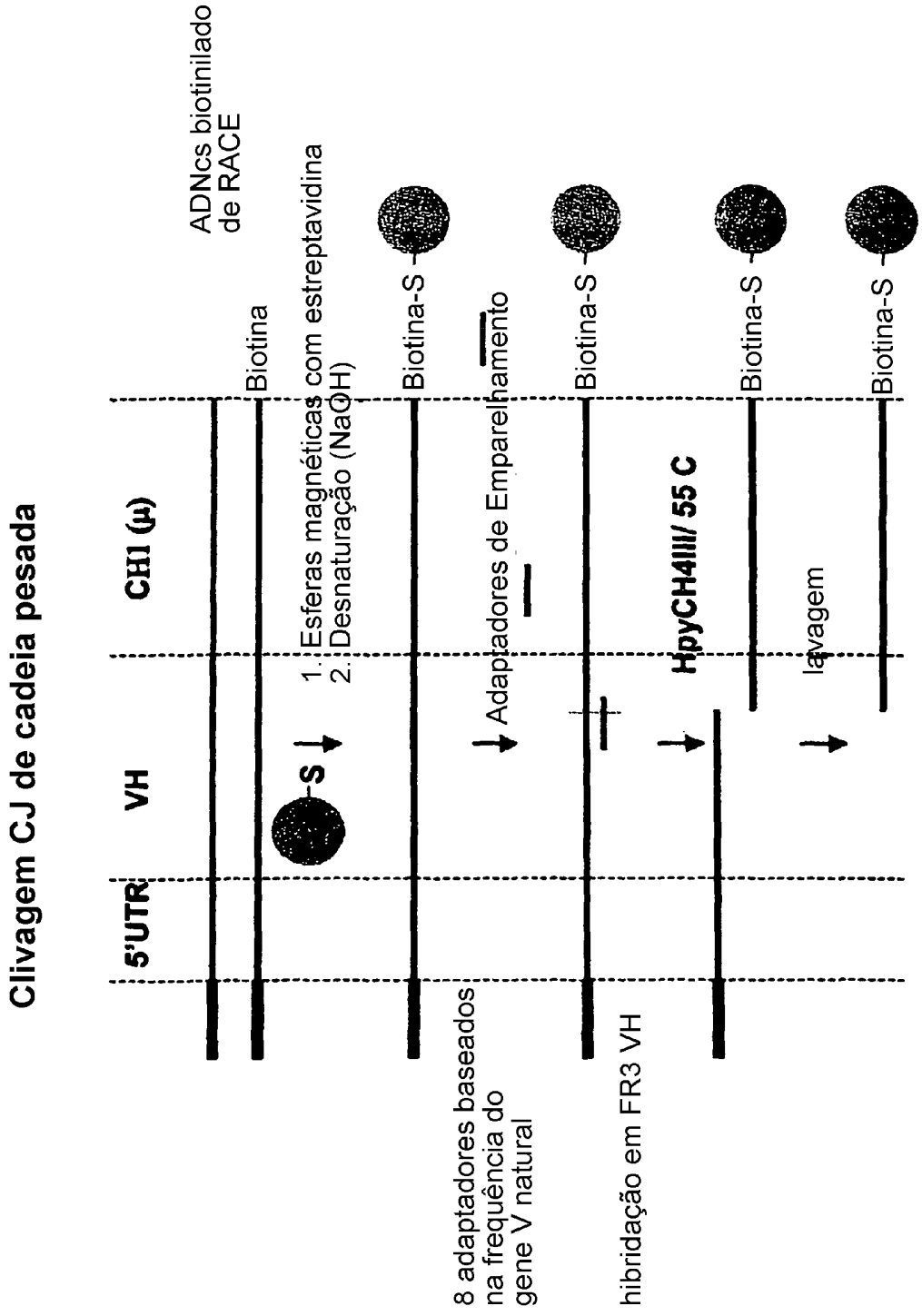

Clivagem CJ de cadeia pesada
5'UTR
VH
CH1 (μ)
ADNcs biotinilado de RACE
Biotina
1. Esferas magnéticas com estreptavidina
2. Desnaturação (NaOH)
S
Biotina-S
Adaptadores de Emparelhamento
8 adaptadores baseados na frequência do gene V natural
hibridação em FR3 VH
HpyCH4III/ 55 C
lavagem

