

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

(22) Data de pedido: 2001.02.16	(73) Titular(es): MAHIDOL UNIVERSITY 999 PHUTTAMONTOL ROAD AMPHUR PHUTTHAMONTOL NAKORNPATTHOM 73170 TH THE GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH & HUMAN SERVICES US
(30) Prioridade(s): 2000.02.16 US 182829 P	
(43) Data de publicação do pedido: 2011.01.26	
(45) Data e BPI da concessão: 2015.08.26 234/2015	(72) Inventor(es): RICHARD M. KINNEY US CLAIRE Y. H. KINNEY US SIRITORN BUTRAPET TH DUANE L. GUBLER US NATTH BHAMARAPRAVATI TH
	(74) Mandatário: JOÃO LUÍS PEREIRA GARCIA RUA CASTILHO, 167 2º 1070-050 LISBOA PT

(54) Epígrafe: **QUIMERAS IMUNOGÉNICAS DO VÍRUS DO DENGUE 2**

(57) Resumo:

SÃO FORNECIDOS FLAVIVÍRUS QUIMÉRICOS QUE SÃO NÃO VIRULENTOS E IMUNOGÉNICOS. OS VÍRUS QUIMÉRICOS SÃO CONSTRUÍDOS PARA CONTEREM MUTAÇÕES DE AMINOÁCIDOS NAS PROTEÍNAS NÃO ESTRUTURAIS DE UM FLAVIVÍRUS. OS VÍRUS QUIMÉRICOS QUE CONTÊM OS GENES NÃO ESTRUTURAIS MUTADOS POR ATENUAÇÃO DO VÍRUS SÃO USADOS COMO UMA ESTRUTURA NA QUAL SÃO INSERIDOS OS GENES DE PROTEÍNAS ESTRUTURAIS DE UMA SEGUNDA ESTIRPE DE FLAVIVÍRUS. ESTES VÍRUS QUIMÉRICOS INDUZEM UMA IMUNOGENICIDADE PRONUNCIADA MAS NÃO POSSUEM OS SINTOMAS CLÍNICOS QUE ACOMPANHAM A DOENÇA VIRAL. OS VÍRUS QUIMÉRICOS ATENUADOS SÃO EFICAZES COMO IMUNOGÉNICOS OU VACINAS E PODEM SER COMBINADOS NUMA COMPOSIÇÃO FARMACÊUTICA PARA CONFERIR IMUNIDADE SIMULTÂNEA CONTRA VÁRIAS ESTIRPES PATOGÉNICAS DE FLAVIVÍRUS.

RESUMO

QUIMERAS IMUNOGÉNICAS DO VÍRUS DO DENGUE 2

São fornecidos flavivírus quiméricos que são não virulentos e imunogênicos. Os vírus quiméricos são construídos para conterem mutações de aminoácidos nas proteínas não estruturais de um flavivírus. Os vírus quiméricos que contêm os genes não estruturais mutados por atenuação do vírus são usados como uma estrutura na qual são inseridos os genes de proteínas estruturais de uma segunda estirpe de flavivírus. Estes vírus quiméricos induzem uma imunogenicidade pronunciada mas não possuem os sintomas clínicos que acompanham a doença viral. Os vírus quiméricos atenuados são eficazes como imunogênicos ou vacinas e podem ser combinados numa composição farmacêutica para conferir imunidade simultânea contra várias estirpes patogênicas de flavivírus.

DESCRIÇÃO

QUIMERAS IMUNOGÉNICAS DO VÍRUS DENGUE 2

CAMPO DA INVENÇÃO

A presente invenção relaciona-se com os campos da imunologia e da virologia e, em particular, com quimeras de flavivírus imunogénicas, não virulentas para a produção de vacinas de flavivírus atenuados, vivos, imunogénicos.

ANTECEDENTES DA INVENÇÃO

Os vírus do dengue são patógenos transmitidos por mosquitos do género *Flavivirus* (família *Flaviviridae*). Foram identificados quatro serotipos do vírus do dengue (muitas vezes abreviado como "DEN"), incluindo o dengue-1, o dengue-2, o dengue-3 e o dengue-4 (DEN-1 a DEN-4). O genoma do flavivírus é um ARN de cadeia simples, de sentido positivo de aproximadamente 11 kb de comprimento, que contém uma região não codificante 5' (5'NC), uma região codificante que codifica para as proteínas estruturais virais, cinco proteínas não estruturais, designadas por NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5, e uma região não codificante 3' (3'NC). As proteínas estruturais virais incluem a cápside, a pré-membrana / membrana e o invólucro. As proteínas estruturais e não estruturais são traduzidas como uma única poliproteína. A poliproteína é então processada por proteases celulares e virais.

Transmitidos pelos mosquitos *Aedes aegypti* para os seres humanos em regiões tropicais e subtropicais do mundo, os vírus do dengue causam milhões de casos de doença cada ano, que vão desde a febre de dengue até à febre hemorrágica de dengue / síndrome de choque de dengue muitas vezes fatal (FHD / SCD). A infeção secundária de seres humanos com um serotipo heterólogo do vírus do DEN

pode induzir uma resposta imunopatológica e é considerada um possível fator de risco para FHD / SCD. Portanto, existe a necessidade do desenvolvimento de uma vacina que confere proteção simultânea contra todas as estirpes do vírus do dengue.

Uma vez que a erradicação de mosquitos *Aedes aegypti* parece ser praticamente inviável, o desenvolvimento de vacinas seguras e eficazes contra os quatro serotipos do vírus do dengue é uma prioridade da Organização Mundial de Saúde. No entanto, não está atualmente disponível nenhuma vacina aprovada, eficaz contra qualquer uma das estirpes do vírus do dengue. Tem sido demonstrado que a passagem em série de flavivírus do tipo selvagem em várias culturas de células, tais como células primárias de rim de cão (PDK), produz variantes do vírus que têm virulência reduzida, retêm a imunogenicidade e não produzem sintomas clínicos adversos.

Foram desenvolvidos vírus do dengue vivos, atenuados de todos os quatro serotipos na Universidade de Mahidol, na Tailândia por passagem dos vírus do tipo selvagem em cultura de células. Estes são atualmente os mais promissores candidatos a vacinas vivas, atenuadas para a imunização contra a infecção e/ou doença do vírus do dengue. Estes candidatos a vacinas têm sido designados por uma combinação do seu serotipo de dengue, da linha de células através da qual foram passados e do número de vezes que foram passados. Assim, um vírus do dengue do serotipo 1 do tipo selvagem passado 13 vezes em células PDK é designado por vírus DEN-1 PDK-13. As outras vacinas candidatas são DEN-2 PDK-53 (sequência de nucleótidos, SEQ ID NO: 15; sequência de aminoácidos, SEQ ID NO: 16), DEN-3 PGMK-30 / FRhL-3 (trinta passagens em células primárias de rim de macaco verde, seguidas por três passagens em células de pulmão de rhesus fetal) e DEN-4 PDK-48. Estes quatro vírus candidatos a vacina foram obtidos por

passagem em cultura de tecidos de vírus do tipo selvagem parental DEN-1 16007, DEN-2 16681 (sequência de nucleótidos, SEQ ID NO: 13; sequência de aminoácidos, SEQ ID NO: 14), DEN-3 16562 e DEN-4 1036, respectivamente.

Os ensaios clínicos humanos preliminares com estes vírus atenuados têm indicado que o DEN-2 PDK-53 tem a menor dose infecciosa (dose infecciosa mínima de 50 % de 5 unidades formadoras de placa ou ufp) em seres humanos, é altamente imunogénico e não produz sintomas clínicos inaceitáveis. Os vírus candidatos a vacina DEN-1 PDK-13, DEN-3 PGMK-30 / FRhL-3 e DEN-4 PDK-48 têm doses mínimas infecciosas de 50 % mais elevadas de 10.000, 3500, e 150 ufp, respectivamente, em seres humanos. As maiores doses infecciosas necessárias para os três últimos candidatos a vacinas levantam preocupações no que respeita à eficácia relativa de cada componente do serotipo numa vacina tetravalente do vírus do dengue. Embora apenas tenha sido necessária uma imunização com o vírus DEN-2 PDK-53 monovalente ou o vírus DEN-4 PDK-48 para alcançar 100 % de seroconversão em indivíduos humanos, foi necessário um reforço para atingir a mesma taxa de seroconversão para os vírus DEN-1 PDK-13 e DEN-3 PGMK-30 / FRhL-3, que têm as duas maiores doses infecciosas para seres humanos.

O vírus DEN-2 PDK-53 candidato a vacina, daqui em diante abreviado como PDK-53, tem vários marcadores biológicos mensuráveis associados à atenuação, incluindo a sensibilidade à temperatura, o tamanho pequeno de placa, a baixa replicação em cultura de células C6/36 de mosquito, a baixa replicação em mosquitos intactos, a perda de neurovirulência em ratinhos lactentes e a baixa incidência de viremia em macacos. Os ensaios clínicos do candidato PDK-53 a vacina demonstraram a sua segurança e imunogenicidade em seres humanos. Para além disso, a vacina de PDK-53 induz respostas de memória de células T específicas do vírus do dengue em seres humanos recetores

da vacina.

Exceto para o vírus DEN-2 PDK-53, são desconhecidos o número e a identidade das mutações genéticas que se acumulam durante várias passagens em cultura de células e que estão associadas com os fenótipos atenuados dos candidatos a vacinas. Não são conhecidos para qualquer um destes quatro candidatos a vacina nem as contribuições relativas de tais mutações associadas à atenuação para o mecanismo real de atenuação, nem o potencial de mutações inversas para reverter qualquer dos candidatos a vacinas para o fenótipo biológico virulento do vírus do dengue do tipo selvagem. Uma compreensão dos marcadores de atenuação de um candidato a vacina é crítica para a previsão da sua estabilidade e segurança.

Consequentemente, existe uma necessidade de vírus do dengue não virulento, mas imunogénico, para ser utilizado no desenvolvimento de vacinas de vírus do dengue para conferir proteção contra um ou mais serotipos do vírus do dengue. O que seria ideal é uma vacina que proteja simultaneamente um indivíduo contra várias estirpes virulentas desta família potencialmente perigosa (Flaviviridae) de vírus. Portanto, é particularmente necessária uma vacina tetravalente que possa ser utilizada para imunizar um indivíduo contra todos os quatro serotipos de dengue. Têm sido descritos vírus quiméricos para utilização em vacinas. O documento W093/06214 descreve um vírus quimérico para utilização numa preparação de vacina que possui um genoma que compreende uma proteína estrutural de um flavivírus e sequências de ácidos nucleicos que codificam para uma proteína não estrutural de outro flavivírus. Chambers et al., (1999) J Virol, vol. 73, n.º 4, pág. 3095-3101, descreve que para uma quimera de ácido nucleico de vírus da febre amarela / encefalite japonesa (YF / JE) não foi obtido nenhum vírus viável quando toda a região estrutural foi derivada da JE.

O documento W098/37911 mostra a produção de uma quimera de ácido nucleico de YF / JE (prM-E).

SUMARIO DA INVENÇÃO

A presente invenção é definida nas reivindicações que a acompanham. São descritas quimeras de flavivírus imunogénicas, uma estrutura de vírus do dengue-2 para a preparação das quimeras de flavivírus e os métodos para a produção das quimeras de flavivírus. As quimeras de flavivírus imunogénicas são fornecidas, por si só ou em combinação, num transportador farmacêuticamente aceitável na forma de composições imunogénicas para minimizar, inibir, ou imunizar indivíduos contra a infeção por um ou mais flavivírus ou estirpes de flavivírus, em particular as estirpes do vírus do dengue de serotipos DEN-1, DEN-2, DEN-3 e DEN-4. Quando combinadas, as quimeras de flavivírus imunogénicas podem ser utilizadas como vacinas multivalentes para conferir proteção simultânea contra a infeção por mais do que uma espécie ou estirpe de flavivírus. De um modo preferido, as quimeras de flavivírus são combinadas numa composição imunogénica útil como uma vacina tetravalente contra os quatro serotipos de vírus do dengue conhecidos. A sequência de ácido nucleico para cada um dos vírus DEN-1, DEN-3 e DEN-4 pode ser utilizada como sondas para detetar o vírus do dengue numa amostra biológica.

As quimeras de flavivírus imunogénicas, não virulentas aqui descritas contêm os genes de proteínas não estruturais do vírus do dengue-2 atenuado, ou o equivalente dos mesmos, e um ou mais dos genes de proteínas estruturais ou de porções imunogénicas dos mesmos do flavivírus contra a qual é conferida a imunogenicidade. Uma quimera pode conter o genoma de PDK-53 do vírus do dengue-2 atenuado como estrutura viral, e os genes de proteínas estruturais que codificam para a pré-membrana / membrana, ou invólucro do genoma de PDK-53,

ou combinações dos mesmos, são substituídos com os genes de proteínas estruturais correspondentes de um flavivírus, contra o qual será feita a proteção, tal como um flavivírus diferente ou um serotipo do vírus do dengue do diferente. A quimera viral resultante tem as propriedades funcionais do vírus do dengue-2 atenuado e é, por conseguinte, não virulenta, mas expressa epítomos antigénicos dos produtos dos genes estruturais e é, portanto, imunogénica.

Uma quimera pode ser uma quimera de ácido nucleico que compreende uma primeira sequência de nucleótidos que codifica para proteínas não estruturais e uma proteína da cápside (C) de um vírus do dengue-2 atenuado, e uma segunda sequência de nucleótidos que codifica para uma proteína estrutural de um segundo flavivírus. O vírus do dengue-2 atenuado pode ser uma estirpe de vacina PDK-53. A proteína estrutural pode ser a proteína prM ou E de um flavivírus. Os exemplos de flavivírus a partir dos quais a proteína estrutural pode ser selecionada incluem, mas não estão limitados a, vírus do dengue-1, vírus do dengue-2, vírus do dengue-3, vírus do dengue-4 ou vírus do Nilo Ocidental.

Ao avaliar o vírus quimérico aqui descrito, descobriu-se inesperadamente que a não virulência da estirpe PDK-53 do vírus atenuado é atribuível à presença de mutações de substituição de aminoácidos específicos nas proteínas não estruturais e uma mutação de substituição de nucleótidos na região não codificante 5'. Esta mutação por substituição de nucleótidos ocorre na haste de uma estrutura de haste e ansa que é conservada em todos os quatro serotipos de dengue. Em particular, uma única mutação em NS1-53, uma mutação dupla em NS1-53 e em 5'NC-57, uma mutação dupla em NS1-53 e em NS3-250, e uma mutação tripla em NS1-53, em 5'NC-57 e em NS3-250, podem proporcionar um vírus do DEN-2 atenuado.

Para além disso, o genoma de qualquer vírus do dengue-2, que contém substituições não conservativas de aminoácidos nestes loci pode ser utilizado como estrutura nas quimeras não virulentas aqui descritas. Para além disso, outros genomas de flavivírus que contém mutações análogas nos mesmos loci, após alinhamento da sequência de aminoácidos ou de nucleótidos e da análise da estrutura de haste também podem ser utilizados como estrutura e são aqui definidos como sendo equivalentes a mutações atenuantes do genoma de PDK-53 do dengue-2.

A estrutura, a região da quimera que compreende as regiões não codificantes 5' e 3' e a região que codifica para as proteínas não estruturais, pode também conter mutações adicionais para manter a estabilidade do fenótipo não virulento e para reduzir a possibilidade de que o vírus não virulento ou quimera possa reverter para o vírus do tipo selvagem virulento. Por exemplo, uma segunda mutação na haste da estrutura haste e ansa na região não codificante 5' irá proporcionar estabilidade adicional, se desejado.

Estes vírus quiméricos podem compreender substituições, deleções ou inserções de nucleótidos e de aminoácidos nas suas proteínas estruturais e não estruturais, para além das que são aqui especificamente descritas.

É aqui descrito um método para a produção de vírus quiméricos utilizando técnicas recombinantes, inserindo as substituições necessárias no genoma da estrutura apropriada.

A presente invenção proporciona composições que compreendem um transportador farmacêuticamente aceitável e vírus quiméricos atenuados que contém sequências de aminoácidos derivadas a partir de outros serotipos do vírus do dengue, ou de outras espécies de flavivírus, tal como definido nas reivindicações. As sequências de

aminoácidos derivadas de outros serotipos do vírus do dengue, de outras espécies de flavivírus ou de outras espécies estreitamente relacionadas, tais como o vírus da hepatite C, podem ser expressas num hospedeiro, numa célula hospedeira ou numa cultura de células. As proteínas ou polipéptidos que compreendem as sequências de aminoácidos derivadas a partir de outros serotipos do vírus do dengue, de outras espécies de flavivírus ou de outras espécies estreitamente relacionadas, podem atuar como imunogénios e, assim, ser utilizadas para induzir uma resposta imunogénica contra outros serotipos do vírus do dengue, outras espécies de flavivírus ou outras espécies estreitamente relacionadas.

A presente invenção também proporciona composições que compreendem um transportador farmacologicamente aceitável, um ou mais vírus quiméricos atenuados da presente invenção e composições imunizantes adicionais, tal como definido nas reivindicações. Os exemplos de tais composições imunizantes adicionais incluem, vacinas de vírus do dengue, vacinas de vírus da febre amarela, vacinas de vírus da encefalite transmitido por carrapatos, vacinas de vírus da encefalite japonesa, vacinas de vírus do Nilo Ocidental, vacinas de vírus da hepatite C e qualquer combinação das mesmas. Tais vacinas podem ser vacinas de vírus vivo atenuado, vacinas de vírus morto, vacinas de subunidade, vacinas de vetor de ADN recombinante ou qualquer combinação das mesmas.

Uma vantagem clara da presente invenção é a de que fornece misturas de quimeras de flavivírus atenuadas numa composição para ser utilizada como vacinas, a fim de conferir imunidade simultaneamente contra várias espécies de flavivírus.

São aqui descritas quimeras de vírus que contêm substituições de aminoácidos ou de nucleótidos que retêm a imunogenicidade do vírus, evitando ao mesmo tempo

quaisquer efeitos patogénicos do vírus.

São fornecidos quimeras de ácido nucleico que compreendem uma sequência de nucleótidos de um vírus do dengue-2 atenuado e uma sequência de nucleótidos de um segundo flavivírus, em que a sequência de nucleótidos do segundo flavivírus dirige a síntese de antígenos de flavivírus.

Um objeto da presente invenção é proporcionar composições para vacinas que compreendem mais do que uma espécie de flavivírus, tal como definido nas reivindicações.

É descrito um método para a preparação de composições imunogénicas ou de vacinas utilizando técnicas recombinantes através da inserção de substituições necessárias num genoma de flavivírus apropriado.

São descritas as composições e os métodos para conferir imunidade contra mais do que um flavivírus em simultâneo.

São fornecidas as sondas de ácido nucleico e os oligonucleótidos iniciadores para utilização em qualquer um de uma série de testes genéticos rápidos que são usados como diagnóstico para cada um dos vírus de vacina aqui descritos. Isto pode ser concretizado em ensaios de reação em cadeia da polimerase, ensaios de hibridação ou outras técnicas de deteção de sequência de ácidos nucleicos conhecidas na tecnologia. Por exemplo, um sistema automatizado de deteção de ácido nucleico com base na PCR.