FIG.14A

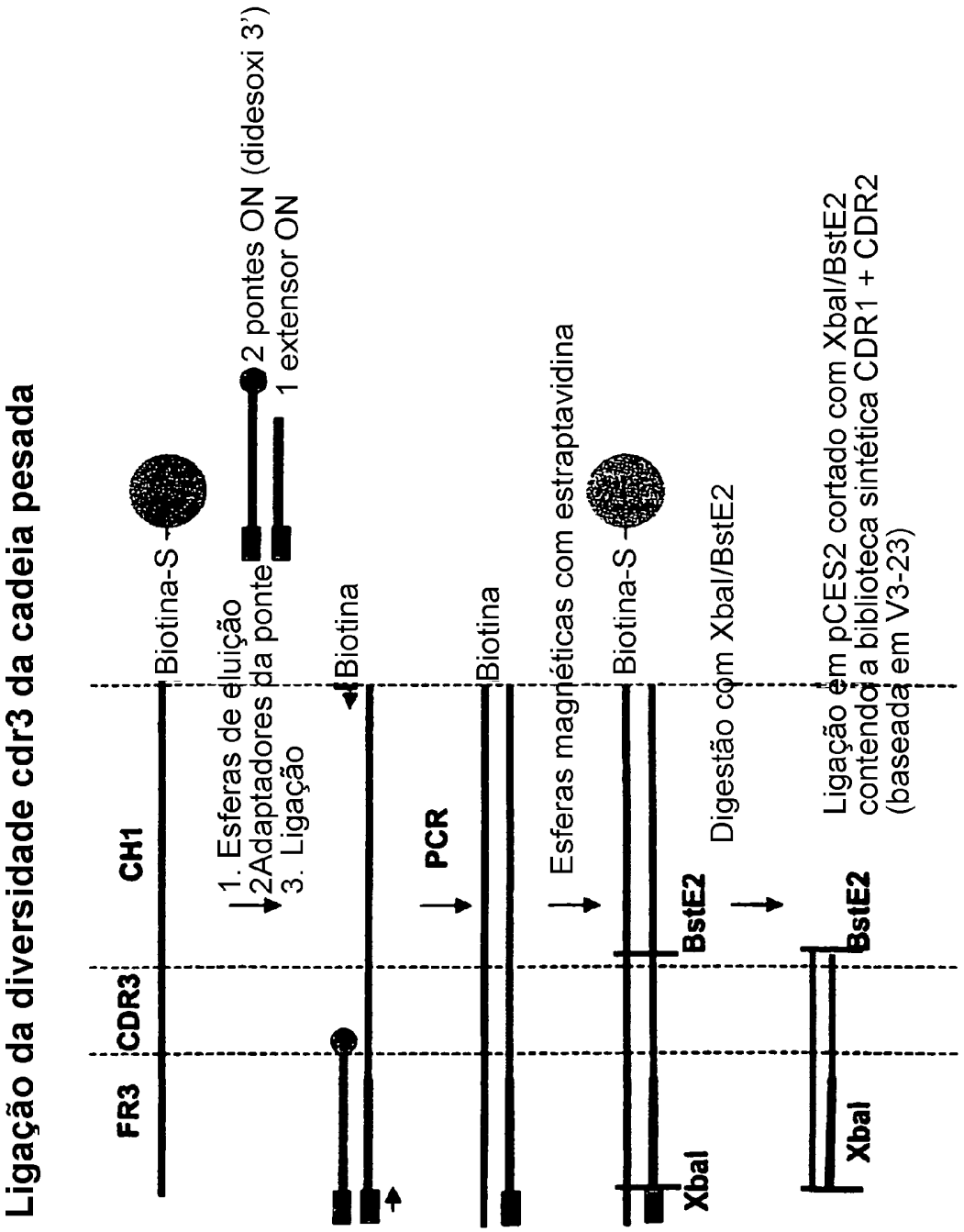
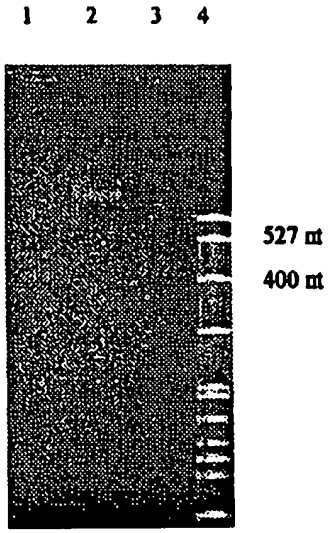