BREVE DESCRIÇÃO DAS FIGURAS

A figura 1 mostra esquematicamente a organização genómica dos vírus quiméricos DEN-2 / DEN-1. As designações das quimeras são baseadas nas estruturas do clone infeccioso específico do vírus do DEN-2 e na inserção dos genes estruturais (C-prM-E) do vírus do DEN-1. São utilizadas nas designações as letras sublinhadas da estrutura e dos vírus de inserção. D2/1 indica a quimera DEN-2 / DEN-1; a

primeira letra a seguir ao hífen é a estrutura viral de DEN-2, parental 16681 (P), PDK53-E (E), ou PDK53-V (V); a última letra indica os genes estruturais da estirpe parental DEN-1 16007 (P) ou do seu derivado de vacina, a estirpe PDK-13 (V). A estrutura de PDK53-E contém três mutações de aminoácidos específicas do vírus DEN-2 PDK-53 (NS1-53-Asp, NS2A-181-Phe, e NS4A-75-Ala), bem como a mutação 5'NC-57 do vírus PDK-53. A estrutura de PDK53-V contém estes mesmos loci específicos do vírus PDK-53 mais o locus adicional NS3-250-Val específico do vírus PDK-53. As figuras 2A e 2B mostram as características de crescimento do vírus DEN-2 / DEN-1 quimérico em células LLC-MK₂. As barras a pontilhado indicam o vírus DEN-1 16007 e os vírus quiméricos que expressam as proteínas estruturais do vírus DEN-1 16007. As barras com riscas indicam o vírus DEN-1 PDK-13 e os vírus quiméricos que expressam as proteínas estruturais do vírus PDK-13. As barras em branco indicam as três estruturas do vírus DEN-2 derivadas de clones infecciosos do vírus DEN-2 16681 (P48) e as duas variantes (PDK53-E e PDK53-V; E48 e V48, respetivamente). Figura 2A: Diâmetro médio da placa ($\pm$ SD). Os valores foram calculados a partir de dez placas individuais de cada vírus no dia 10 após a infeção. pp: placas de identificação de tamanho inferior a 1 mm. Figura 2B: Sensibilidade à temperatura (*ts*) e títulos de pico de vírus quiméricos no dia 8 ou 10 após a infeção. As pontuações *ts* foram baseadas na redução dos títulos de vírus a 38,7 °C versus aqueles a 37 °C (-, +, 2+ e 3+ indicam uma redução do título inferior ou igual a 60 %, 61 - 90 %, 91 - 99 %, > 99 %, respetivamente, calculadas a partir de pelo menos três experiências). As alturas das barras do gráfico representam os títulos $\log_{10}$ dos vírus a 37 °C. A multiplicidade de infeção (*moi*) foi de aproximadamente 0,001 ufp/célula.

A figura 3 mostra as curvas de crescimento dos vírus DEN-

1 16007, DEN-1 PDK-13, DEN-2 16681-P48, DEN-2 PDK53-E48, DEN-2 PDK53-V48 e do vírus DEN-2 / DEN-1 quimérico em células C6/36. As células foram infetadas a uma *moi* aproximada de 0,001 ufp/ml.

Figuras 4A - C. Figura 4A: Diâmetro médio da placa $\pm$ SD em milímetros (n = 12) dos vírus DEN-2 16681, PDK-53 e do vírus 16681 / PDK-53 recombinante aos nove dias após a infecção em células LLC-MK₂. Figura 4B: Títulos de replicação do pico a 6 - 8 dias após a infecção de células LLC-MK₂ a uma *moi* de aproximadamente 0,001 ufp/célula numa única experiência. As pontuações de sensibilidade à temperatura (ts) para os vírus cultivados a 37 °C ou a 38,7 °C em células LLC-MK₂ são mostradas acima das barras do gráfico para os títulos de replicação do pico. As pontuações de (-), (+/-) e (+) indicam menos de 81 %, 81 - 89 % e 90 - 97 % de redução no título viral, respetivamente, a 38,7 °C. As pontuações foram determinadas aos oito dias após a infecção. Figura 4C: Média dos títulos de replicação do pico aos 12 dias após a infecção de células C6/36 a uma multiplicidade de aproximadamente 0,001 ufp/célula em duas experiências independentes. Os títulos de pico individuais das duas experiências são indicados por linhas verticais em cada barra do gráfico. As designações numéricas para os vírus Px e Vx recombinantes (em que x = loci 5'NC, NS 1, e/ou NS3) indicam os loci específicos do vírus 16681 parental (P em designação de vírus) manipulados no clone de ADNc infeccioso específico do vírus PDK-53 ou os loci recíprocos específicos do vírus PDK-53 candidato a vacina (V em designação de vírus) manipulados no clone 16681, respetivamente. Os vírus cognatos são indicados em todos os três gráficos por barras de gráfico de padrão idêntico sólido ou de quadrícula transversal. O cognato para o vírus P5 é o vírus V13, assumindo que o fenótipo viral é determinada predominantemente pelos loci 5'NC-57, NS1-53

e NS3-250. Ambos os vírus P5 e V13 contêm os loci 5'NC-57-C (16681), NS1-53-Asp (PDK-53) e NS3-250-Val (PDK-53) nos fundos genéticos dos vírus PDK-53 e 16681, respetivamente.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A descrição que se segue inclui o melhor modo presentemente contemplado de realizar a invenção. Esta descrição é feita com a finalidade de ilustrar os princípios gerais das invenções e não deve ser tomada num sentido limitativo.

São aqui descritas quimeras de flavivírus imunogénicas, um vírus do dengue-2 ou uma estrutura equivalente ao vírus do dengue-2 para a preparação das quimeras de flavivírus e os métodos para a preparação das quimeras de flavivírus. As quimeras de flavivírus imunogénicas são úteis, por si só ou em combinação, num transportador farmacêuticamente aceitável na forma de composições imunogénicas para minimizar, inibir, ou imunizar indivíduos contra a infeção por um ou mais flavivírus ou estirpes de flavivírus, em especial as estirpes do vírus do dengue dos serotipos DEN-1, DEN-2, DEN-3 e DEN-4. Quando combinadas, as quimeras de flavivírus imunogénicas podem ser utilizadas como uma vacina multivalente para conferir proteção simultânea contra a infeção. De um modo preferido, as quimeras do vírus do dengue são combinadas numa composição imunogénica útil como uma vacina tetravalente contra os quatro serotipos de vírus do dengue conhecidos.

As quimeras de flavivírus imunogénicas aqui descritas também são úteis, em combinação com as estirpes de vírus não virulentas, num transportador farmacêuticamente aceitável, na forma de composições imunogénicas para minimizar, inibir ou imunizar indivíduos contra a infeção por múltiplas espécies patogénicas. Por exemplo, uma ou mais das quimeras de flavivírus imunogénicas podem ser

combinadas com serotipos do vírus não virulentos de flavivírus selecionados para proporcionar uma vacina tetravalente segura e eficaz contra os quatro serotipos de vírus do dengue conhecidos. Numa outra forma de realização, as quimeras de flavivírus aqui descritas podem ser combinadas com as vacinas de vírus não virulentos para fornecer uma vacina segura e eficaz contra a infecção por múltiplas espécies patogénicas.

A presente descrição também se refere a composições que compreendem um transportador farmacêuticamente aceitável, um ou mais vírus quiméricos atenuados aqui descritos e composições imunizantes adicionais. Os exemplos de tais composições imunizantes adicionais incluem, vacinas de vírus do dengue, vacinas da febre amarela, vacinas da encefalite transmitida por carrapatos, vacinas da encefalite japonesa, vacinas de vírus do Nilo Ocidental, vacinas de vírus da hepatite C e qualquer combinação das mesmas. Tais vacinas podem ser vacinas de vírus vivo atenuado, vacinas de vírus morto, vacinas de subunidade, vacinas de vetor de ADN recombinante ou qualquer combinação das mesmas.

É também descrita a sequência de ácido nucleico para cada um dos vírus DEN-1, DEN-3 e DEN-4 para utilização como sondas para detetar o vírus do dengue numa amostra biológica.

As quimeras aqui descritas podem compreender a estrutura do vírus do dengue-2 de um vírus do dengue-2 atenuado e outras sequências de nucleótidos selecionadas a partir de mais de um serotipo de vírus do dengue, outras espécies de flavivírus, outras espécies estreitamente relacionadas, tais como o vírus da hepatite C, ou qualquer combinação das mesmas. Estas quimeras podem ser usadas para induzir uma resposta imunogénica contra mais do que uma espécie selecionada dos serotipos do vírus do dengue, espécies de flavivírus, outras espécies estreitamente

relacionadas ou qualquer combinação das mesmas.

Uma quimera de ácido nucleico aqui descrita pode compreender uma primeira sequência de nucleótidos que codifica para proteínas não estruturais de um vírus do dengue-2 atenuado, e uma segunda sequência de nucleótidos que codifica para uma proteína estrutural de um segundo flavivírus. O vírus do dengue-2 atenuado pode ser a estirpe de vacina PDK-53. A proteína estrutural pode ser a proteína C de um flavivírus, a proteína prM de um flavivírus, a proteína E de um flavivírus, ou qualquer combinação das mesmas. Os exemplos de flavivírus a partir dos quais a proteína estrutural pode ser selecionada incluem, mas não estão limitados a, vírus do dengue-1, vírus do dengue-2, vírus do dengue-3, vírus do dengue-4, vírus do Nilo Ocidental, vírus da encefalite japonesa, vírus da encefalite de St. Louis, vírus da febre amarela e vírus da encefalite transmitida por carrapatos. Num outro aspeto, a proteína estrutural pode ser selecionada a partir de espécies que não flavivírus que estão intimamente relacionadas com os flavivírus, tal como o vírus da hepatite C.

Os termos "um", "uma" e "o", "a", tal como são aqui utilizados, são definidos para significar um ou mais, e incluem o plural a menos que o contexto seja inadequado.

O termo "resíduo" é aqui utilizado para se referir a um aminoácido (D ou L) ou um mimético de aminoácido que é incorporado num péptido por uma ligação amida. Como tal, o aminoácido pode ser um aminoácido que ocorre naturalmente ou, a menos que de outro modo limitado, pode englobar análogos conhecidos de aminoácidos naturais que funcionam de um modo semelhante aos aminoácidos de ocorrência natural (ou seja, miméticos de aminoácidos). Para além disso, uma ligação amida mimética inclui modificações da estrutura do péptido bem conhecidas dos peritos na tecnologia.

Para além disso, um perito na tecnologia irá reconhecer que substituições, supressões ou adições individuais na sequência de aminoácidos, ou na sequência de nucleótidos que codifica para os aminoácidos, que alteram, adicionam ou eliminam um único aminoácido ou uma pequena percentagem de aminoácidos (tipicamente menos de 5 %, mais tipicamente menos de 1 %) numa sequência codificada são variações modificadas de forma conservativa, em que as alterações resultam na substituição de um aminoácido por um aminoácido quimicamente similar. As tabelas de substituição conservativa que fornecem aminoácidos funcionalmente semelhantes são bem conhecidas na tecnologia. Os seis grupos que se seguem contêm cada um aminoácidos que são substituições conservativa entre si:

- 1) alanina (A), serina (S), treonina (T);
- 2) ácido aspártico (D), ácido glutâmico (E);
- 3) asparagina (N), glutamina (Q);
- 4) arginina (R), lisina (K);
- 5) isoleucina (I), leucina (L), metionina (M), valina (V); e
- 6) fenilalanina (F), tirosina (Y), triptofano (W).

Tal como são aqui utilizados, os termos "quimera de vírus", "vírus quimérico", "quimera de flavivírus" e "flavivírus quiméricos" significam uma construção infecciosa aqui descrita que compreende uma porção da sequência de nucleótidos de um vírus do dengue-2 e sequência de nucleótidos adicional que não é do mesmo vírus do dengue-2. A sequência de nucleótidos adicional pode ser, sequências do vírus do dengue-1, do vírus do dengue-3, do vírus do dengue-4, ou do vírus do Nilo Ocidental.

Tal como é aqui utilizado, "construção infecciosa" indica um vírus, uma construção viral, uma quimera viral, um ácido nucleico derivado a partir de um vírus ou

qualquer porção do mesmo, que podem ser utilizados para infectar uma célula.

Tal como é aqui utilizado, "quimera de ácido nucleico" significa uma construção que compreende ácido nucleico que compreende uma porção da sequência de nucleótidos de um vírus do dengue-2 e sequência de nucleótidos adicional que não é da mesma origem que a sequência de nucleótidos do vírus do dengue-2. Correspondentemente, qualquer flavivírus quimérico ou quimera de flavivírus da descrição é para ser reconhecido como um exemplo de uma quimera de ácido nucleico.

As sequências de nucleótidos do genoma de ARN do vírus e das quimeras aqui descritos são enumeradas nas listagens de sequências em termos de ADN.

A menos que definido de outro modo, todos os termos técnicos e científicos aqui utilizados têm o mesmo significado que o normalmente entendido por um vulgar perito na tecnologia à qual esta invenção pertence. Embora possam ser utilizados na prática ou no teste da presente invenção outros materiais e métodos semelhantes ou equivalentes aos aqui descritos, os métodos e os materiais preferidos são agora descritos.

Quimeras de flavivírus

Os vírus do dengue dos tipos 1 - 4 (DEN-1 a DEN-4) são flavivírus patogénicos transmitidos por mosquitos. O genoma do flavivírus contém uma região não codificante 5' (5'-NC), seguida por uma região codificante da proteína da cápside (C), seguida por uma região codificante da proteína de pré-membrana / membrana (prM), seguida por uma região codificante da proteína do invólucro (E), seguida pela região que codifica para as proteínas não estruturais (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) e, finalmente, uma região não codificante de 3' (3'NC). As proteínas estruturais virais são a C, a prM e a E, e as proteínas não estruturais são a NS1-NS5. As proteínas estruturais e

não estruturais são traduzidas como uma única poliproteína e processadas por proteases celulares e virais.

As quimeras de flavivírus aqui descritas são construções formadas por fusão de genes de proteínas não estruturais de um tipo, ou serotipo, do vírus do dengue ou de espécies de vírus da família Flaviviridae, com genes de proteínas, por exemplo, genes de proteínas estruturais, de um tipo, ou serotipo diferente, do vírus do dengue ou de espécies de vírus da família Flaviviridae. Alternativamente, uma quimera de flavivírus pode ser uma construção formada por fusão de genes de proteínas não estruturais de um tipo, ou serotipo, do vírus do dengue ou de espécies de vírus da família Flaviviridae, com sequências de nucleótidos adicionais que dirigem a síntese de polipéptidos ou de proteínas selecionados de outros serotipos de vírus do dengue ou de outros vírus da família Flaviviridae.

As quimeras de flavivírus imunogênicas, não virulentas aqui proporcionadas contêm os genes de proteínas não estruturais do vírus do dengue-2 atenuado, ou um equivalente do mesmo, e um ou mais dos genes de proteínas estruturais ou porções antigênicas dos mesmos, do flavivírus contra o qual se pretende conferir imunogenicidade. Os flavivírus adequados incluem, mas não estão limitados aos listados na tabela 1.

Outros flavivírus adequados para a utilização na construção de quimeras de flavivírus são DEN-1 16007, DEN-2 16681 (SEQ ID NO: 13; SEQ ID NO: 14), DEN-3 16562 e DEN-4 1036 do tipo selvagem, virulento e DEN-1 PDK-13, DEN-2 PDK-53 (SEQ ID NO: 15; SEQ ID NO: 16), DEN-3 PMK-30 / FRhL-3 e DEN-4 PDK-48 estirpe de vacina, atenuada. São aqui descritos flavivírus adequados adicionais, ou variantes dos flavivírus mencionados anteriormente. As diferenças genéticas entre os pares DEN-1, DEN-2, DEN-3 e DEN-4 do tipo selvagem / vírus atenuado são mostradas nas

tabelas 2 - 5, juntamente com as alterações nas sequências de aminoácidos codificadas pelos genomas virais.

As listagens de sequências para DEN-2 PDK-53 aqui fornecidas (SEQ ID NO: 15; SEQ ID NO: 16) correspondem à variante DEN-2 PDK-53-V, em que a posição do nucleótido 5270 no genoma está mutada de uma A para uma T e, a posição do aminoácido 1725 da poliproteína ou a posição do aminoácido 250 da proteína NS3 contém um resíduo de valina. A variante DEN-2 PDK-53 sem esta mutação de nucleótido, DEN-2 PDK-53-E, difere de PDK-53-V apenas nesta única posição. DEN-2 PDK-53-E tem um A na posição do nucleótido 5270 e um glutamato na posição do aminoácido 1725 da poliproteína, posição do aminoácido 250 da proteína NS3 (tabela 3).

O DEN-3 16562 aqui descrito corresponde à variante em que a posição do nucleótido 1521 no genoma é um T e a posição do aminoácido 476 da poliproteína ou a posição do aminoácido 196 da proteína E contém uma leucina. Uma segunda variante, presente em culturas de DEN-3 16562 tem um T na posição do nucleótido 1521 e a posição do aminoácido 476 da poliproteína ou a posição do aminoácido 196 da proteína E contém uma serina (tabela 4).

O DEN-4 PDK-48 aqui descrito corresponde à variante em que as posições do nucleótido 6957 no genoma é um T e a posição do aminoácido 2286 da poliproteína e a posição do aminoácido 44 da proteína NS4B é uma fenilalanina, 7546 é um T e a posição do aminoácido 2366 da poliproteína e a posição do aminoácido 240 da proteína NS4B é uma valina, e 7623 é um T e a posição do aminoácido 2508 da poliproteína e a posição do aminoácido 21 da proteína NS5 é uma tirosina (tabela 5).

Ao longo do texto, as designações das quimeras são baseadas nas estruturas do clone infeccioso específico do vírus do DEN-2 e na inserção de genes estruturais (prM-E ou C-prM-E) de outros flavivírus. Cada designação começa

com DEN-2 para a estrutura do dengue-2, seguida da estirpe a partir da qual os genes estruturais são inseridos. A variante de estrutura particular é refletida na seguinte letra. A variante da estrutura de DEN-2 particular a partir da qual a quimera foi construída é indicada pela letra seguinte colocada depois de um hífen, parental 16681 (P), PDK53-E (E), ou PDK53-V (V); a última letra indica os genes estruturais de C-prM-E da estirpe parental (P) ou do seu derivado de vacina (V) ou os genes estruturais de prM-E da estirpe parental (P1) ou do seu derivado de vacina (V1). Por exemplo, DEN-2/1-VP (SEQ ID NO: 5; SEQ ID NO: 6) indica a quimera que compreende a estrutura de DEN-2 PDK-53V atenuado que compreende uma valina em NS3-250 e os genes de C-prM-E do DEN-1 16007 do tipo selvagem; DEN-2/1-VV (SEQ ID NO: 7; SEQ ID NO: 8) indica a estrutura de DEN-2 PDK-53V com a estirpe de vacina do dengue-1, DEN-1 PDK-13; DEN-2/1-VP1 (SEQ ID NO: 27; SEQ ID NO: 28) indica a estrutura de DEN-2 PDK-53V e os genes de prM-E do DEN-1 16007 do tipo selvagem; DEN-2/3-VP1 (SEQ ID NO: 9; SEQ ID NO: 10) indica a estrutura de DEN-2 PDK-53V e os genes de prM-E do DEN-3 16562 do tipo selvagem; DEN-2/4-VP1 (SEQ ID NO: 11; SEQ ID NO: 12) indica a estrutura de DEN-2 PDK-53V e os genes de prM-E do DEN-4 1036 do tipo selvagem; e DEN-2 / WN-PP1 (SEQ ID NO: 17; SEQ ID NO: 18) indica a estrutura de DEN-2 16681 e os genes de prM-E do Nilo Ocidental NY99. Outros quimeras da presente divulgação indicadas da mesma maneira, são aqui claramente definidas pelas sequências divulgadas. Por exemplo, DEN-2/1-PV é aqui definido como consistindo da estrutura do dengue-2 do tipo selvagem, DEN-2 16681, e os genes de C-prM-E da estirpe de vacina do dengue-1, DEN-1 PDK-13.

As quimera aqui descritas podem conter o genoma do vírus do dengue-2 PDK-53 atenuado como estrutura viral, em que os genes de proteínas estruturais que codificam para as proteínas C, prM e E do genoma de PDK-53, ou

combinações dos mesmos, são substituídos com os genes da proteína estrutural correspondente de um flavivírus contra o qual se pretende a proteção, tais como um flavivírus diferente ou uma estirpe diferente do vírus do dengue. Também podem ser incorporados na estrutura de DEN-2 flavivírus recentemente descobertos ou agentes patogênicos de flavivírus. A recombinação genética com os vírus relacionados, tais como o vírus da hepatite C (HCV) também pode ser utilizada para produzir as quimeras da presente divulgação. A quimera viral resultante tem as propriedades funcionais do vírus do dengue-2 atenuado e é, por conseguinte, não virulenta, mas expressa epítomos antigênicos dos produtos de genes estruturais e é, portanto, imunogênica.

São aqui identificadas nove mutações de nucleótidos entre os genomas do vírus DEN-2 16681 do tipo selvagem e duas estirpes de vírus PDK-53 atenuadas (tabela 3). Três destas mutações são mutações silenciosas na medida em que não resultam na produção de um aminoácido que seja diferente do aminoácido na mesma posição no vírus do tipo selvagem. A primeira mutação é uma mutação nucleotídica de C para T (do tipo selvagem para PDK-53) na posição do nucleótido 57 (nt 57) no genoma na região não codificante 5'. A segunda mutação é uma mutação de A para T na posição do nucleótido 524 no genoma, que codifica para a substituição de aminoácidos Asp para Val na região da proteína estrutural de pré-membrana, prM-29.

Nas regiões da proteína não estrutural, foi descoberta uma mutação Gly para Asp (do tipo selvagem para PDK-53) na proteína não estrutural NS1-53 (posição do nucleótido 2579 no genoma); foi descoberta uma mutação Leu para Phe (tipo selvagem para PDK-53) na proteína não estrutural NS2A-181 (posição do nucleótido 4018 no genoma); foi descoberta uma mutação Glu para Val (do tipo selvagem para PDK-53) na proteína não estrutural NS3-250

(posição do nucleótido 5270 no genoma); e foi descoberta uma mutação Gly para Ala (do tipo selvagem para PDK-53) na proteína não estrutural NS4A-75 (posição do nucleótido 6599 no genoma).

A estirpe do vírus PDK-53 atenuada tem um genótipo misto no nt 5270 no genoma. Uma porção significativa (aproximadamente 29 %) da população de vírus codifica para NS3-250-Glu não mutado que está presente no vírus DEN-2 16681 do tipo selvagem em vez da mutação NS3-250-Val. Como ambas as variantes genéticas são não virulentas, esta mutação pode não ser necessária numa quimera não virulenta.

Foi descoberto inesperadamente que a não virulência da estirpe do vírus PDK-53 atenuada pode ser atribuída à presença de mutações específicas na sequência de nucleótidos que codifica para proteínas não estruturais e na região não codificante 5' (exemplo 5). Em particular, uma única mutação em NS1-53, uma mutação dupla em NS 1-53 e em 5'NC-57, uma mutação dupla em NS1-53 e em NS3-250 e uma mutação tripla em NS 1-53, em 5'NC-57 e em NS3-250, resultam na atenuação do vírus do DEN-2. Portanto, o genoma de qualquer vírus do dengue-2 que contenha tais substituições de aminoácidos não conservativas ou substituições de nucleótidos nestes loci pode ser utilizado como estrutura nas quimeras não virulentas aqui descritas. A estrutura também pode conter mutações adicionais para manter a estabilidade do fenótipo não virulento e para reduzir a possibilidade de que o vírus não virulento ou quimera possa reverter para o vírus do tipo selvagem virulento. Por exemplo, uma segunda mutação na haste da estrutura haste - ansa na região não codificante 5' irá proporcionar estabilidade fenotípica não virulenta adicional, se desejado. As mutações nesta região perturbam as estruturas secundárias potenciais importantes para a replicação viral. Em particular, as

mutações na região não codificante 5' têm a capacidade de perturbar o funcionamento da cadeia de ARN de sentido positivo e a função da cadeia de sentido negativo durante a replicação. Uma única mutação nesta curta (apenas 6 resíduos de nucleótidos de comprimento) estrutura de haste em ambos os vírus de DEN e da encefalite equina venezuelana perturba a formação da estrutura em gancho (Kinney et al., Virology 67, 1269-1277, (1993)). Mutações adicionais nesta estrutura de haste diminuem a possibilidade de reversão neste locus, mantendo ao mesmo tempo a viabilidade do vírus. Para além disso, os genomas de flavivírus que contêm uma estrutura de haste análoga que consiste em sequências de nucleótidos curtas (hastes que consistem em 6 ou mais pares de bases na estrutura de haste e ansa) localizadas na região não codificante 5' e que têm uma ou mais mutações na estrutura de haste podem ser igualmente úteis como estrutura da presente divulgação.

Tais mutações podem ser alcançadas através de mutagénesse dirigida ao local, utilizando técnicas conhecidas dos peritos na tecnologia. Para além disso, podem ser utilizados como estrutura da quimera aqui descrita outros genomas de flavivírus que contenham mutações análogas nos mesmos loci após alinhamento da sequência de aminoácidos, e são aqui definidos como sendo equivalentes ao genoma de dengue-2 PDK-53. Será entendido por aqueles peritos na tecnologia que os ensaios de rastreio de virulência, tal como são aqui descritos e tal como é bem conhecido na tecnologia, podem ser usados para distinguir entre estruturas virulentas e não virulentas.

Construção de quimeras de flavivírus

As quimeras de flavivírus aqui descritas podem ser produzidas por "splicing" de um ou mais dos genes de proteínas estruturais do flavivírus contra o qual se pretende conferir imunidade numa estrutura do genoma do

vírus do dengue PDK-53, ou o equivalente da mesma, tal como é descrito anteriormente, utilizando técnicas de engenharia recombinante bem conhecidas dos peritos na tecnologia, para remover o gene PDK-53 correspondente e substituí-lo com o gene desejado. Alternativamente, usando as sequências proporcionadas na listagem de sequências, as moléculas de ácido nucleico que codificam para as proteínas de flavivírus podem ser sintetizadas utilizando técnicas de síntese de ácido nucleico conhecidas, e inseridas num vetor apropriado. O vírus imunogénico, não virulento é, por conseguinte, produzido utilizando técnicas de engenharia recombinante conhecidas dos peritos na tecnologia.

Tal como mencionado anteriormente, o gene a ser inserido na estrutura codifica para uma proteína estrutural do flavivírus. De preferência, o gene de flavivírus a ser inserido é um gene que codifica para uma proteína C, uma proteína PrM e/ou uma proteína E. A sequência inserida na estrutura de dengue-2 pode codificar tanto a proteína estrutural prM como a E. A sequência inserida na estrutura de dengue-2 pode codificar para as proteínas estruturais C, prM e E. A estrutura do vírus do dengue é o genoma do vírus dengue-2 PDK-53 e inclui quer os genes excisados que codificam para as proteínas estruturais PrM e E de dengue-1 (DEN-2/1), quer os genes excisados que codificam para as proteínas estruturais PrM e E de dengue-3 (DEN-2/3), quer os genes excisados que codificam para as proteínas estruturais PrM e E de dengue-4 (DEN-2/4). O gene excisado que codifica para a proteína estrutural do vírus do dengue-3 pode dirigir a síntese de uma proteína E que contém uma leucina na posição do aminoácido 345.

Numa forma de realização particular, a quimera aqui divulgada codifica para a proteína estrutural C do vírus do dengue-2 e dirige a síntese de uma proteína C que

contém uma serina na posição do aminoácido 100, e compreende um gene excisado que codifica para as proteínas estruturais de dengue-4, que dirige a síntese de uma proteína E que contém uma leucina na posição do aminoácido 447.

Numa outra forma de realização, a quimera aqui divulgada codifica para a proteína estrutural C do vírus do dengue-2 e dirige a síntese de uma proteína C que contém uma serina na posição do aminoácido 100, e compreende um gene excisado que codifica para as proteínas estruturais de dengue-4, que dirige a síntese de uma proteína E que contém uma leucina na posição do aminoácido 447 e uma valina na posição do aminoácido 364. As proteínas estruturais aqui descritas podem estar presentes como a única proteína estrutural do flavivírus ou em qualquer combinação de proteínas estruturais do flavivírus numa quimera viral da presente invenção.

As quimeras aqui descritas podem ser concebidas por recombinação de clones de ADNc do genoma de tamanho completo derivado do vírus DEN-2 16681 do tipo selvagem de e de qualquer uma das variantes do vírus PDK-53 dengue-2 (-E ou -V (SEQ ID NO: 15)). A vacina de PDK-53 não clonado contém uma mistura de duas variantes genotípicas, aqui designadas como PDK53-E e PDK53-V. A variante PDK53-V contém todas as nove mutações de nucleótidos específicas da vacina de PDK-53, incluindo a mutação Glu para Val na posição do aminoácido NS3-250. A variante PDK53-E contém oito das nove mutações da vacina de PDK-53 e NS3-250-Glu do vírus parental 16681. Os clones de ADNc infecciosos são construídos para ambas as variantes, e os vírus derivados a partir de ambos os clones são atenuados em ratinhos. Os marcadores fenotípicos de atenuação do vírus DEN-2 PDK-53 incluem o tamanho pequeno placa, a sensibilidade à temperatura (em particular em células LLC-MK₂), a replicação limitada (em particular em células C6/36), a

atenuação de ratinhos recém-nascidos (especificamente perda de neurovirulência para ratinhos lactentes) e a diminuição da incidência de viremia em macacos. As quimeras que são úteis como candidatos a vacina são construídas nos fundos genéticos das duas variantes de DEN-2 PDK-53, que contêm todas mutações nas regiões não estruturais do genoma, incluindo 5'NC-57 C para T (16681 para PDK-53) na região não codificante 5', bem como mutações na sequência de aminoácidos das proteínas não estruturais, tais como, por exemplo, NS1-53 Gli para Asp e NS3-250 Glu para Val.

Os vírus quiméricos adequados ou quimeras de ácidos nucleicos que contêm sequências de nucleótidos que codificam para as proteínas estruturais de outros flavivírus ou serotipos de vírus do dengue podem ser avaliados para a sua utilidade como vacinas, por rastreio para os marcadores fenotípicos precedentes de atenuação que indicam a não virulência e por rastreio para a imunogenicidade. A antigenicidade e a imunogenicidade podem ser avaliadas utilizando a reatividade *in vitro* ou *in vivo* com anticorpos de flavivírus ou soro imunorreativo utilizando os procedimentos de rastreio de rotina conhecidos dos peritos na tecnologia.

Vacinas de flavivírus

Os vírus quiméricos preferidos e as quimeras de ácidos nucleicos fornecem vírus vivos, atenuados úteis como imunogénios ou vacinas. As quimeras podem exibir elevada imunogenicidade, enquanto ao mesmo tempo não produzem efeitos patogénicos perigosos ou letais.

Até à data não está disponível uma vacina eficaz contra todas as estirpes de vírus do dengue. Foram desenvolvidos candidatos a vacinas vivas, atenuadas, individuais para os quatro serotipos através da passagem em série de vírus do tipo selvagem em células primárias de rim de cão (PDK) ou em outros tipos de células. Tal como

descrito anteriormente, o vírus PDK-53 é um candidato a vacina do dengue útil. No entanto, uma vacina derivada do PDK-53 só iria fornecer imunidade contra o serotipo DEN-2.

Para evitar a possível ocorrência de FHD / SCD em pacientes vacinados contra apenas um serotipo do vírus do dengue, é necessária uma vacina tetravalente para proporcionar imunidade simultânea para todos os quatro serotipos do vírus. Uma vacina tetravalente pode ser produzida através da combinação de dengue-2 PDK-53 com as quimeras de dengue-2/1, dengue-2/3 e dengue-2/4 descritas anteriormente, num transportador farmacêutico adequado para a administração como uma vacina multivalente.

Os vírus quiméricos ou as quimeras de ácido nucleico aqui descritos podem compreender os genes estruturais quer de um vírus do tipo selvagem quer de um vírus atenuado numa estrutura do vírus de DEN-2 virulento ou atenuado. Por exemplo, a quimera pode expressar os genes das proteínas estruturais do vírus DEN-1 16007 do tipo selvagem ou do seu derivado candidato a vacina PDK-13 em qualquer um dos fundos DEN-2 PDK-53.

Tal como descrito no exemplo 1, todos os vírus DEN-2/1 quiméricos que contêm as proteínas C, prM e E quer do vírus DEN-1 16007 (quimeras DEN-2/1-EP e -VP) quer do vírus PDK-13 (quimeras DEN-2/1-EV e -VV) nas estruturas de DEN-2 PDK-53, retêm todos os marcadores fenotípicos de atenuação do vírus DEN-2 PDK-53. Os vírus DEN-2/1-EP e -VP quiméricos, que contêm as proteínas C, prM e E do vírus DEN-1 16007 são geneticamente mais estáveis após a passagem em cultura de células do que os vírus DEN-2/1-EV e -VV. A imunogenicidade dos vírus quiméricos que expressam as proteínas estruturais do vírus DEN-1 16007 foi mais elevada em comparação com os títulos de anticorpos neutralizantes induzidos pelo vírus de vacina PDK-13 e de quimeras que expressam as proteínas estruturais do vírus PDK-13. Assim, os vírus DEN-2/1-EP e

-VP quiméricos, que expressam os genes estruturais do vírus DEN-1 16007 do tipo selvagem no fundo genético das duas variantes de DEN-2 PDK-53, são candidatos potenciais a vacina de DEN-1, que são superiores ao candidato a vacina de PDK-13. Estas duas quimeras replicam-se bem em células LLC-MK₂ e retêm os marcadores de atenuação associados com o vírus DEN-2 PDK-53, incluindo o tamanho pequeno de placa, a sensibilidade à temperatura, a replicação restrita em células de mosquito e a atenuação para ratinhos. São pelo menos tão imunogênicas como o vírus DEN-1 16007 do tipo selvagem em ratinhos.

Outros exemplos, tais como as quimeras DEN-2/3 e DEN-2/4, nos exemplos 2 - 4, também mostraram que os vírus quiméricos que contêm os genes das proteínas estruturais do vírus DEN-3 ou DEN-4 do tipo selvagem nas estruturas de DEN-2 PDK-53, são candidatos adequados a vacinas que retêm todos os marcadores fenotípicos atenuados do vírus DEN-2 PDK-53 (tabela 14), ao mesmo tempo que fornecem imunogenicidade contra o vírus DEN-3 ou DEN-4. A estratégia aqui descrita da utilização de um fundo genético que contém os determinantes de atenuação em regiões não estruturais do genoma para expressar os genes das proteínas estruturais dos vírus heterólogos levou ao desenvolvimento de candidatos a vacinas de flavivírus vivos, atenuados que expressam os genes da proteína estrutural do tipo selvagem de imunogenicidade ótima. Assim, podem ser concebidos os candidatos a vacinas para variantes imunogênicas de vários patógenos de flavivírus.

Os vírus utilizados nas quimeras aqui descritas são normalmente cultivados utilizando técnicas conhecidas na tecnologia. As titulações de placas de vírus são então realizadas, e as placas são contadas a fim de avaliar as características de viabilidade e fenotípicas das culturas em crescimento. Os vírus do tipo selvagem são passados através de linhas de células cultivadas para se obterem os

materiais de partida dos candidatos atenuados.

Os clones infecciosos quiméricos são construídos a partir de vários clones de serotipo do dengue disponíveis. Se desejado, também pode ser realizada a clonagem de fragmentos de ADNc específicos do vírus. Os fragmentos de ADNc que contêm os genes da proteína estrutural ou da proteína não estrutural são amplificados por reação em cadeia da polimerase por transcriptase reversa (RT-PCR) a partir do ARN do vírus do dengue com vários oligonucleótidos iniciadores. Os fragmentos amplificados são clonados nos locais de clivagem de outros clones intermédios. Os clones intermédios do vírus do dengue quimérico são então sequenciados para verificar a precisão da inserção do ADNc específico do vírus do dengue.

Os plasmídeos quiméricos do genoma de comprimento completo construídos através da inserção da região do gene da proteína estrutural ou da proteína não estrutural de vírus de serotipo do dengue em vetores, são obtidos utilizando técnicas recombinantes bem conhecidas dos peritos na tecnologia.

Análise de nucleótido e de aminoácidos

As sequências de nucleótidos para o DEN-2 16681 e a PDK-53-V correspondente são fornecidas na SEQ ID NO: 13 e SEQ ID NO: 15, respetivamente. As sequências de aminoácidos para o DEN-2 16681 e a PDK-53-V correspondente são fornecidas na SEQ ID NO: 14 e SEQ ID NO: 16. É ainda descrita na tabela 3 a variante -E de PDK-53, que varia da PDK-53-V na posição do nucleótido 5270 e na posição do aminoácido 1725 da poliproteína. É mostrada na tabela 3 uma comparação das substituições de nucleótidos e de aminoácidos críticos que foram descobertas entre a estirpe parental e o vírus atenuado. A sequência dos amplicões de ADNc do DEN-2 foi amplificada a partir do ARN genómico viral do DEN-2 por reação em cadeia da polimerase por transcriptase reversa (RT-PCR).

Ao contrário do PDK-53, que não contém mutações de aminoácidos na proteína E em comparação com o vírus do dengue-2 do tipo selvagem, os vírus DEN-1, DEN-3 e DEN-4 atenuados da Mahidol têm todas mutações de aminoácidos na proteína E (tabelas 2, 4 e 5). O DEN-3 16562 do tipo selvagem aqui descrito, foi demonstrado compreender vestígios de uma variante que compreende um T na posição do nucleótido 1521, que dirige a incorporação de uma leucina na posição 476 da poliproteína, posição do resíduo de aminoácido 476 da proteína E.