Ligação da diversidade cdr3 da cadeia pesada
FR3
CDR3
CH1
Biotina-S
1. Esferas de eluição
2Adaptadores da ponte
3. Ligação
2 pontes ON (didesoxi 3')
1 extensor ON
Biotina
PCR
Biotina
Esferas magnéticas com estraptavidina
Biotina-S
BstE2
Digestão com XbaI/BstE2
XbaI
BstE2
XbaI
Ligação em pCES2 cortado com XbaI/BstE2
contendo a biblioteca sintética CDR1 + CDR2
(baseada em V3-23)

FIG. 14B

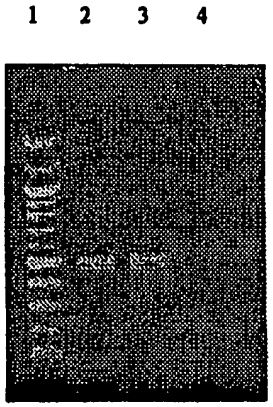
Clivagem e ligação da Cadeia Pesada

A) Clivagem com HpyCH4III



1 = ADN clivado eluído a partir da coluna PN
2 = Esferas após digestão com HpyCH4III
3 = Sobrenadante após clivagem
4 = Digestão de pBR322 com MspI

B) PCR



1 = Marcador de 100 pb NEB
2 = 5 uL/produto de PCR de 100 uL 20 ciclos; amostra A
3 = 5 uL/produto de PCR de 100 uL 20 ciclos; amostra B
4 = Sem molde

FIG. 15

Fago DyF31 (Sistema 3 + 3)