Cada um dos três vírus anteriores possui uma mutação Glu para Lys (parental para vacina) na proteína E, embora a mutação esteja localizada num resíduo de aminoácido diferente na proteína E. Esta substituição provoca uma mudança de um aminoácido carregado negativamente para um carregado positivamente. A substituição de Glu para Lys na proteína E do vírus de vacina DEN-4 foi a única mutação presente na proteína E, enquanto que as proteínas E dos vírus de vacina DEN-1 e DEN-3 tinham cinco e três mutações de aminoácidos, respetivamente.

A mutação NS1-53 no vírus de vacina DEN-2 PDK-53 é significativa para o fenótipo atenuado deste vírus, porque a NS1-53-Gly do vírus DEN-2 16681 é conservada em quase todos os flavivírus, incluindo os vírus transmitidos por carrapatos, sequenciados até ao momento. O vírus de vacina DEN-4 da Mahidol também contém uma mutação de aminoácidos na proteína NS1 na posição 253. Este locus, o qual é uma mutação Gln para His no vírus de vacina DEN-4 PDK-48 é Gln em todos os quatro serotipos do tipo selvagem do vírus do dengue. Este resíduo Gln é único para os vírus do dengue dentro do género flavivírus. A proteína NS1 é uma glicoproteína que é segregada a partir de células infectadas com flavivírus. Está presente na superfície da célula infectada e os anticorpos específicos de NS1 estão presentes no soro de indivíduos infectados com o vírus. Tem

sido relatada proteção de animais imunizados com a proteína NS1 ou passivamente com o anticorpo específico de NS1. A proteína NS1 parece participar na replicação precoce do ARN viral.

As mutações que ocorreram nas proteínas NS2A, NS2B, NS4A, e NS4B das estirpes DEN-1, -2, -3 e -4 atenuadas foram todas conservativas em natureza. As mutações NS4A-75 e NS4A-95 dos vírus de vacina DEN 2 e DEN-4, respetivamente, ocorreram nos locais de conservação de aminoácidos entre os vírus do dengue, mas não entre os flavivírus em geral.

A proteína NS3 flaviviral possui pelo menos duas funções reconhecidas: a proteinase viral e a ARN helicase / NTPase. A proteína NS3 com 698-aa de comprimento (vírus do DEN-2) contém um domínio de protease de serina do terminal amino (triáde catalítica NS3-51-His, -75-Asp, -135-Ser) que é seguido por motivos de sequência de funções de ARN helicase / NTPase. Nenhuma das mutações nas proteínas NS3 dos vírus DEN-1, DEN-2, ou DEN-3 ocorreu dentro de um motivo reconhecido. A mutação NS3-510 Tyr para Phe no vírus DEN-1 PDK-13 foi conservativa. Uma vez que os vírus DEN-2, -3 e -4 do tipo selvagem contêm Phe nesta posição, não é provável que a mutação de Tyr para Phe desempenhe um papel na atenuação do vírus DEN-1. A mutação NS3-182 Glu para Lys no vírus DEN-1 PDK-13 ocorreu numa posição que é conservada como Asp ou Glu na maioria dos flavivírus transmitidos por mosquitos, e pode desempenhar algum papel na atenuação. Esta mutação estava localizada 15 resíduos de aminoácidos a montante do motivo de helicase GAGKT. Tal como observado em relatórios anteriores, o NS3-250-Glu no vírus DEN-2 16681 é conservado em todos os flavivírus transmitidos por mosquitos, exceto para o vírus da febre amarela.

Modo de administração

As quimeras virais aqui descritas são combinadas

individualmente ou em conjunto com um transportador ou veículo farmacêuticamente aceitável para a administração como um imunogénio ou uma vacina para seres humanos ou animais. Os termos "transportador farmacêuticamente aceitável" ou "veículo farmacêuticamente aceitável" são aqui utilizados para indicar qualquer composição ou composto, incluindo, mas não se limitando a, água ou uma solução salina, um gel, uma pomada, um solvente, um diluente, uma base de pomada fluida, lipossomas, micelas, micela gigante e similares, que é adequado para utilização em contato com tecido vivo humano ou animal sem causar respostas fisiológicas adversas e que não interage de modo prejudicial com os outros componentes da composição.

As formulações imunogénicas ou de vacinas podem ser convenientemente apresentadas em forma de dosagem unitária de ufp virais e preparadas através de técnicas farmacêuticas convencionais. Tais técnicas incluem a etapa de colocar em associação o ingrediente ativo e o(s) transportador(es) ou excipiente(s) farmacêutico(s). Em geral, as formulações são preparadas por colocação em associação de forma uniforme e íntima o ingrediente ativo com os transportadores líquidos. As formulações adequadas para a administração parentérica incluem soluções aquosas e não aquosas para injeção estéril que podem conter antioxidantes, tampões, bacteriostáticos e solutos que tornam a formulação isotónica com o sangue do recetor pretendido, e suspensões estéreis aquosas e não aquosas que podem incluir agentes de suspensão e agentes espessantes. As formulações podem ser apresentadas em recipientes de dose unitária ou de multidosagem, por exemplo, ampolas e frascos selados, e podem ser armazenadas numa condição liofilizada requerendo apenas a adição do transportador líquido estéril, por exemplo, água para injeções, imediatamente antes da utilização. As soluções e suspensões injetáveis extemporâneas podem ser preparadas a

partir de pós estéreis, grânulos e comprimidos vulgarmente utilizados por um perito vulgar na tecnologia.

As formulações de dosagem unitária preferidas são aquelas que contêm uma dose ou unidade, ou uma fração apropriada da mesma, do ingrediente administrado. Deve ser entendido que para além dos ingredientes particulares mencionados anteriormente, as formulações podem incluir outros agentes normalmente utilizados por um perito vulgar na tecnologia.

A composição imunogénica ou de vacina pode ser administrada através de diferentes vias, tais como oral ou parentérica, incluindo, mas não se limitando a, bucal e sublingual, retal, parentérica, aerossol, nasal, intramuscular, subcutânea, intradérmica e tópica. A composição pode ser administrada sob diferentes formas, incluindo, mas não se limitando a, soluções, emulsões e suspensões, microesferas, partículas, micropartículas, nanopartículas e lipossomas. Espera-se que possam ser necessárias a partir de cerca de 1 até cerca de 5 doses por regime de imunização. As doses iniciais podem variar desde cerca de 100 até cerca de 50.000 ufp, com uma gama de dosagem preferida de cerca de 500 até cerca de 20.000 ufp, uma gama de dosagem mais preferida de desde cerca de 1000 até cerca de 12.000 ufp e uma gama de dosagem ainda mais preferida de cerca de 1000 até cerca de 4000 ufp. As injeções de reforço podem variar na dosagem desde cerca de 100 até cerca de 20.000 ufp, com uma gama de dosagem preferida de cerca de 500 até cerca de 15.000, uma gama de dosagem mais preferida de cerca de 500 até cerca de 10.000 ufp, e uma gama de dosagem ainda mais preferida de cerca de 1000 até cerca de 5000 ufp. Por exemplo, o volume de administração irá variar consoante a via de administração. As injeções intramusculares podem variar em volume desde cerca de 0,1 ml até 1,0 ml.

A composição pode ser armazenada a temperaturas desde

cerca de -100 °C até cerca de 4 °C. A composição pode também ser armazenada num estado liofilizado a diferentes temperaturas, incluindo a temperatura ambiente. A composição pode ser esterilizada através de meios convencionais conhecidos por um perito vulgar na tecnologia. Tais meios incluem, mas não estão limitados a, filtração. A composição pode também ser combinada com agentes bacteriostáticos, tal como o timerosal, para inibir o crescimento bacteriano.

Calendário de administração

A composição imunogénica ou de vacina aqui descritas podem ser administradas a seres humanos, em especial a indivíduos que viajam para as regiões onde está presente a infeção pelo vírus do dengue, e também a habitantes dessas regiões. O tempo ótimo para a administração da composição é de cerca de um a três meses antes da infeção inicial. No entanto, a composição também pode ser administrada depois da infeção inicial para melhorar a progressão da doença, ou depois da infeção inicial para tratar a doença.

Adjuvantes

Pode ser administrada uma variedade de adjuvantes conhecidos de um vulgar perito na tecnologia em conjunto com o vírus quimérico numa composição imunogénica ou de vacina. Tais adjuvantes incluem, mas não estão limitados aos seguintes: polímeros, copolímeros, tais como copolímeros de polioxietileno - polioxipropileno, incluindo copolímeros em bloco, o polímero P1005, o adjuvante completo de Freund (para animais), o adjuvante incompleto de Freund; monooleato de sorbitano, esqualeno, adjuvante CRL-8300, alúmen, QS 21, dipéptido de muramilo, motivos de oligonucleótidos CpG e combinações de motivos de oligonucleótidos CpG, trealose, extratos bacterianos, incluindo extratos nicobacterianos, endotoxinas destoxificadas, lípidos de membrana, ou combinações dos mesmos.

Sequências de ácido nucleico

As sequências de ácido nucleico dos vírus DEN-1 16007, DEN-1 PDK-13, DEN-2 16681 (SEQ ID NO: 13), DEN-2 PDK-53 (SEQ ID NO: 15), DEN-3 16562, DEN-3 PGMK-30 / FRhL-3, DEN-4 1036 e DEN-4 PDK-13 são úteis para a concepção de sondas de ácido nucleico e de oligonucleótidos iniciadores para a detecção do vírus do dengue numa amostra ou espécime, com elevada sensibilidade e especificidade. As sondas ou oligonucleótidos iniciadores que correspondem a cada subtipo viral podem ser utilizados para detetar a presença do vírus do DEN-1, do vírus do DEN-3 e do vírus do DEN-4, respetivamente, para detetar a infeção pelo vírus do dengue, em geral, na amostra, para o diagnóstico de infeção com o vírus do dengue, para distinguir entre os vários subtipos do vírus do dengue, para quantificar a quantidade de vírus do dengue na amostra, ou para monitorizar o progresso das terapias utilizadas para o tratamento de uma infeção pelo vírus do dengue. O ácido nucleico e as sequências de aminoácidos correspondentes são também úteis como ferramentas de investigação de laboratório para estudar os organismos e as doenças e para desenvolver terapias e tratamentos para as doenças.

As sondas de ácidos nucleicos hibridam seletivamente com moléculas de ácido nucleico que codificam para os vírus DEN-1, DEN-3 e DEN-4 ou sequências complementares dos mesmos. Por "seletivo" ou "seletivamente" entende-se uma sequência que não hibrida com outros ácidos nucleicos para impedir a detecção adequada do vírus do dengue. Portanto, na concepção de ácidos nucleicos para hibridação, a seletividade irá depender dos outros componentes presentes numa amostra. O ácido nucleico que hibrida deve ter pelo menos 70 % de complementaridade com o segmento de ácido nucleico ao qual se hibrida. Tal como é aqui utilizado para descrever os ácidos nucleicos, o termo "hibrida seletivamente" exclui os ácidos nucleicos

ocasionais que hibridam aleatoriamente, e, portanto, tem o mesmo significado que "hibrida especificamente". O ácido nucleico que hibrida seletivamente pode ter pelo menos 70 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 %, e 99 % de complementaridade com o segmento da sequência ao qual se hibrida, de preferência 85 % ou mais.

Também são divulgadas sequências, sondas e oligonucleótidos iniciadores que hibridam seletivamente com o ácido nucleico codificante ou com a cadeia complementar, ou oposta do ácido nucleico. A hibridação específica com o ácido nucleico pode ocorrer com pequenas modificações ou substituições no ácido nucleico, desde que seja mantida a capacidade de hibridação específica da espécie funcional. Por "sonda" entende-se sequências de ácidos nucleicos que podem ser utilizadas como sondas ou oligonucleótidos iniciadores para a hibridação seletiva com sequências de ácidos nucleicos complementares para a sua detecção ou amplificação, em que as sondas podem variar em comprimento de cerca de 5 a 100 nucleótidos, ou de preferência a partir de cerca de 10 a 50 nucleótidos, ou mais preferencialmente de cerca de 18 - 24 nucleótidos. Por conseguinte, os termos "sonda" ou "sondas" tal como são aqui utilizados são definidos para incluir "oligonucleótidos iniciadores". São aqui descritos os ácidos nucleicos isolados que hibridam seletivamente com os ácidos nucleicos específicos da espécie em condições rigorosas e devem ter, pelo menos, cinco nucleótidos complementares à sequência de interesse, tal como descrito em *Molecular cloning: A laboratory manual*, 2^a ed., Sambrook, Fritsch e Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nova Iorque, (1989).

Se utilizada como oligonucleótidos iniciadores, a composição inclui de preferência, pelo menos, duas moléculas de ácido nucleico que hibridam com regiões diferentes da molécula alvo de forma a amplificar uma

região desejada. Dependendo do comprimento da sonda ou do oligonucleótido iniciador, a região alvo pode oscilar entre 70 % de bases complementares e complementaridade completa, e mesmo assim hibridar sob condições rigorosas. Por exemplo, para o propósito de detetar a presença do vírus do dengue, o grau de complementaridade entre o ácido nucleico de hibridação (sonda ou oligonucleótido iniciador) e a sequência à qual se hibrida é pelo menos suficiente para distinguir a hibridação com um ácido nucleico a partir de outros organismos.

As sequências de ácidos nucleicos que codificam para os vírus DEN-1, DEN-3 ou DEN-4 podem ser inseridas num vetor, tal como um plasmídeo, e expressas de forma recombinante num organismo vivo para produzir os péptidos e/ou polipéptidos recombinantes do vírus do dengue.

Métodos de detecção de ácidos nucleicos

É aqui descrito um teste genético rápido que é diagnóstico para cada um dos vírus de vacina aqui descritos. Isto melhora as análises dos vírus isolados a partir do soro de seres humanos vacinados que desenvolveram uma viremia, bem como aumenta a caracterização de viremia em primatas não humanos imunizados com o vírus candidato a vacina da Mahidol.

Um sistema automatizado de detecção de sequências de ácido nucleico com base na PCR recentemente desenvolvido, é o ensaio TaqMan (Applied Biosystems), que é cada vez mais amplamente utilizado nos laboratórios de diagnóstico. O ensaio TaqMan é um ensaio altamente específico e sensível que permite a visualização e quantificação em tempo real, automatizada de amplicões gerados por PCR a partir de um molde de ácido nucleico da amostra. O ensaio TaqMan pode determinar a presença ou a ausência de uma sequência específica. Neste ensaio, são concebidos um oligonucleótido iniciador no sentido direto e outro no sentido inverso para hibridarem a montante e a jusante do

local da mutação alvo, respetivamente. Uma sonda de deteção específica, a qual foi concebida para ter uma temperatura de fusão de cerca de 10 °C mais elevada do que qualquer um dos amplímeros, e que contém a mutação nucleotídica específica do vírus de vacina ou o seu complemento (dependendo da cadeia do amplicão de RT/PCR que está a ser detetada), constitui o terceiro componente primário do presente ensaio.

A sonda é um detetor ou um oligonucleótido repórter fluorescente que contém um corante repórter-5' e um corante de extinção-3'. A extremidade 5' do nucleótido está ligada a um de uma série de diferentes corantes repórter fluorescentes, tais como FAM (6-carboxifluoresceína) ou TET (tetracloro-6-carboxifluoresceína). Na extremidade 3' da sonda, o corante de extinção TAMRA (6-carboxitetrametilrodamina) está ligado através de um ligante. O corante de extinção suprime a fluorescência do corante repórter na sonda intacta, em que ambos os corantes estão em estreita proximidade. Se a sequência alvo específica da sonda está presente no amplicão de RT/PCR, a sonda vai hibridar entre os locais do oligonucleótido iniciador de sentido direto e reverso no amplicão. Se a sonda hibrida com a sequência alvo, a atividade de nuclease 5'-3' da polimerase de ADN AmpliTaq Gold (Applied Biosystems) cliva a sonda entre o corante repórter e o corante de extinção. A polimerase não vai digerir a sonda livre. Uma vez que a polimerase desloca a sonda, a polimerização e a amplificação por PCR continuam. Uma vez separado do corante de extinção, o corante repórter produz uma fluorescência que é medida pelo sistema de deteção de sequência ABI PRISM. Se a sequência alvo específica da sonda está presente no amplicão, o nível de fluorescência aumenta, e é medido automaticamente em cada ciclo de amplificação por PCR.

Uma sonda concebida para detetar especificamente um

locus mutado num dos genomas virais de vacina da Mahidol irá conter os nucleótidos específicos da vacina no meio da sonda. Esta sonda irá resultar em fluorescência detetável no ensaio TaqMan se o molde de ARN viral é específico do vírus de vacina. No entanto, os moldes de ARN genómico de vírus DEN do tipo selvagem terão uma diminuição da eficiência da hibridação da sonda devido ao desfasamento de um único nucleótido (no caso dos vírus DEN parentais) ou possivelmente mais do que um desfasamento (tal como pode ocorrer em outros vírus DEN do tipo selvagem), e não irá resultar em fluorescência significativa. A polimerase de ADN é mais provável que desloque uma sonda com desfasamento do molde do amplicão de RT/PCR do que clivar a sonda com desfasamento para libertar o corante repórter (ensaio TaqMan de discriminação alélica, Applied Biosystems).

Uma estratégia desenvolvida mais recentemente para testes genéticos de diagnóstico faz uso de sinais moleculares (Tyagi e Kramer, *Mature Biotechnology* 14: 303-308 (1996)). A estratégia de sinal molecular também utiliza oligonucleótidos iniciadores para a amplificação por RT/PCR de amplicões, e a deteção de uma sequência específica no interior do amplicão por uma sonda que contém os corantes repórter e de extinção nos terminais da sonda. Neste ensaio, a sonda forma uma estrutura de haste e ansa. O corante repórter do terminal-5' e -3' e o corante de extinção, respetivamente, estão localizados nos terminais da estrutura de haste curta, o que coloca o corante de extinção em estreita justaposição com o corante repórter. A estrutura de haste funde-se durante a etapa de desnaturação do ensaio de RT/PCR. Se o ARN viral alvo contém a sequência alvo e é amplificado pelo amplímero de sentido direto e reverso, a ansa aberta da sonda hibrida com a sequência alvo durante a etapa de emparelhamento do ciclo. Quando a sonda é emparelhada a qualquer das cadeias

do molde do amplicão, os corantes de extinção e repórter são separados, e é detectada a fluorescência do corante repórter. Este é um ensaio de identificação e quantificação em tempo real que é muito semelhante ao ensaio TaqMan. O ensaio de sinais moleculares emprega corantes de extinção e repórter que diferem dos que são utilizados no ensaio TaqMan.

A presente invenção é ainda ilustrada pelos exemplos não limitativos que se seguem, que não são para ser interpretados de modo algum como impondo limitações ao âmbito da mesma. A invenção é definida pelas reivindicações em anexo. O objeto do exemplo 5 é incluído para fins comparativos.

Exemplo 1

Reparação de uma vacina de dengue-1 quimérica: PDK-53 / Dengue-1

Vírus e culturas de células

Os vírus DEN-1 16007 e DEN-2 16681 do tipo selvagem estavam disponíveis na coleção de vírus nos Centers for Disease Control and Prevention (Atlanta, GA). O vírus DEN-1 16007 foi recuperado a partir do soro de um paciente com FHD / SCD em 1964 na Tailândia. O vírus foi isolado após três passagens em células BS-C-1 de rim de macaco Grivet e de uma passagem em células LLC-MK₂, passadas duas vezes em mosquitos *Toxorhynchites amboinensis* e, em seguida passadas em células primárias de rim de cão (PDK) no Center for Vaccine Development, Universidade de Mahidol, na Tailândia, para derivar o vírus candidato a vacina DEN-1 PDK-13 (Yoksan et al., "Dengue virus vaccine development: study on biological markers of uncloned dengue 1-4 viruses serially passaged in primary kidney cells", pp. 35-38, em Arbovirus research na Austrália, Proceedings of the Fourth Symposium. CSIRO / QIMR, Brisbane, Austrália (1986); Bhamarapavati & Sutee, Vaccine 18: 44-47 (2000)). Salvo indicação em contrário,

foi usada uma única passagem em LLC-MK₂ deste vírus candidato a vacina (lote 10 de março de 1989). Após as passagens em mosquitos mencionadas anteriormente, o vírus 16007 foi passado uma vez em células LLC-MK₂ para utilização.

Os vírus foram cultivados em células LLC-MK₂ e C6/36 em meio essencial mínimo modificado por Dulbecco (DMEM) contendo penicilina / estreptomicina e 5 % de soro fetal de bovino (FBS). Foram realizadas titulações de placas virais em placas de 6 poços de células Vero ou LLC-MK₂ confluentes, tal como descrito anteriormente (Miller et al., Am. J. Trop. Med. & Hyg. 35: 1302-1309 (1986)). O primeiro meio de sobreposição de 4 ml - contendo 1 % de agarose SeaKem LE (FMC BioProducts, Rockland, Maine), em meio nutriente (0,165 % de hidrolisado de lactalbumina [Difco Laboratories, Detroit, Mich.]), 0,033 % de extrato de levedura [Difco], solução salina equilibrada de Earl, 25 mg/l de sulfato de gentamicina [Bio Whittaker, Walkersville, Md.], 1,0 mg/l de anfotericina B [Fungizone, E. R. Squibb & Sons, Princeton, N.J.], e 2 % de FBS) - foi adicionado após a adsorção de 200 µl do inóculo de vírus, durante 1,5 h a 37 °C. Após a incubação a 37 °C durante 7 dias, foi adicionado um segundo meio de sobreposição de 2 ml, contendo 80 µg/ml adicionais de corante vital vermelho neutro (GIBCO-BRL, Gaithersburg, Md.). As placas foram contadas 8 a 11 dias após a infecção.

Construção de clones infecciosos D2/1 quiméricos

Vetores pD2-16681-P48, pD2-PDK53-E48, pD2-PDK53-V48

Foram utilizados três vetores DEN-2 para a construção dos clones D2/1 quiméricos. Estes foram modificados a partir dos clones infecciosos DEN-2 relatados por Kinney et al., (Virology 230: 300-308 (1997)). O clone pD2-16681-P48 foi modificado a partir de pD2 / IC-30P-A para conter locais de clonagem MluI e NgoMIV nas posições do nucleótido 451 e 2380, respectivamente. Os mesmos locais de

clonagem foram introduzidos em ambos os clones específicos do vírus DEN-2 PDK-53, pD2 / IC-130Vx-4 e -130Vc-K, e os clones modificados foram designados por pD2-PDK53-E48 e pD2-PDK53-V48, respectivamente. Foram encontrados dois erros de clonagem no pD2 / IC-130Vx-4 original e -130Vc-K nos nt-6665 e nt-8840. Estes defeitos foram corrigidos em pD2-PDK53-E48 e -V48. O local de clonagem NgoMIV introduzido resultou em duas mutações de nucleótidos (nt 2381 e 2382; TG para CC), que codificou uma substituição de Val para Ala em E-482. As alterações de nucleótidos introduzidas no local MluI foram silenciosas. O local MluI introduzido na junção C / prM foi utilizado para clonar os genes prM-E dos vírus heterólogos.

pD2 / 1-PP, -EP, -VP, -PV, -EV, e -VV quiméricos

Foram construídos dois clones intermédios de DEN-2, pD2I-P e pD2I-E através da exclusão dos fragmentos de HpaI (nt-2676) a XbaI (terminal 3' do ADNc do genoma viral) de pD2-16681-P48 e pD2-PDK53-E48, respectivamente. Estes clones intermédios foram utilizados para subclonar os fragmentos de ADNc específicos do vírus DEN-1. Os fragmentos de ADNc que contêm os genes C-prM-E dos vírus DEN-1 16007 ou PDK-13 foram amplificados por reação em cadeia da polimerase por transcriptase reversa (RT-PCR) a partir do ARN viral de DEN-1 com oligonucleótidos iniciadores DEN-Bgl1.5NC (5'-TAGAGAGCAGATCTCTG-3' (SEQ ID NO: 31); sequência conservada no 5'NCR de genomas virais do dengue, a sequência sublinhada é um local BglII) e cD1-2394.Ngo: (5'-TGTGACCATGCCGCTGCGATGCACTCACCGA-3' (SEQ ID NO: 32); o local NgoMIV sublinhado seguido pela sequência complementar próxima da extremidade 3' do gene E do vírus DEN-1). Os fragmentos amplificados foram clonados nos locais BglII - NgoMIV dos clones intermédios pD2I-P e pD2I-E. Os clones pD2/1 intermédios, quiméricos foram sequenciados para verificar a precisão da inserção do ADNc específico do vírus DEN-1. Os fragmentos excisados a

partir dos clones intermédios pD2/1 com SstI (que precede o promotor de T7) e NgoMIV foram clonados nos vetores de genoma de comprimento completo de DEN-2, pD2-16681-P48, pD2-PDK53-E48, e pD2-PDK53-V48. Foram construídos seis plasmídeos de genoma de comprimento completo de pD2/1 quimérico através da inserção da região do gene C-prM-E do vírus DEN-1 16007 ou PDK-13 nestes três vetores (fig. 1). Os plasmídeos foram designados de pD2/1-XY e os seus derivados de vírus foram designados de DEN2/1-XY, em que X = fundo do clone DEN-2 infeccioso (P = clone 16881 parental, E = variante PDK53-E, V = variante PDK53-V) e Y = inserção de C-prM-E específica do vírus DEN-1 (P = estirpe 16007 parental, V = candidato a vacina de PDK-13). A estrutura de PDK53-E contém três mutações de aminoácidos específicas do vírus DEN-2 PDK-53 (NS1-53-Asp, NS2A-181-Phe e NS4A-75-Ala), bem como a mutação 5'NC-57 do vírus PDK-53. A estrutura de PDK53-V contém esses mesmos loci específicos do vírus PDK-53 mais o local adicional NS3-250-Val específico do vírus PDK-53.

Recuperação de vírus recombinantes

Todos os plasmídeos recombinantes foram crescidos em células XL1-Blue de *Escherichia coli*. O ARN viral recombinante foi transcrito e tapado com o análogo de tampa ("cap") m⁷ GpppA de 200 - 400 ng de ADNC linearizado com XbaI, e transfetado em 3 - 4 x 10⁶ células LLC-MK₂ ou BHK-21 por eletroporação. As células transfetadas foram transferidas para frascos de 75 cm² em meio DMEM contendo 10 % de FBS. As proteínas virais expressas nas células transfetadas foram analisadas por ensaio de imunofluorescência indireta (IFA). As células infectadas com vírus foram fixadas em acetona gelada durante 30 minutos. Foram usados no ensaio os anticorpos monoclonais 1F1 e 3H5 específicos do vírus DEN-1 e DEN-2, respetivamente, e a ligação foi detetada com o anticorpo de cabra anti-ratinho marcado com fluoresceína. Os vírus

foram recolhidos após 8 a 10 dias, e depois foram passados uma vez em células LLC-MK₂ (DEN-2-16681-P48, DEN-2 PDK53-E48 e -V48, DEN-2/1-PP, - EP, e -VP) ou duas vezes (DEN-2/1-PV, -EV, e - VV) para se obter sementes de trabalho. Os vírus D2/1-EV e - VV foram passados uma terceira vez em células LLC-MK₂ para se obter títulos virais mais elevados necessários para o desafio ou a imunização de ratinhos.

Caraterização do fenótipo de replicação de vírus quiméricos em culturas de células

Os tamanhos das placas foram medidos 10 dias após a infecção em células LLC-MK₂. Os diâmetros médios de placa foram calculados a partir de 10 placas de cada vírus. As curvas de crescimento virais foram realizadas em frascos de 75 cm² em células LLC-MK₂ ou C6/36 a aproximadamente 0,001 multiplicidade de infecção (moi). Após adsorção durante 2 h, foram adicionados 30 ml de meio DMEM (para células LLC-MK₂) ou meio nutriente de sobreposição (para células C6/36) contendo 5 % de FBS e penicilina / estreptomicina, e as culturas foram incubadas em 5 % de CO₂ a 37 °C ou 29 °C, respetivamente. Foram recolhidas alíquotas de meio de cultura em intervalos de 48 h durante 12 dias, ajustadas a 12,5 % de FBS, e armazenadas a -80 °C antes da titulação.

A sensibilidade à temperatura foi testada em células LLC-MK₂. As células cultivadas em dois conjuntos de frascos de 75 cm² foram infetadas e incubadas tal como descrito para o estudo da curva de crescimento. Um conjunto de culturas foi incubado durante 8 dias a 37 °C, o outro a 38,7 °C. Foi calculada a razão de título do vírus a 38,7 °C versus o título a 37 °C. O vírus foi designado como sensível à temperatura se o título de vírus a 38,7 °C foi reduzido de 60 % ou superior, em relação ao seu título a 37 °C.

Sequenciamento do ADNc viral

O ARN viral foi extraído a partir de sementes de

vírus ou utilizando o kit de ARN viral QIAmp (Qiagen, Valencia, Calif.). Os oligonucleótidos iniciadores específicos do vírus DEN-1 basearam-se nos dados publicados da estirpe S275/90 de Singapura (Fu et al., Virology 88: 953-958 (1992)). Foram amplificados por RT-PCR cinco a sete fragmentos de ADNc viral que se sobrepõem com o sistema de RT-PCR Titan One-Tube (Roche Molecular Biochemicals, Indianapolis, Ind.). Foram sequenciadas diretamente ambas as cadeias dos amplicões de ADNc. Para a sequenciação do genoma viral de DEN-1 PDK-13, o molde de ARN genômico foi extraído diretamente de um frasco do candidato a vacina de DEN-1 PDK-13 (lote de 10 de março de 1989). As sequências terminais 5' e 3' dos genomas virais de DEN-1 16007 e DEN-1 PDK-13 foram determinadas com o kit 5' RACE (GIBCO BRL) e colocando uma cauda de poli (A) no ARN genômico. A sequenciação automática foi realizada tal como recomendado num sequenciador PRISM 377 (Perkin-Elmer / Applied Biosystems, Foster City, Calif.).

Estudos com ratinhos

As ninhadas de recém-nascidos com 1 dia de ratinhos ICR brancos não consanguíneos foram inoculadas por via intracraniana com 5.000 ufp de vírus num volume de 30 µl. Os ratinhos foram observados diariamente para paralisia e morte, e os ratinhos que sobreviveram foram pesados individualmente uma vez por semana durante 5 semanas.

As respostas de anticorpos neutralizantes foram testadas em ratinhos ICR com 3 semanas. Os ratinhos foram inoculados por via intraperitoneal com 10^4 ufp de vírus, e foram reforçados com a mesma quantidade de vírus 3 semanas ou 6 semanas mais tarde. Os ratinhos foram sangrados 2 dias antes do reforço e 3 semanas após o reforço.

A estirpe de ratinhos ICR consanguíneos utilizada anteriormente não é normalmente suscetível fatalmente ao desafio com o vírus DEN-1 do tipo selvagem. Por conseguinte, para testar completamente a capacidade da

quimera viral DEN-2/1 para induzir uma resposta imune eficaz, foi necessário utilizar ratinhos AG-129 consanguíneos, que não têm recetores tanto para o interferão alfa / beta como para o interferão gama (Muller et al., Science 264: 1918-1921 (1994)), uma vez que estes ratinhos foram encontrados serem suscetíveis ao desafio intraperitoneal com doses elevadas do vírus DEN-1 do tipo selvagem, estirpe Mochizuki. Portanto, os ratinhos AG-129 consanguíneos com 3,5 - 4,5 semanas foram imunizados por via intraperitoneal com 10^4 ufp de vírus DEN-1 16007 do tipo selvagem, candidato a vacina DEN-1 PDK-13 da Mahidol, DEN-2/1-EP quimérico ou DEN-2/1-VP quimérico. Estes ratinhos imunizados foram desafiados por via intraperitoneal com uma dose letal de vírus DEN-1 Mochizuki.

Ensaio de neutralização

As amostras de soro de ratinho foram testadas para os anticorpos neutralizantes pelo teste de neutralização de diluição de soro - redução de placas (PRNT). Foram incubadas sessenta ufp do vírus DEN-1 16007 com diluições em série de 2 vezes de amostras de soro de ratinho inativado pelo calor (56 °C durante 30 minutos) durante a noite a 4 °C. O título do anticorpo neutralizante foi identificado como a diluição de soro mais elevada que reduziu por 50 % ou mais o número de placas de vírus no ensaio.

Resultados

Para avaliar a capacidade dos clones de ADNc infecciosos derivados das duas variantes do vírus DEN-2 16681 PDK-53 (PDK-53-E e PDK-53-V) para servirem como vetores para o desenvolvimento de vacinas, nós concebemos clones de ADNc de DEN-2 / DEN-1 quiméricos (D2/1-EP, D2/1-VP, D2/1-EV, e D2/1-VV) contendo os genes estruturais (C-prM-E) do vírus DEN-1 16007 do tipo selvagem ou do seu derivado de vacina, estirpe PDK-13, dentro da estrutura

destes dois vetores (fig. 1). Foram também construídos para comparação dois outros clones quiméricos, D2/1-PP e D2/1-PV, contendo os genes estruturais (C-prM-E) dos vírus DEN-1 16007 ou PDK-13 na estrutura do vírus DEN-2 16681 do tipo selvagem (fig. 1). Sequenciámos todo o ADNc genómico de comprimento completo em todos os clones infecciosos. Foi incorporado um artefacto de ADNc silencioso nos clones quiméricos no nt-297 (T para C). Foi concebida uma mutação silenciosa no nt-1575 (T para C) em todos os clones quiméricos para remover o local natural de XbaI no gene E do vírus DEN-1.

Os títulos após a transfeção de células LLC-MK₂ ou BHK-21 foram de 10^4 - 10^6 ufp/ml para os vírus quiméricos D2/1-PP, -EP, e -VP contendo o C-prM-E do vírus DEN-1 16007. Estes títulos aumentaram para $10^{6,5}$ - $10^{7,5}$ ufp/ml após uma única passagem em células LLC-MK₂, comparáveis aos títulos obtidos para os vírus parentais. Foram obtidos títulos inferiores de 10^2 - 10^4 ufp/ml em células transfectadas para os vírus DEN-2/1-PV, -EV, e -VV quiméricos contendo o C-prM-E do vírus DEN-1 PDK-13. O vírus D2/1-PV atingiu 10^6 ufp/ml depois de 2 passagens em células LLC-MK₂, enquanto que o vírus D2/1-EV e -VV atingiu títulos de $10^{3,3}$ - $10^{5,3}$ ufp/ml após duas ou três passagens. As células infetadas com qualquer um dos vírus DEN-2/1 quiméricos foram positivas por IFA com o anticorpo monoclonal 1F1 (específico para a proteína E do DEN-1) e negativas com o anticorpo monoclonal 3H5 (específico para a proteína E do DEN-2), indicando que foram expressas pelas quimeras as proteínas E de DEN-1 adequadas. Os genomas virais de DEN-2/1-PP, DEN-2/1-EP, e DEN-2/1-VP foram sequenciados completamente através da análise direta de fragmentos de RT-PCR que se sobrepõem amplificados a partir do ARN genómico viral extraído de sementes mestre. Todos os três genomas tinham a sequência esperada.

Crescimento dos vírus quiméricos em cultura de células

LLC-MK₂ e C6/36

Todos os vírus DEN-2/1 quiméricos produziram placas mais pequenas, em relação à placa de $6,8 \pm 0,4$ mm do vírus DEN-1 16007 do tipo selvagem em células LLC-MK₂ (fig. 2A). Ambas as placas virais de DEN-2/1-EP ($3,1 \pm 0,3$ mm) e -VP ($2,8 \pm 0,3$ mm) foram semelhantes em tamanho às do vírus DEN-1 PDK-13 ($2,9 \pm 0,3$ mm). Os vírus quiméricos DEN-2/1-PV, DEN-2/1-EV, e DEN-2/1-VV contendo o C-prM-E do vírus DEN-1 PDK-13 formaram placas pequenas ($1,3 \pm 0,3$ mm) ou de cabeça de alfinete (< 1 mm). O vírus DEN-2 16681-P48 produziu placas de $3,5 \pm 0,3$ mm que eram similares às placas do vírus DEN-2 16681 do tipo selvagem. O vírus DEN-2 PDK53-V48 formou placas que eram mais pequenas e menos nítidas do que as do vírus DEN-2 PDK53-E48. As placas de $5,1 \pm 0,3$ mm do vírus DEN-2/1-PP foram maiores do que aquelas dos outros vírus quiméricos, mas mais pequenas do que as do vírus DEN-1 16007.

Os vírus foram testados quanto à sensibilidade à temperatura em células LLC-MK₂. A sensibilidade à temperatura foi determinada no dia 8 ou 10 após a infecção, tal como indicado na fig. 2B. A sensibilidade à temperatura foi baseada na redução dos títulos de vírus a 38,7 °C daqueles a 37 °C. A sensibilidade à temperatura foi calculada a partir de medições feitas em pelo menos 3 experiências.

A variante DEN-2 PDK53-V (D2-PDK53-V48) foi mais sensível à temperatura do que o vírus DEN-2 PDK53-E (D2-PDK53-E48) e o vírus DEN-2 16681 (DEN-2-16681-P48) foi um pouco sensível à temperatura (70 - 87 % de redução do título a 38,7 °C). Vários testes de sensibilidade à temperatura para o vírus DEN-2-PDK53-E48 resultaram em 83 % - 97 % de redução do crescimento. O título do vírus DEN-1 16007 foi reduzido em 40 % ou menos a 38,7 °C, tornando-o o vírus menos sensível à temperatura no presente estudo. Todos os vírus quiméricos foram sensíveis à temperatura em

relação ao vírus DEN-1 16007, e foram pelo menos tão sensíveis à temperatura como o PDK-13.