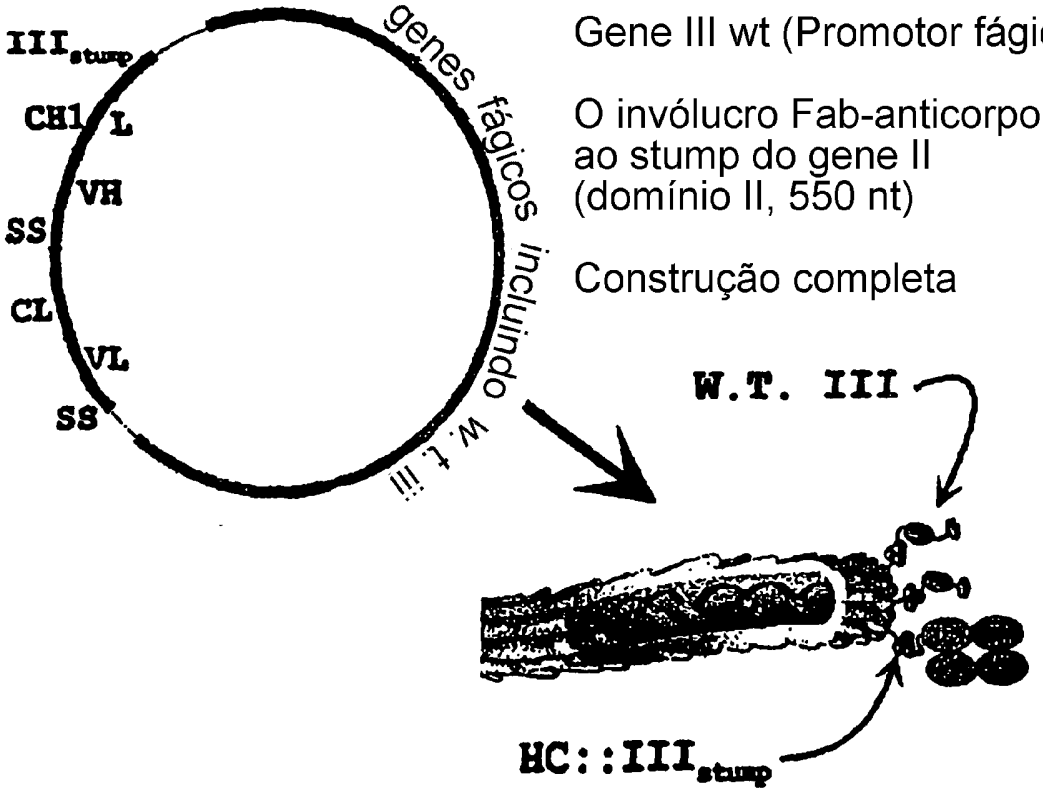


FIG. 16

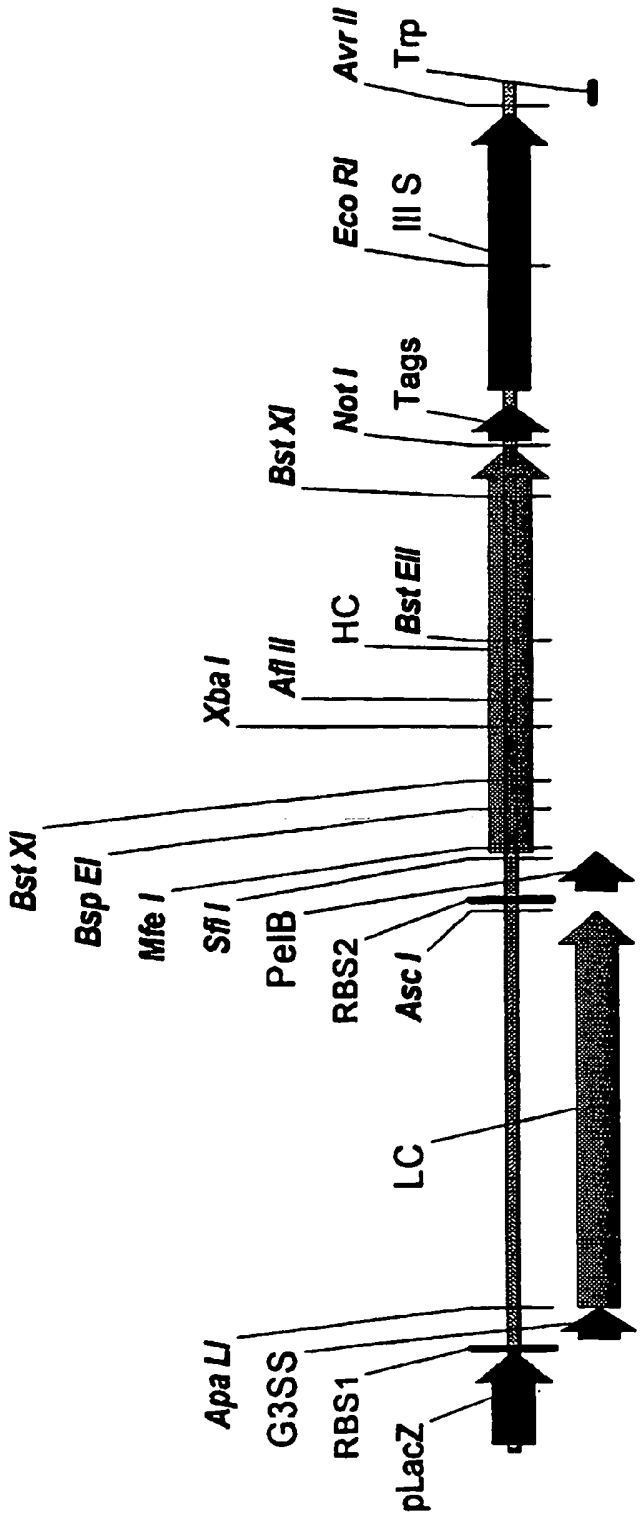

Apa LI
G3SS
RBS1
pLacZ
Bst XI
Bsp EI
Mfe I
Sfi I
PelB
RBS2
Asc I
LC
Xba I
Afl II
HC
Bst EII
Bst XI
Not I
Tags
Eco RI
III S
Avr II
Trp
Invólucro Fab
2263 pb

FIG. 17

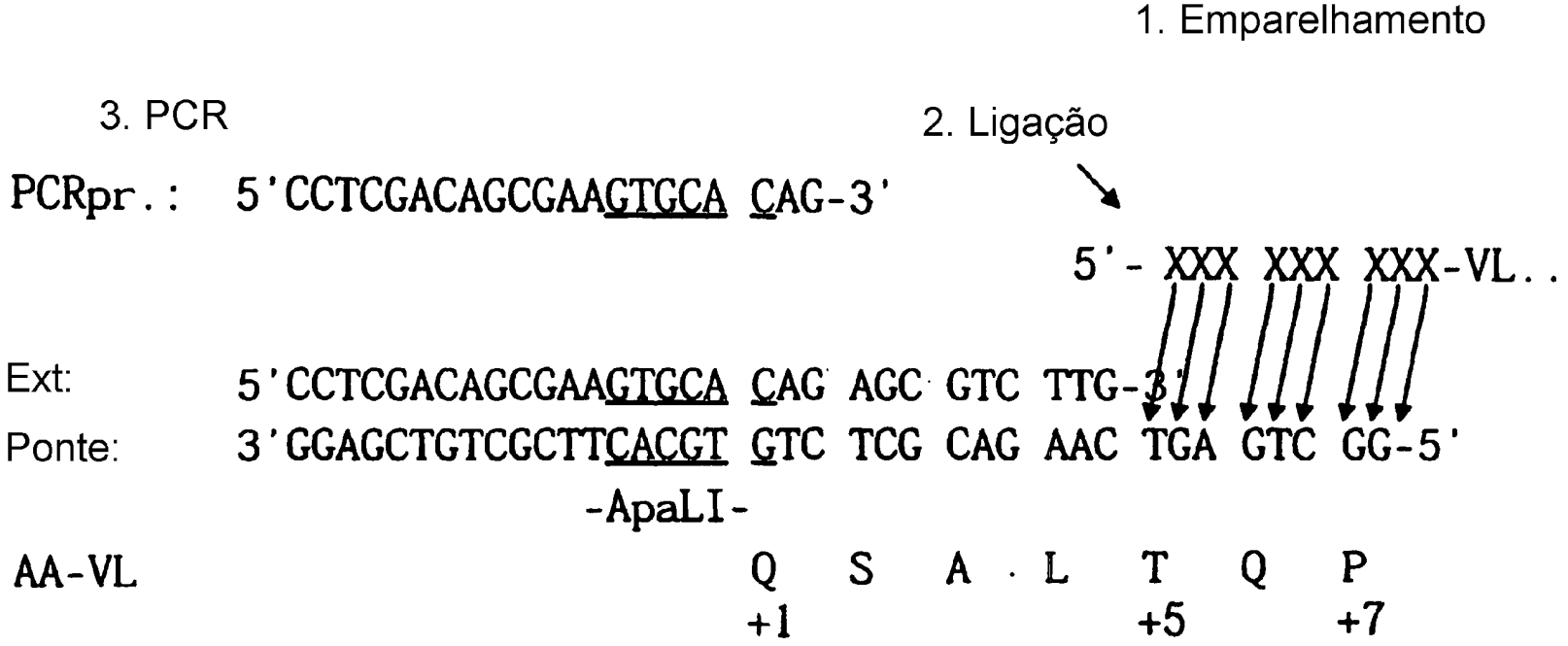

1. Emparelhamento
3. PCR
2. Ligação
PCRpr.: 5'CCTCGACAGCGAAGTGCA CAG-3'
5'- XXX XXX XXX-VL..
Ext: 5'CCTCGACAGCGAAGTGCA CAG AGC GTC TTG-3'
Ponte: 3'GGAGCTGTCGCTTCACGT GTC TCG CAG AAC TGA GTC GG-5'
-ApaLI-
AA-VL Q S A L T Q P
+1 +5 +7