Todos os vírus DEN-1, DEN-2, e D2/1 quiméricos atingiram títulos de pico entre os 8 e 10 dias após a infecção em células LLC-MK₂ (fig. 2B). Os vírus derivados de clones IDEN-2-16681-P48, DEN-2-PDK53-E48, e DEN-2-PDK53-V48 replicaram-se para $10^{7,0}$ ufp/ml ou mais, tal como os o fizeram os vírus DEN-2 16681 e PDK-53. Embora tenha alcançando um título de pico semelhante, o vírus DEN-2-PDK53-V48 replicou-se mais lentamente do que o vírus DEN-2-PDK53-E48 durante os primeiros 4 dias após a infecção. Os vírus DEN-2/1-PP, DEN-2/1-EP, DEN-2/1-VP, e DEN-2/1-PV quiméricos alcançaram títulos de pico de mais de $10^{6,7}$ ufp/ml, comparáveis aos títulos de pico do seu vírus DEN-1 e DEN-2 parental. Os vírus DEN-2/1-EV e -VV quiméricos, que tiveram títulos de pico de $10^{5,6}$ - $10^{5,9}$ ufp/ml ou inferiores em várias experiências separadas, replicaram-se menos eficientemente do que os outros vírus.

As quimeras específicas do vírus PDK-13 resultam em títulos de vírus mais baixos recuperados a partir de células transfectadas, em relação às quimeras específicas do vírus 16007. As experiências anteriores com as quimeras DEN-2 / DEN-1 e DEN-2 / DEN-4 indicam que os vírus quiméricos que exibem replicação deficiente durante a transfeção e depois desenvolvem títulos de vírus altos após a passagem em cultura de células, acumulam muitas vezes mutações inesperadas. Os vírus com aumento da capacidade replicativa podem surgir através da seleção de subpopulações de variantes do vírus, resultantes da incorporação de erros durante a transcrição *in vitro* do ADNc. Os vírus quiméricos que se replicam bem em células transfectadas eram mais estáveis geneticamente após a passagem em células LLC-MK₂. A replicação eficiente com uma passagem mínima em cultura de células de mamífero pode ser um critério importante de estabilidade genética e

adequação para um vírus de vacina derivado do clone infeccioso.

As curvas de crescimento viral em células C6/36 foram monitorizadas após a infecção de células C6/36 a uma multiplicidade de aproximadamente 0,001 ufp/ml (fig. 3). Os três vírus de estrutura DEN-2, DEN-2-16681-P48, DEN-2-PDK53-E48 e DEN-2-PDK53-V48, replicaram-se como o vírus DEN-2 16681 original e as duas variantes PDK-53, respetivamente. Ambos os vírus DEN-2-PDK53-E48 e DEN-2-PDK53-V48 replicaram-se cerca de 4000 vezes menos eficientemente do que os vírus DEN-2-16681-P48, DEN-1 16007 e PDK-13 em células C6/36. Os vírus DEN-1 16007 e PDK-13 replicaram-se para títulos elevados de $10^{8,7}$ e $10^{8,4}$ ufp/ml, respetivamente. O vírus DEN-2/1-PP quimérico replicou-se para $10^{5,2}$ ufp/ml, o que foi equivalente aos títulos de pico das duas variantes de DEN-2-PDK53. A replicação dos vírus DEN-2/1-EP e DEN-2/1-VP quiméricos foi muito ineficiente em células C6/36. Estes vírus atingiram títulos de pico inferiores a 10^2 ufp/ml.

Neurovirulência em ratinhos lactentes

Grupos de ratinhos ICR de recém-nascidos a um dia de idade (n = 16) foram inoculados por via intracraniana com 5000 ufp do vírus. O vírus DEN-2 16681 do tipo selvagem foi 100 % fatal com o tempo de sobrevivência médio aos $14,1 \pm 1,6$ dias, enquanto ambos os vírus DEN-2-PDK53-E48 e DEN-2-PDK53-V48 derivados de clone não conseguiram matar nenhum ratinho ICR. Ao contrário do vírus DEN-2 16681, que normalmente mata 50 % ou mais dos ratinhos desafiados, o vírus DEN-1 16007 do tipo selvagem causou apenas uma única letalidade (21 dias após o desafio) nos ratinhos ICR. O vírus DEN-1 PDK-13 não matou nenhum dos ratinhos ICR. Os ratinhos ICR infetados com o vírus DEN-1 16007 tiveram pesos corporais significativamente menores ($p < 0,00003$, teste t de Student), em relação ao grupo de controlo inoculado com diluente, entre 21 - 35 dias após o desafio.

Todos os grupos de ratinhos ICR desafiados com cinco vírus DEN-2/1 quiméricos (o vírus DEN-2/1-VV não foi testado) apresentaram pesos médios menores ($p < 0,02$) quando comparados com o grupo de controlo, mas os seus pesos médios foram significativamente maiores ($p < 0,004$) do que o grupo de DEN-1 16007 aos 28 dias após a infeção. Os pesos corporais médios dos grupos de ratinhos ICR desafiados com 10^4 ufp dos vírus DEN-1 16007 ou PDK-13 foram quase idênticos aos dos ratinhos ICR desafiados com 5000 ufp do vírus DEN-1 16007 entre 7 - 35 dias após o desafio.

Imunogenicidade de vírus DEN-2/1 quiméricos em ratinhos

Ratinhos ICR não consanguíneos: para testar a imunogenicidade dos vírus quiméricos foram imunizados grupos de ratinhos ICR de 3 semanas ($n = 8$), por via intraperitoneal com 10^4 ufp de vírus na experiência 1 e na experiência 2 (tabela 7). Na experiência 1, os ratinhos foram sangrados 20 dias após a imunização primária e depois foram reforçados 2 dias mais tarde. Na experiência 2, os ratinhos foram sangrados 41 dias após a imunização primária e foram reforçados 2 dias mais tarde. A tabela 7 mostra os títulos recíprocos de PRNT de ponto final de 50 % de redução de placas das amostras de conjuntos de soro obtidas a partir de cada grupo imunizado. É também mostrada a gama de títulos individuais para os oito ratinhos em muitos dos grupos. Em ambas as experiências, o título recíproco do conjunto de soro dos ratinhos imunizados com vírus 16007 foi de 80 antes do reforço e 2560 depois do reforço. Em ambas as experiências, os ratinhos imunizados com os vírus DEN-2/1-PP, -EP, ou DEN-2/1-VP quiméricos tiveram um título de conjunto de soro que foi pelo menos tão alto como o dos grupos imunizados com o vírus 16007 antes (títulos recíprocos de 40 - 160) e depois (títulos recíprocos de 2560 - 5210) do reforço. As respostas imunitárias dos ratinhos nestes grupos de vírus

foram quase equivalentes em ambas a experiência 1 e a experiência 2.

O vírus PDK-13 e todas as três quimeras que contêm os genes estruturais de PDK-13 induziram títulos recíprocos de PRNT mínimos ou baixos de 10 - 20 contra o vírus DEN-1 16007, aos 20 dias após a imunização primária na experiência 1. Um título recíproco de PRNT um pouco mais elevado de 40 foi induzido por cada um destes vírus aos 41 dias após a imunização na experiência 2. O desenvolvimento de anticorpos neutralizantes foi mais lento e de magnitude inferior a seguir à imunização com os vírus PDK-13, DEN-2/1-PV, -EV, ou -VV do que com os vírus 16007, DEN-2/1-PP, -EP, ou -VP nestes ratinhos. Um ratinho no grupo DEN-2/1-PV na experiência 1 e um ratinho em cada um dos grupos DEN-2/1-EV em ambas as experiências falhou em produzir um título de PRNT detetável antes do reforço. Os títulos de reforço também foram maiores para os grupos imunizados com PDK-13, DEN-2/1-PV, -EV, e -VV na experiência 2 (conjunto de títulos recíprocos de 160 - 2560) do que na experiência 1 (conjuntos de títulos recíprocos de 80). Exceto para o grupo DEN-2/1-PV na experiência 2, estes títulos de reforço foram menores do que os títulos de PRNT de reforço induzidos pelo vírus 16007 do tipo selvagem e os vírus DEN-2/1-PP, -EP e -VP quiméricos que contêm os genes estruturais do vírus DEN-1 16007 do tipo selvagem (tabela 7). O elevado título de PRNT obtido com o conjunto de soro de ratinhos reforçados com o vírus DEN-2/1-PV resultou do soro de dois ratinhos que tinham títulos recíprocos de 2560 e 10.240. Os 6 ratinhos restantes neste grupo apresentaram títulos recíprocos de 20 a 640, que foram semelhantes aos títulos individuais de ratinhos nos grupos PDK-13, DEN-2/1-EV, e -VV. Os vírus PDK-13, DEN-2/1-PV, DEN-2/1-EV, e DEN-2/1-VV pareciam ser menos imunogênicos do que os vírus 16007, DEN-2/1-PP, -EP, e -VP nestes ratinhos não consanguíneos. As amostras de conjunto de

soro obtidas a partir de ratinhos imunizados com os vírus DEN 2-16681-P48, DEN-2-PDK-53-E48, ou DEN-2-PDK-53-V48 não continham anticorpo neutralizante cruzado detetável contra o vírus DEN-1 16007.

Ratinhos AG-129 consanguíneos: para testar a imunogenicidade dos vírus quiméricos, foram imunizados grupos de ratinhos AG-129 consanguíneos com 3,5 - 4,5 semanas de idade por via intraperitoneal com 10^4 ufp de vírus DEN-1 16007 do tipo selvagem, candidato a vacina de DEN-1 PDK-13 da Mahidol, DEN-2/1-EP quimérico, ou DEN-2/1-VP quimérico. Aos 26 dias após a imunização primária, o conjunto de soros de ratinhos imunizados com os vírus DEN-2/1-EP ou DEN-2/1-VP quiméricos tinham títulos recíprocos de anticorpos neutralizantes de 70 % de diluição de soro - redução de placas (PRNT₇₀) de 80 - 160. Estes títulos foram equivalentes (dentro de 2 vezes) da resposta de anticorpos neutralizantes desencadeada pelo candidato a vacina de DEN-1 PDK-13, mas menores do que o título de 640 provocado pelo vírus DEN-1 16007 do tipo selvagem. Os ratinhos de controlo injetados por via intraperitoneal com solução salina tamponada com fosfato apresentaram títulos inferiores a 10. Todos os ratinhos foram desafiados por via intraperitoneal com 10^7 ufp de vírus DEN-1 do tipo selvagem, virulento, estirpe Mochizuki, aos 28 dias após a imunização primária. Todos os onze ratinhos de controlo, não imunizados morreram no espaço de 21 dias (tempo médio de sobrevivência de $11,5 \pm 4,2$ dias [média $\pm$ desvio padrão]) após o desafio, enquanto que todos os ratinhos imunizados com os vírus de DEN-1 (16007, PDK-13, DEN-2/1-EP, DEN-2/1-VP) sobreviveram o desafio com o vírus DEN-1 Mochizuki. Aos 34 dias após o desafio com o DEN-1 Mochizuki, todos os ratinhos imunizados com o vírus tinham títulos recíprocos de PRNT₇₀ de 1280 - 2560. Os vírus DEN-2/1-EP e DEN-2/1-VP quiméricos eram altamente imunogénicos e protetores para ratinhos AG-129.

Análises de sequência de nucleótidos de genomas virais de DEN-1 16007 e PDK-13

Nós sequenciámos os genomas do vírus DEN-1 16007 do tipo selvagem (número de acesso GenBank AF180817) e do seu derivado de vacina PDK-13 (número de acesso AF180818). Havia diferenças de 14 nucleótidos e 8 aminoácidos codificados entre os vírus 16007 e PDK-13 (tabela 2). Ocorreram mutações silenciosas no genoma nas posições de nucleótido 1567, 2695, 2782, 7330, e 9445 nos genes E, NS1, NS4B, e NS5. Ao contrário do vírus candidato a vacina de DEN-2 PDK-53, que não tem mutações de aminoácidos na proteína E, o vírus DEN-1 PDK-13 tinha cinco mutações de aminoácidos em E, incluindo E-130 Val para Ala, E-203 Glu para Lys, E-204 Arg para Lys, E-225 Ser para Leu, e E-447 Met para Val. As mutações de aminoácidos nos genes não estruturais incluíram NS3-182 Glu para Lys, NS3-510 Tyr para Phe, e NS4A-144 Met para Val. A E-477-Val específica do vírus PDK-13 foi incorporada em todas as construções quiméricas.

Imunização de macacos com vírus DEN-2/1 quiméricos

Foi testada a imunogenicidade dos vírus DEN-2/1-EP e DEN-2/1-VP quiméricos em macacos adultos comedores de caranguejos (*Macaca fascicularis*). A imunização dos macacos foi realizada por meio de injeção subcutânea com 10^6 ufp de vírus DEN-2/1-EP ou DEN-2/1-VP quimérico. Os soros obtidos a partir dos macacos imunizados foram analisados para a presença de viremia por ensaio direto de placas de alíquotas de soro em monocamadas de células LLC-MK₂ mantidas sob sobrecamada de agarose em placas de 6 poços, e por inoculação de alíquotas de soro em culturas de células LLC-MK2 mantidas em meio líquido. Não foi identificado nenhum vírus infeccioso em qualquer dos soros de macaco por qualquer destes dois métodos clássicos de ensaio. Através destes dois ensaios de vírus não foi detetável viremia após a imunização com qualquer vírus

DEN-2/1-EP ou DEN-2/1-VP quimérico. Os soros de macaco foram testados para anticorpos neutralizantes específicos do vírus DEN-1. Aos 30 dias após a imunização primária, os soros de três macacos individuais imunizados com o vírus DEN-2/1-EP quimérico tinham títulos recíprocos de anticorpos neutralizantes de 50 % de diluição de soro - redução de placas (PRNT₅₀) de 80, 160 e 1280. Os soros de três macacos individuais imunizados com DEN-2/1-VP quimérico tinham títulos recíprocos de PRNT₅₀ de 160, 160 e 640. Os macacos injetados com diluente como um controle para a experiência tinham títulos recíprocos de PRNT₅₀ inferiores a 10, assim como todos os macacos imediatamente antes da imunização. Os vírus DEN-2/1-EP e -VP quiméricos induziram anticorpos neutralizantes específicos do vírus DEN-1 em primatas não humanos.

Exemplo 2

Construção de clones infecciosos de DEN-2/3 quimérico

Foi utilizada uma estratégia de ligação *in vitro* para os clones infecciosos de DEN-2/3 de genoma de comprimento completo.

(i) Clones intermédios de DEN-2/3 da extremidade 5': pD2I / D3-P1-Asc e pD2I / D3-E1-Asc

Foram usados clones intermédios de DEN-2, pD2I-P e pD2I-E para subclonar o fragmento de ADNc específico do vírus DEN-3 16562. O fragmento de ADNc que contém os genes prM-E do vírus DEN-3 16562 do tipo selvagem foi amplificado pela reação em cadeia da polimerase por transcriptase reversa (RT-PCR) a partir de ARN viral de DEN-3 com os oligonucleótidos iniciadores D3-435.Mlu: (5'-TGCTGGCCACTTA^{ACT}TACGCGTGATGGAGAGCCGCGCA-3' (SEQ ID NO: 33); local MluI sublinhado seguido pela sequência do vírus DEN-3 perto da extremidade 5' do gene prM) e cD3-2394.Ngo: (5'-TGTAATGATGCCGCGCGGCGATGCATGAAAATGA-3' (SEQ ID NO: 34); local NgoMIV sublinhado seguido pela sequência complementar próxima da extremidade 3' do gene E do vírus

DEN-3). O local MluI continha uma mutação silenciosa para o vírus DEN-2 no aminoácido prM-5-Thr. Esta posição é Ser no vírus DEN-3, mas Thr no vírus DEN-2/3 quimérico. O local NgoMIV resultou numa substituição Val para Ala em E-482 do vírus DEN-2, e numa substituição Ile para Ala em E-480 do vírus DEN-3. O fragmento amplificado foi clonado nos locais MluI - NgoMIV dos clones intermédios pD2I-P e pD2I-E. Foi introduzido um local de restrição, AscI, a jusante do local NgoMIV por mutagénese dirigida ao local para facilitar a ligação *in vitro*. Este local único AscI estava apenas a 16 nt a jusante do local do NgoMIV e foi excisado antes da ligação *in vitro* dos clones DEN-2/3 de comprimento completo. Foi introduzida uma mutagénese adicional no nt 1970 (A para T), que mudou o aminoácido E-345 de His para Leu, para permitir a derivação de vírus quiméricos viáveis em células LLC-MK₂, tal como explicado em seguida. Estes clones intermédios de DEN-2/3, pD2I / D3-P-Asc e pD2I / D3-E-Asc quiméricos foram sequenciados para verificar a precisão da inserção do ADNc específico do vírus DEN-3. Foi incorporada uma mutação silenciosa no nt 552 (C para T), em ambos os clones intermédios quiméricos.

(ii) Clones intermédios de DEN-2 da extremidade 3': pD2-Pm^b-Asc, pD2-Em^b-Asc, e pD2-Vm^b-Asc

Os clones intermédios de DEN-2 que contêm a sequência de ADNc específica do vírus DEN-2 truncado do nt 2203 - 10723 (extremidade 3') dos vírus DEN-2 16681, PDK53-E, ou PDK-53-V foram obtidos pela deleção da extremidade 5' (incluindo a sequência do promotor de T7 e os nt 1 - 2202 de DEN-2) do ADNc específico do vírus nos clones de comprimento completo, pD2-16681-P48, pD2-PDK53-E48, e pD2-PDK53-V48, respetivamente. Também foi introduzido um local AscI no nt 2358 (22 nt a montante do local NgoMIV) para facilitar a ligação *in vitro*. Este local AscI foi excisado antes de se efetuar a ligação *in vitro* dos clones

infeciosos de DEN-2/3 quimérico de genoma de comprimento completo.

(iii) ADNc de DEN-2/3 quimérico de comprimento completo: DEN-2/3-PP1, DEN-2/3-EP1, e DEN-2/3-VP1

Foram digeridas de três a dez mg da extremidade 5' de pD2I/D3 e da extremidade 3' de clones intermédios D2 por AscI, tratadas com fosfatase de intestino de vitela (CIP), e, em seguida, digeridas com NgoMIV. Os pequenos fragmentos AscI - NgoMIV excisados foram removidos por passagem do ADN digerido através das colunas spin de purificação de PCR da Qiagen. As grandes extremidades 5' e 3' dos clones intermédios linearizados foram então ligadas em conjunto (razão molar de 5' : 3' = 1:3) para se obter DEN-2/3-PP1 de genoma de comprimento completo (pD2I / D3-P1-Asc ligado com pD2-Pm^b-Asc), DEN-2/3-EP1 (pD2I / D3-E1-Asc ligado com pD2-Em^b-Asc), e DEN-2/3-VP1 (pD2I / D3-E1-Asc ligado com pD2-Vm^b-Asc). Estes ADN ligados foram em seguida cortados com XbaI para produzir a extremidade 3' linearizada do ADNc viral requerida para a transcrição do ARN viral recombinante.