FIG. 18

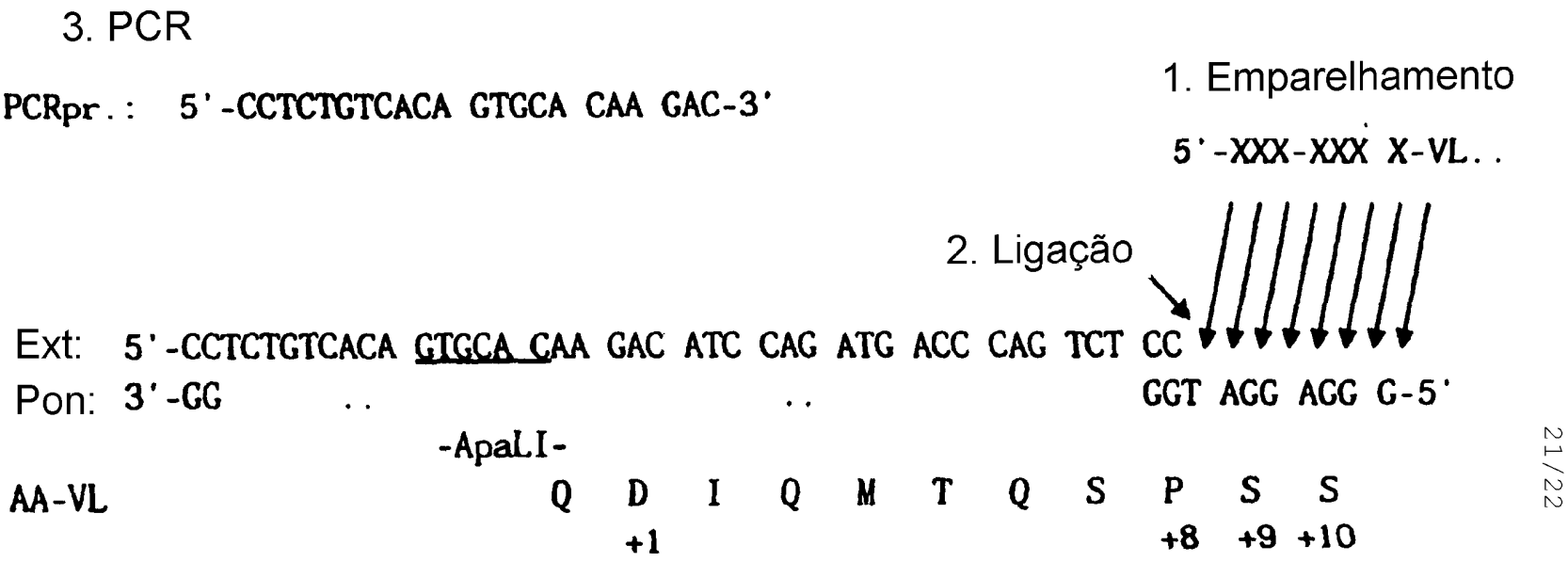

3. PCR
PCRpr.: 5'-CCTCTGTCACA GTGCA CAA GAC-3'
1. Emparelhamento
5'-XXX-XXX X-VL..
2. Ligação
Ext: 5'-CCTCTGTCACA GTGCA CAA GAC ATC CAG ATG ACC CAG TCT CC
Pon: 3'-GG .. .. GGT AGG AGG G-5'
-ApaLI-
AA-VL Q D I Q M T Q S P S S
+1 +8 +9 +10

FIG. 19