Recuperação de vírus DEN-2/3 quiméricos

O ARN viral recombinante foi transcrito a partir do ADNc linearizado com XbaI e tapado com o análogo de tampa ("cap") m⁷ GpppA. As células LLC-MK₂ ou BHK-21 (3 - 5 x 10⁶ células) foram transfetadas por eletroporação, tal como descrito por Kinney et al., (J. Virol 230: 300-308 (1997)). As células transfetadas foram transferidas para frascos de 75 cm² em meio DMEM contendo 10 % de FBS. As proteínas virais expressas nas células transfetadas foram analisadas por ensaio de imunofluorescência indireta (IFA). As células infetadas com vírus foram fixadas em acetona gelada durante 30 minutos. Foram usados no ensaio os anticorpos monoclonais 8A1 e 3H5 específicos dos vírus DEN-3 e DEN-2, respetivamente, e a ligação foi detetada com o anticorpo de cabra anti-ratinho marcado com

fluoresceína. Os vírus foram recolhidos depois de 5 a 11 dias, e foram depois passados uma vez em células LLC-MK₂ para se obter sementes de trabalho. Os genomas virais extraídos a partir destas sementes foram sequenciados para confirmar os genótipos dos vírus de progenitura. Uma estratégia anterior utilizando clones intermédios da extremidade 5' de D2I/D3 contendo os genes prM-E autênticos de DEN-3 16562 resultou em mutações em várias posições diferentes nos genomas dos vírus recuperados a partir de células LLC-MK₂ ou BHK-21 transfectadas e passadas uma vez em células LLC-MK₂. Foi encontrada uma mutação no nt 1970 de A para T, a qual mudou o aminoácido E-345 de His para Leu, em sete de nove vírus recombinantes recuperados independentemente. Era óbvio que os vírus DEN-2/3 quiméricos originais eram instáveis em células LLC-MK₂ e/ou BHK-21. A única mutação em E-345 foi a única mutação que ocorreu nos genomas dos três vírus recuperados, indicando que esta mutação pode ajudar a estabilizar os vírus em cultura. Introduzimos esta mutação em todos os clones infecciosos de DEN-2/3, e os vírus recombinantes recuperados a partir de tais clones mutagenizados provaram ser estáveis em cultura de células.

Exemplo 3

Construção de clones infecciosos de DEN-2/4 quiméricos

Vetores pD2-16681-P48, pD2-PDK53-E48, pD2-PDK53-V48

Os três vetores de estrutura de DEN-2 utilizados para a construção dos clones de DEN-2/4 quiméricos foram modificados tal como descrito anteriormente. Os clones infecciosos de DEN-2 foram relatados previamente em Kinney et al., 1997. O clone pD2-16681-P48 foi modificado a partir de pD2 / IC-30P-A para conter os locais de clonagem MluI e NgoMIV nas posições de nucleótido 451 e 2380, respetivamente. Foram introduzidos os mesmos locais de clonagem em ambos os clones específicos do vírus DEN-2 PDK-53, pD2 / IC-130Vx-4 e -130Vc-K, e os clones

modificados foram designados como pD2-PDK53-E48 e pD2-PDK53-V48, respetivamente. Foram encontrados dois erros de clonagem no pD2 / IC-130Vx-4 original e -130Vc-K nos nt-6665 e nt-8840. Estes defeitos foram corrigidos no pD2-PDK53-E48 e -V48. O local de clonagem NgoMIV introduzido resultou em duas mutações de nucleótidos (nt 2381 e 2382; TG para CC), que codificaram para uma substituição Val para Ala em E-482 do vírus DEN-2. As alterações de nucleótidos introduzidas no local MluI foram silenciosas para o vírus DEN-2 e os vírus DEN-2/4 quiméricos. O local MluI (perto da junção C/prM) e o local NgoMIV (perto da junção E/NS1) foram utilizados para clonar os genes prM-E dos vírus heterólogos.

pDEN-2/4-PP1, -EP1, e -VP1 quiméricos

Foram construídos dois clones intermédios de DEN-2, pD2I-P e pD2I-E, através da exclusão dos fragmentos de HpaI (nt-2676) a XbaI (terminal 3' do ADNc do genoma viral) de pD2-16681-P48 e pD2-PDK53-E48, respetivamente. Estes clones intermédios foram utilizados para subclonar os fragmentos de ADNc específicos do vírus DEN-4 1036. O fragmento de ADNc que contém os genes prM-E do vírus DEN-4 1036 foi amplificado pela reação em cadeia da polimerase por transcriptase reversa (RT-PCR) a partir do ARN viral de DEN-4 com os oligonucleótidos iniciadores D4-453Mlu: (5'-GGCGTTTCACTTGTCAACGCGTGATGGCGAACCCCTCA-3' (SEQ ID NO: 35); local MluI sublinhado seguido pela sequência próxima da extremidade 5' do genoma viral de DEN-4 1036) e cD4-2394.Ngo: (5'-AGTGATTCCGCCGGCAGCTATGCACGTCATAGCCAT-3' (SEQ ID NO : 36); local NgoMIV sublinhado seguido pela sequência complementar próxima da extremidade 3' do gene E do vírus DEN-4). Os fragmentos amplificados foram clonados nos locais de MluI - NgoMIV dos clones intermédios pD2I-P e pD2I-E. Os clones intermédios de DEN-2/4 quiméricos foram sequenciados para verificar a precisão da inserção do ADNc específico do vírus DEN-4. Foi incorporada uma

mutação silenciosa no nt 1401 (A para G) em ambos os clones intermédios. Os fragmentos excisados dos clones DEN-2 / IC-P48, -VE48, e -VV48 com NgoMIV e ScaI (a jusante da sequência de ADNc de DEN-2 nestes plasmídeos) foram clonados nos clones intermédios D2/4 quiméricos cortados com NgoMIV / ScaI para se obter os clones D2/4-PP, -EP, e -VP quiméricos de comprimento total. No entanto, o ARN transcrito a partir destes clones quiméricos só produziu vírus quiméricos DEN-2/4 viáveis em células C6/36 transfetadas, mas não em células LLC-MK₂ transfetadas. Após passar mais uma vez os vírus quiméricos em células C6/36 para se obter títulos mais elevados, estes vírus foram passados cinco vezes em células LLC-MK₂ para se obter vírus DEN-2/4 quiméricos estáveis que se replicam de forma eficiente em células LLC-MK₂. Os títulos das sementes de vírus aumentaram de 200 ufp/ml na primeira LLC passagem em células LLC-MK₂ para mais de 10⁶ ufp/ml na quinta passagem em células LLC-MK₂.

Foram sequenciados os genomas dos vírus da primeira, segunda, terceira e quinta passagens em células LLC-MK₂ dos vírus DEN-2/4-PP quiméricos. Foram identificadas quatro mutações, C-100 (Arg para Ser) específica do vírus DEN-2, E-364 (Ala para mistura de Ala / Val) específica do vírus DEN-4, E-447 (Met para Leu) específica do vírus DEN-4 e NS4B-239 (Ile para Leu) específica do vírus DEN-2 (posições de aminoácidos com base em sequências do vírus DEN-2/4 quimérico). As três mutações localizadas nos genes estruturais (C e E) foram introduzidas nos clones de ADNc infeccioso de DEN-2/4-PP, -EP, e -VP quiméricos para se obter os clones DEN-2/4-PP1, -EP1, e -VP1, respetivamente. Todos os três clones DEN-2/4 quiméricos produziram vírus quiméricos de elevados títulos, viáveis em células LLC-MK₂ imediatamente após a transfeção, indicando que estas três mutações ajudaram os vírus a se replicarem em células LLC-MK₂. Os clones DEN-2/4 quiméricos foram também

mutagenizados para conter combinações diferentes das quatro mutações para determinar quais as mutações que são necessárias para a eficiência da replicação dos vírus DEN-2/4 quiméricos em células LLC-MK₂. A mutação DEN-4 E-447 (Met para Leu) sozinha ou em conjunto com a mutação DEN-4 E-364, em combinação com a mutação DEN-2 C-100 (Arg para Ser) foi adequada para permitir a derivação do vírus DEN-2/4 em células LLC-MK₂.

Recuperação de vírus recombinantes.

Os plasmídeos recombinantes pD2/4-PP1, -EP1, e -VP1 foram todos cultivados em células XL1-Blue de *Escherichia coli*. O ARN viral recombinante foi transcrito e tapado com o análogo de tampa ("cap") m⁷ GpppA de 200 - 400 ng de ADNc linearizado com XbaI, e transfetado em 3 - 5.10⁶ células LLC-MK₂. As células transfetadas foram transferidas para frascos de 75 cm² em meio DMEM contendo 10 % de FBS. As proteínas virais expressas nas células transfetadas foram analisadas pelo ensaio de imunofluorescência indireta (IFA). As células infetadas com vírus foram fixadas em acetona gelada durante 30 minutos. Foram usados no ensaio os anticorpos monoclonais 1H10 e 3H5 específicos do vírus DEN-4 e DEN-2, respetivamente, e a ligação foi detetada com o anticorpo de cabra anti-ratinho marcado com fluoresceína. Os vírus foram recolhidos após 8 a 10 dias, e foram então passados uma vez em células LLC-MK₂ para se obter sementes de trabalho. Os genomas de todas estas três sementes de trabalho foram completamente sequenciados, e todos os genomas virais de DEN-2/4 quiméricos continham as sequências esperadas. As sequências de nucleótidos e de aminoácidos para a quimera DEN 2/4 são aqui proporcionadas em SEQ ID NO: 11 e SEQ ID NO: 12, respetivamente.

Exemplo 4

Caraterização dos vírus DEN-2/3 e DEN-2/4 quiméricos

Os vírus infecciosos DEN-2/3 e DEN-2/4 quiméricos, viáveis, os quais expressam a região do gene prM/E do

vírus DEN-3 16562 do tipo selvagem ou do vírus DEN-4 1036 do tipo selvagem, respetivamente, expressam os epítomos apropriados da proteína do invólucro (E) específica do vírus DEN-3 ou DEN-4, tal como analisado por imunofluorescência indireta de células LLC-MK₂ infetadas com vírus com anticorpos monoclonais anti-E específicos do vírus. Estes vírus quiméricos, assim como o vírus DEN-2/1 quimérico, também expressam epítomos de neutralização apropriados específicos do serotipo de DEN, quando testados contra fluidos ascíticos polivalentes padrão de ratinho ou anticorpos monoclonais nos testes de neutralização de diluição de soro - redução de placas (tabela 8). Os vírus DEN-2/1-EP, DEN-2/3-EP1, e DEN-2/4-EP1 quiméricos foram neutralizados por estes anticorpos padrão específicos do vírus DEN para títulos recíprocos de PRNT₅₀ que foram pelo menos tão elevados como aqueles que ocorreram para os vírus DEN-1, DEN-3 e DEN-4 do tipo selvagem, respetivamente (tabela 8). Estes dados de neutralização indicaram que os vírus DEN-2/1, DEN-2/3, e DEN-2/4 quiméricos expressam epítomos de neutralização apropriados específicos do serotipo de DEN de vírus DEN-1, DEN-3 e DEN-4, respetivamente.

Replicação em células LLC-MK2

A replicação e a sensibilidade à temperatura das quimeras de DEN-2/3 em células LLC-MK2 foram monitorizadas como a replicação e a sensibilidade à temperatura das quimeras DEN-2/1 foram analisadas no exemplo 1. Os vírus DEN-2/3-PP1, -EP1, e -VP1 quiméricos e os vírus DEN-2/4-PP1, -EP1, e -VP1 quiméricos, que expressam a região do gene prM/E do vírus DEN-3 16562 ou do vírus DEN-4 1036 no fundo genético de DEN-2 16681 (-PP1), variante DEN-2 PDK-53-E (-EP1), ou a variante DEN-2 PDK-53-V (-VP1), respetivamente, replicaram-se todos a títulos de pico elevados de pelo menos 10^{6,3} ufp/ml em células LLC-MK₂. Tal como definido em termos da redução dos títulos de vírus a

38,7 °C versus a 37 °C (-, +, 2+, 3+ indicam a redução do título inferior ou igual a 60 %, 61 - 90 %, 91 - 99 %, > 99 %, respetivamente, calculadas a partir de pelo menos 3 experiências). Os vírus DEN-2/3-PP1 e DEN-2/4-PP1 quiméricos exibiram quer uma sensibilidade à temperatura limítrofe (DEN-2/3-PP1) quer nenhuma sensibilidade à temperatura (DEN-2/4-PP1). Os vírus DEN-2/3-EP1 e -VP1 quiméricos, bem como os vírus DEN-2/4-EP1 e -VP1 quiméricos, foram todos construídos no fundo genético da variante DEN-2 PD-53-E ou -V, retiveram os fenótipos sensíveis à temperatura que foram exibidos pelas duas variantes virais de PDK-53 (vírus DEN-2-PDK53-E48 e -V48, respetivamente). As quimeras construídas no fundo da variante PDK-53-V exibiram um maior grau de sensibilidade à temperatura do que as construídas no fundo da variante PDK-53-E.

Tamanhos das placas em células LLC-MK2

O tamanho da placa resultante da inoculação de células LLC-MK2 colocadas sob sobrecamada de agarose, utilizado como marcador biológico para determinar a atenuação tal como descrito no exemplo 1, também foi examinado para os vírus DEN-2/3 e DEN-2/4 quiméricos. O tamanho médio de placa em mm segue cada vírus ou quimera entre parêntesis: DEN-3-16562 (6,6), DEN-2/3-PP1 (5,4), DEN-2/3-EP1 (4,5), DEN-2/3-VP1 (2,3), DEN-4-1036 (8,6), DEN-2/4-PP1 (3), DEN-2/4-EP1 (1,5), DEN-2/4-VP1 (1,2), DEN-2-16681-P48 (4,2), DEN-2-PDK53-E48 (2,5) e DEN-2-PDK-53-V48 (1,8).

Os tamanhos das placas dos vírus DEN-2/3-PP1 e DEN-2/4-PP1 exibiram diâmetros médios de placa que foram maiores do que os dos vírus DEN-2 16681 ou PDK-53 (tal como descrito no exemplo 1), mas menores do que os dos vírus DEN-3 16562 e DEN-4 1036 do tipo selvagem, respetivamente. Isto indica que os genes estruturais dos vírus dadores DEN-3 e DEN-4 e da cápside e/ou regiões de

genes não estruturais no fundo genético recetor do vírus DEN-2 16681 afetaram ambos o tamanho da placa. Os vírus DEN-2/3-EP1 e -VP1 e DEN-2/4-EP1 e -VP1 quiméricos exibiram reduções significativas no tamanho da placa, em relação aos vírus DEN-3 16562 e DEN-4 1036 do tipo selvagem, respetivamente. O efeito sobre o tamanho da placa específico do fundo de DEN-2 PDK-53 pode resultar de interação sinérgica das mutações nos loci NS1-53 e NS3-250 do vírus DEN-2 PDK-53. Consequentemente, as quimeras -VP1 apresentaram reduções maiores no tamanho da placa do que as quimeras -EP1. Os vírus DEN-2/1, DEN-2/3, e DEN-2/4 quiméricos construídos no fundo genético do vírus candidato a vacina de DEN-2 PDK-53 retiveram o fenótipo da diminuição do tamanho da placa, tal como exibido pelo vírus DEN-2 PDK-53.

Replicação em células C6/36

A capacidade do vírus se replicar em cultura de células C6/36 de mosquito, usada como marcador biológico para determinar a atenuação tal como descrito no exemplo 1, também foi examinada para os vírus DEN-2/3 e DEN-2/4 quiméricos. Os títulos de pico médios em unidades de $\log_{10}$ ufp/ml seguem cada vírus ou quimera entre parênteses; D3-16562 (7,3), DEN-2/3-PP1 (7,6), DEN-2/3-EP1 (5,2), DEN-2/3-VP1 (5,7), D4-1036 (8,7), DEN-2/4-PP1 (7,8), DEN-2/4-EP1 (4,5), DEN-2/4-VP1 (4,4), DEN-2-16681-P48 (8,3), DEN-2-PDK-53-E48 (5,4) e DEN-2-PDK-53-V48 (5).

Ambas as variantes candidatas de DEN-2 PDK-53 da Mahidol exibiram diminuição da eficiência de replicação, em relação ao vírus DEN-2 16681 do tipo selvagem, em cultura de células C6/36 de mosquito (vírus DEN-2-PDK-53-E48 e -V48 versus DEN-2-16681-P48). A diminuição da capacidade de replicação em células C6/36 foi atribuída às mutações nos dois loci 5'-NC-57 e NS1-53 no vírus DEN-2 PDK-53; coerente com esta visão, ambas as variantes replicaram-se a títulos de pico equivalentemente reduzidos

nestas células. Este fenótipo de replicação deficiente em células C6/36 foi retido nos vírus DEN-2/3-EP1 e -VP1 e DEN-2/4-EP1 e -VP1 quiméricos, replicando-se todos a títulos de pico inferiores do que os vírus DEN-3 16562 ou DEN-4 1036 do tipo selvagem, respectivamente. Os vírus DEN-2/3-PP1 e DEN-2/4-PP1 quiméricos, construídos no fundo genético do vírus DEN-2 16681 do tipo selvagem, replicaram-se essencialmente na mesma medida (DEN-2/3-PP1) ou um pouco inferior (DEN-2/4-PP1) do que os vírus DEN-3 16562 e DEN-4 1036 do tipo selvagem, respectivamente. Estes resultados indicam que o efeito de replicação deficiente dos loci 5'-NC-57 e NS1-53 no fundo específico do vírus DEN-2 PDK-53 foi preservado naqueles vírus quiméricos que foram construídos dentro do fundo genético de DEN-2 PDK-53.

Neurovirulência para ratinhos recém-nascidos

Ratinhos ICR brancos, recém-nascidos, não consanguíneos (n = 16 para cada grupo) foram desafiados por via intracraniana com 10^4 ufp dos vírus DEN-3 e DEN-4 do tipo selvagem, vírus candidatos a vacinas de DEN-3 e DEN-4 da Mahidol, vírus DEN-2/3-PP1, -EP1, -VP1 quiméricos, e vírus DEN-2/4-PP1, -EP1, e -VP1 quiméricos (tabela 9). Os vírus DEN-3 16562 e DEN-4 1036 do tipo selvagem, que causaram com fiabilidade 100 % de letalidade em ratinhos recém-nascidos, constituíram um modelo mais sensível para a atenuação de neurovirulência viral do que o modelo de desafio DEN-2 16681, que resulta em 50 - 100 % de letalidade em ratinhos recém-nascidos desafiados com este vírus. Curiosamente, o vírus candidato a vacina de DEN-4 PDK-48 da Mahidol também resultou em 100 % de letalidade em ratinhos desafiados, embora o tempo médio de sobrevivência destes ratinhos tenha sido de cerca de dois dias a mais do que para os ratinhos desafiados por via intracraniana com o vírus DEN-4 1036 do tipo selvagem (tabela 9). Tal como o vírus de vacina de DEN-2 PDK-53 e

ambas as suas populações de variantes, os vírus DEN-2/3-EP1 e -Vp1 quiméricos e os vírus DEN-2/4-EP1 e -VP1 quiméricos foram atenuados (sem mortes) para os ratinhos recém-nascidos. Esta atenuação pode ser devida, pelo menos em parte, ao fundo genético de DEN-2 PDK-53 atenuado destes vírus quiméricos, porque o vírus DEN-2/4-PP1 quimérico exibiu neurovirulência significativa (mortalidade de 62,5 %) nestes ratinhos. Esta última quimera foi construída no fundo genético do vírus DEN-2 16681 do tipo selvagem.