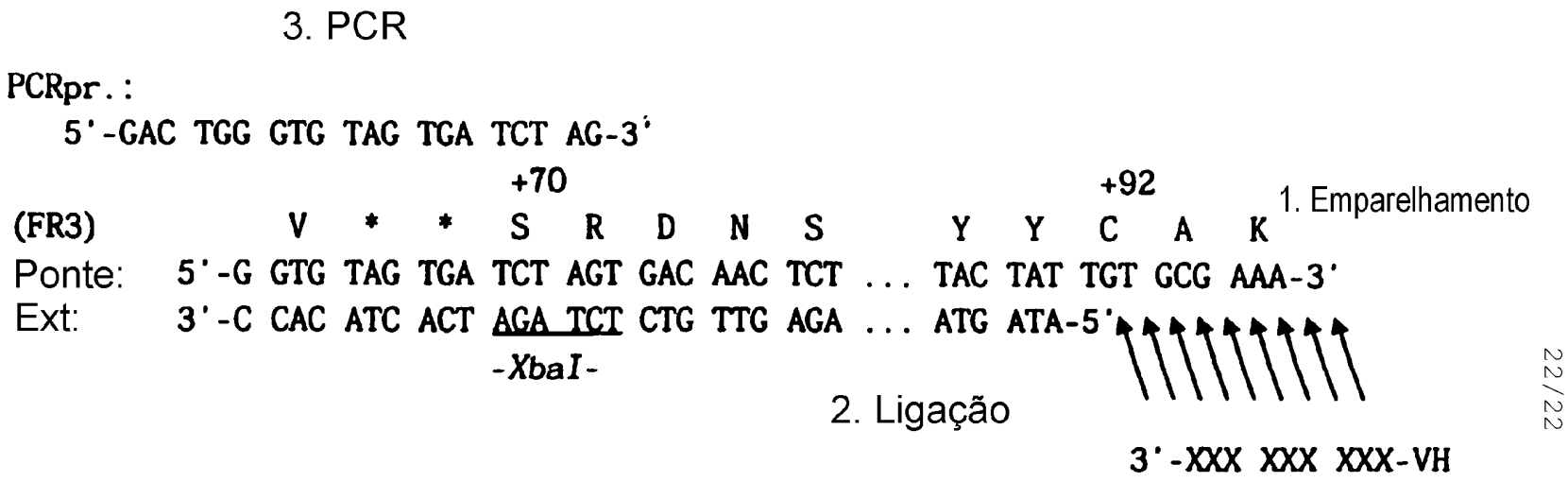

3. PCR
PCRpr.:
5'-GAC TGG GTG TAG TGA TCT AG-3'
+70
+92
1. Emparelhamento
(FR3) V * * S R D N S Y Y C A K
Ponte: 5'-G GTG TAG TGA TCT AGT GAC AAC TCT ... TAC TAT TGT GCG AAA-3'
Ext: 3'-C CAC ATC ACT AGA TCT CTG TTG AGA ... ATG ATA-5'
-XbaI-
2. Ligação
3'-XXX XXX XXX-VH

FIG. 20