Imunização de ratinhos AG-129 com os vírus DEN-2/3 e DEN-2/4 quiméricos

Os ratinhos AG-129 consanguíneos foram imunizados por via intraperitoneal com 10^5 ufp do vírus candidato a vacina de DEN-3 ou DEN-4 do tipo selvagem ou da Mahidol ou com 10^5 ufp do vírus DEN-2/3-EP1 ou DEN-2/4-EP1 quimérico (tabela 10). O vírus DEN-2/3-EP1 quimérico provocou títulos recíprocos de PRNT₇₀ de 80 - 320 às 4 - 6 semanas após a imunização primária. Estes títulos foram essencialmente equivalentes aos provocados pelo vírus de vacina de DEN-3 da Mahidol (estirpe PGMK-30 / FRhL-3). No entanto, o vírus DEN-2/4-EP1 quimérico provocou um título de anticorpos neutralizantes de 20 - 40, que foi menor do que o título de 80 - 320 provocado pelo vírus candidato a vacina da Mahidol do vírus de vacina de DEN-4 PDK-48. No entanto, ambos os vírus DEN-2/3 e DEN-2/4 quiméricos induziram respostas de anticorpos neutralizantes em ratinhos AG-129.

Exemplo 5

Atenuação de um vírus de vacina de dengue-2: estirpe 16681 (PDK-53)

Vírus e culturas de células

Foram investigados o vírus DEN-2 16681 parental, várias passagens intermediárias de PDK (PDK-5, -10, -14, -35, e -45) do vírus 16681, os vírus recombinantes 16681 / PDK-53, e a passagem caracterizada geneticamente em LLC-MK₂-

1 do vírus candidato a vacina de PDK-53.

As culturas de células BHK-21 (clone 15), LLC-MK₂, Vero e C6/36 foram cultivadas em meio mínimo essencial modificado por Dulbecco (DMEM) suplementado com 10 % de soro fetal bovino (FBS; HyClone Laboratories, Inc., Logan, Utah) inativado pelo calor (56 °C durante 30 minutos), 3,7 g/l de bicarbonato de sódio (GIBCO-BRL, Life Technologies, Gaithersburg, Md.), 100 unidades / ml de penicilina G, e 100 mg/ml de sulfato de estreptomicina (GIBCO-BRL).

As titulações de placa foram realizadas em monocamadas confluentes de células Vero em placas de plástico de 6 poços, tal como descrito anteriormente (Miller et al., Am. J. Trop. Med. & Hyg. 35: 1302-1309 (1986)). Um inoculo de 200 µl de vírus foi adsorvido durante 1,5 h a 37 °C, seguido pela adição de 4 ml de meio de sobreposição de agarose contendo 1 % de agarose SeaKem LE (FMC BioProducts, Rockland, Md.) em meio de nutriente (0,165 % de hidrolisado de lactalbumina [Difco Laboratories, Detroit, Mich.], 0,033 % de extrato de levedura [Difco], solução de sal equilibrada de Earl, 25 mg/l de sulfato de gentamicina [Bio Whittaker, Walkersville, Md.], 1,0 mg/l de anfotericina B [Fungizone®, E. R. Squibb & Sons, Princeton, N.J.], e 2 % de FBS). Após a incubação a 37 °C durante 7 dias, foi adicionada uma segunda camada de 2 ml de agarose contendo 80 µg/ml de corante vital vermelho neutro (GIBCO-BRL).

Construção de vírus recombinantes DEN-2 16681 / PDK-53

Durante a validação genética dos vírus DEN-2 derivados de clones no presente estudo, foram descobertos dois erros de clonagem de ADNc, nt-6665 A para G (NS4A-97 Tyr para Cys) e nt-8840 A para G (NS5-424 Glu para Gly), no clone pD2 / IC-130Vc-K (variante NS3-250-Val) específico do vírus PDK-53 relatado anteriormente (Kinney et al., Virology 230: 300-308 (1997)). Estes defeitos foram corrigidos num clone específico do vírus PDK-53

recém-derivado (variante NS3-250-Val), pD2 / IC-VV45R.

Em estudos preliminares, foram utilizados vírus recombinantes 16681 / PDK-53 contendo regiões de genes específicas do vírus PDK-53 no fundo genético do vírus 16681 para investigar os loci genéticos envolvidos nos marcadores de atenuação do vírus PDK-53. As análises destes vírus indicaram que a mutação de PDK-53 no nt 57 na região 5'NC e as mutações de aminoácidos em NS1-53 (analisado de uma maneira ligada com a mutação NS2A-181) e NS3-250 eram os determinantes do fenótipo específico do vírus PDK-53. A mutação prM-29 tem pouco efeito sobre a virulência. Com base na análise de sequência e comparação, as mutações 5'NC, NS1 e NS3 foram submetidas a análise mutacional adicional. A mutação 5'NC ocorreu numa possível estrutura de haste. As mutações NS1 e NS3 ocorreram ambas nos loci conservados em alguns flavivírus. Foram construídos 14 plasmídeos recombinantes pD2 / IC-16681 / PDK-53 através da troca de fragmentos de ADNc entre pD2 / IC-30P-A (clone 16681) e pD2 / IC-VV45R (clone PDK-53) nos locais das enzimas de restrição SstI (que precede o promotor de T7), SalI (nt-165), SphI (nt-1380), SpeI (nt-2370 e nt-3579), KpnI (nt-4493), XhoI (nt-5426), e XbaI (extremidade 3' do clone). Todos os plasmídeos recombinantes foram crescidos em *Escherichia coli*, estirpe XLI-blue, e foram linearizados no local XbaI único concebido na extremidade 3' do ADNc. As células BHK-21 foram transfectadas com o ARN viral transcrito pelo método de Liljestrom et al., (J. Virology 63: 4107-4113 (1991)).

Os genótipos dos vírus D2 / IC-Px recombinante (em que x = 5, 1, e/ou 3 para indicar a incorporação dos loci 5'NC-57, NS 1-53, e/ou NS3-250 específicos do vírus 16681 parental [P na designação de vírus] na estrutura de pD2 / IC-VV45R [PDK-53]) e D2 / IC-Vx (em que x indica a incorporação recíproca dos loci específicos dos três vírus candidatos a vacina de PDK-53 [V na designação de vírus]

dentro da estrutura de pD2 / IC-30P-A [16681]) são apresentados na tabela 11. Se os loci 5'NC, NS1, e NS3 são os principais determinantes do fenótipo específico do vírus PDK-53, então os vírus D2 / IC-P5 e D2 / IC-V13 devem ser equivalentes (cognato), porque ambos os vírus contêm 5'NC-57-C, NS1-53-Asp, e NS3-250-Val. Os pares de vírus cognatos derivados a partir de mutagenese recíproca dos clones específicos do vírus 16681 e PDK-53 são indicados na tabela 11. Para investigar adicionalmente o locus prM-29, movemos o locus prM-29-Asp do vírus DEN-2 16681 para pD2 / IC-VV45R e pD2 / IC-P5 para derivar os vírus recombinantes D2 / IC-Pp e -P5p, respetivamente. As recombinações recíprocas produziram os vírus D2 / IC-Vp e -V5p.

Cada vírus derivado de clone (sementes BHK-21 transfetadas) foi propagado uma vez em células LLC-MK₂. Os genótipos de todos os vírus recombinantes 16681 / PDK-53, passados em LLC-MK₂-1 foram confirmados por análises completas de sequência de nucleótidos dos seus genomas. Uma vez que todos os vírus tinham as sequências de nucleótidos esperadas, nós deduzimos que os seus clones de ADNc eram também corretos. Todos os vírus derivados de clone continham o locus de nt-8571-C específico do vírus 16681, que é o local de uma mutação silenciosa no vírus PDK-53. Foi utilizada a sequenciação direta de amplicões de ADNc de sobreposição gerados a partir de ARN genómico viral de DEN-2 utilizando a reação em cadeia da polimerase por transcriptase reversa (RT-PCR), para determinar a sequência de todos exceto dos terminais do ADNc. As sequências da extremidade 5' e 3' dos 30 nucleótidos do genoma foram determinadas por sequenciação direta do clone de ADNc infeccioso no plasmídeo pBRUC-139. O prefixo D2 / IC- é eliminado nas designações de vírus na tabela 11 e no texto que se segue.

Caraterização dos fenótipos de replicação dos vírus

recombinantes 16681 / PDK-53.

Os vírus foram analisados para o tamanho da placa, a sensibilidade à temperatura, e a replicação em células LLC-MK₂ e C6/36. Os tamanhos das placas foram avaliados após 9 dias de incubação sob agarose em monocamadas de células LLC-MK₂ cultivadas em placas de 6 poços. As curvas de crescimento viral foram realizadas em frascos de 75 cm² de células LLC-MK₂ ou C6/36 inoculadas a uma multiplicidade de infecção (moi) de aproximadamente 0,001 ufp/célula. Após a adsorção a 37 °C durante 2 h, foram adicionados 30 ml de meio DMEM (células LLC-MK₂) ou meio de nutriente de sobreposição (células C6/36) contendo penicilina / estreptomicina e 5 % de FBS, e as culturas foram incubadas em 5 % de CO₂ a 37 °C ou a 29 °C, respectivamente. Foram removidas alíquotas de meio de cultura a intervalos de 48 h, foram ajustadas a 12,5 % de FBS, e foram armazenadas a -80 °C antes da titulação do vírus.

Os ensaios de sensibilidade à temperatura foram realizados em células LLC-MK₂ cultivadas em frascos de 75 cm². As células foram inoculadas a uma moi de aproximadamente 0,001 ufp/célula. Após a adsorção durante 2 h a 37 °C, foram adicionados 30 ml de meio DMEM contendo 5 % de FBS. Um conjunto de culturas foi incubado durante 8 dias a 37 °C, o outro a 38,7 °C. Foi calculada a razão do título do vírus a 38,7 °C versus o título a 37 °C.

Ensaio de neurovirulência em ratinho

As ninhadas de ratinhos ICR brancos recém-nascidos, não consanguíneos foram inoculadas por via intracraniana com 104 ufp de vírus num volume de 30 µl. Os ratinhos foram pesados individualmente uma vez por semana e foram observados para a paralisia ou morte durante 35 dias.

Fenótipos de placa dos vírus recombinantes 16681 / PDK-53

Foram medidos os diâmetros médios de placas de vírus (n = 12) aos 9 dias após a infecção sob sobrecamada de agarose em células LLC-MK₂ (fig. 4). As placas maiores

(diâmetro médio de 3,2 - 3,4) foram produzidas pelo vírus 16681 do tipo selvagem, pelo vírus 30P-A derivado do seu clone, e pelo vírus recombinante P513, que continham os loci 5'NC, NS1, e NS3 específicos do vírus 16681 no fundo genético do vírus VV45R (PDK-53) (tabela 11). Estes três loci específicos de 16681 dentro do fundo genético de PDK-53 foram suficientes para reconstituir o fenótipo de placa grande do vírus 16681.

As mutações 5'-NC-57-T, NS1-53-Asp, e NS3-250-Val específicas do vírus PDK-53 incorporadas no genótipo de 16681 resultaram cada uma em tamanhos de placa significativamente menores ($p < 0,00003$) de 2,1 - 2,3 mm em vírus V5, V1, e V3, respectivamente, versus aqueles do vírus 16681. Os fenótipos de placa dos vírus P13 e P51 foram semelhantes aos dos seus vírus cognatos V5 e V3 (indicado por barras de gráfico de padrão idêntico sólido ou de quadrícula transversal na fig. 4). O tamanho da placa de 2,8 mm do vírus P53 diferiu num grau menor ($p < 0,03$) das placas virais de 16681 do que as placas de 2,1 mm do seu vírus cognato V1. As placas de 1,1 - 1,6 mm dos pares de vírus cognatos do vírus V13 (P5) (contendo os loci NS1 e NS3 específicos do vírus PDK-53), V53 (P1) (loci 5'NC e NS3 específicos do vírus PDK-53) e V51 (P3) (loci 5'NC e NS1 específicos do vírus PDK-53) foram essencialmente equivalentes às placas de 1,3 mm do vírus PDK-53. Estes resultados indicam que todas as combinações de pares destes três loci específicos do vírus PDK-53 na estrutura de 16681 geraram um vírus de placas pequenas semelhantes ao vírus PDK-53. Embora o vírus V51 tenha produzido um fenótipo de placa que era semelhante ao do vírus PDK-53, todos os três loci 5'NC-C, NS1-53-Gly, e NS3-250-Glu específicos do vírus 16681 eram necessários dentro do fundo de PDK-53 do vírus P513 para reconstituir o fenótipo de placa do vírus 16681. A presença de todos os três loci de PDK-53 na estrutura de 16681 gerou o vírus

V513, que tinha um fenótipo de placa menor ($p < 0,001$, teste t de Student) do que o vírus PDK-53 ou W45R. A diferença entre o tamanho da placa do vírus PDK-53 e do vírus W45R não foi significativa.

O vírus recombinante Vp, que continha o locus prM-29-Val específico do vírus PDK-53 no fundo de 16681, tinha um tamanho de placa significativamente reduzido ($p < 0,002$) de 2,6 mm. No entanto, o fenótipo de placa de 1,0 mm do vírus Pp (locus prM-29-Asp 16681 no fundo de PDK-53) foi essencialmente idêntico às placas de 1,1 mm e 1,3 mm dos vírus VV45R e PDK-53 (fig. 4A), respetivamente. O locus prM-29 não afetou os fenótipos de placa dos vírus P5p ($1,0 \pm 0,0$ mm) e V5p ($2,0 \pm 0,4$ mm), que produziram placas semelhantes às de P5 ($1,1 \pm 0,3$ mm) e V5 ($2,3 \pm 0,4$ mm), respetivamente.

Replicação dos vírus recombinantes 16681 / PDK-53 em células LLC-MK₂ e C6/36

Todos os vírus replicaram-se bem em células LLC-MK₂, atingindo títulos de pico de $10^7 - 10^{7,9}$ ufp/ml aos 6 - 8 dias após a infeção. O PDK-53 e o vírus VV45R derivado do seu clone replicaram-se a uma taxa reduzida durante os primeiros quatro dias após a infeção, em relação aos outros vírus. Para determinar as sensibilidades à temperatura, as células LLC-MK₂ infetadas com o vírus foram incubadas a 37 °C ou a 38,7 °C em 2 - 5 experiências (fig. 4B). As pontuações de sensibilidade à temperatura foram determinadas aos 8 dias após a infeção. Todos os vírus exibiram algum grau de sensibilidade à temperatura sob estas condições. Em experiências individuais, o vírus 16681 do tipo selvagem mostrou uma redução do título de 75 - 80 % a 38,7 °C. O vírus V1 e os vírus cognatos P3 (V51), que mostraram reduções médias do título de 84 - 86 %, foram ligeiramente mais sensíveis à temperatura do que o vírus 16681. No entanto, apenas PDK-53, VV45R, V513, P5p, Pp, e os vírus recombinantes cognatos V13 e P5, os quais

continham ambos os loci NS1-53-Asp e NS3-250-Val específicos do vírus PDK-53, mostraram de forma reprodutível uma redução média do título de 90 - 97 % a 38,7 °C.

O vírus 16681 e o vírus 30P-A derivado do seu clone replicaram-se para títulos de pico médios de $10^{8,6}$ - $10^{8,8}$ ufp/ml aos 12 dias após a infecção em dois experiências independentes de curva de crescimento em células C6/36 (fig. 4C). A replicação do vírus PDK-53 (título de pico de $10^{4,5}$ ufp/ml) e do vírus VV45R derivado do seu clone (título de pico de $10^{4,6}$ ufp/ml) foi aproximadamente 15.000 vezes menos eficiente em células C6/36. Os loci 5'NC e NS1 específicos do vírus 16681 dentro do fundo de PDK-53 do vírus P51 reconstituíram totalmente a eficácia de replicação para a do vírus 16681 do tipo selvagem. Por outro lado, os loci 5'NC e NS1 específicos do vírus PDK-53 dentro do fundo de 16681 do vírus V51 foram suficientes para estabelecer o fenótipo de replicação deficiente do vírus PDK-53. Os pares de vírus cognatos recombinantes V5 (P13) e V1 (P53), que continham a região 5'NC ou o locus NS1 específicos do vírus PDK-53, respetivamente, replicaram-se para títulos de pico médios de $10^{5,9}$ - $10^{6,7}$ ufp/ml. Embora o título de pico médio do vírus V53 tenha sido cerca de 40 vezes maior do que o do vírus V5 em células C6/36, os títulos de pico dos vírus P513, P13, P53, V3, V13, V513, e P3 foram muito semelhantes aos dos vírus P51, P1, P5, 30P-A, V1, V51, e VV45R, respetivamente. Estes dados indicaram que o locus NS3-250 teve pouco ou nenhum efeito observável na replicação em células C6/36 (fig. 4C). Os vírus Vp e 16681 tiveram títulos de pico médios quase iguais em células C6/36, assim como os vírus P5 e P5p. Os vírus Pp e V5p produziram títulos de pico médios que foram ligeiramente superiores (8 e 40 vezes, respetivamente) do que os dos vírus PDK-53 e V5, respetivamente. O locus prM-29 pareceu ter pouco ou

nenhum efeito sobre a replicação viral em células C6/36.

Neurovirulência dos vírus recombinantes 16681 / PDK-53 em ratinhos recém-nascidos

Para investigar a neurovirulência dos vírus recombinantes, foram infetadas duas ninhadas de ratinhos ICR brancos recém-nascidos, oito ratinhos por ninhada, por via intracraniana com 10^4 ufp do vírus. O vírus DEN-2 16681 e o vírus 30P-A derivado do seu clone provocaram 50 % - 100 % de mortalidade nestes ratinhos. Os tempos de sobrevivência médios (AST) para os ratinhos que sucumbem ao desafio com o vírus 16681 ou 30P-A variaram de 15,2 a 16,8 dias em várias experiências. Os ratinhos foram pesados individualmente a cada 7 dias após a infecção. Morreu um único ratinho no dia 1 após a infecção, provavelmente como resultado de trauma da inoculação, em cada um dos grupos de P53 e Vp (tabela 12). Estes dois ratinhos foram excluídos das análises. Não houve mortes e nenhuma perda de peso no grupo de controle, inoculado com diluente.

Foram observados três fenótipos de neurovirulência em ratinho (tabela 12). O primeiro fenótipo consistiu dos vírus virulento de ratinho DEN-2 16681, 30P-A, P513, P51, V3, e Vp, que provocou pelo menos 50 % de mortalidade com AST de 13,2 - 17,0 dias (tabela 12). Em duas outras experiências independentes, o vírus Vp provocou 46,67 % de mortalidade com AST de $17,4 \pm 1,4$ dias ($n = 16$) e 56,25 % de mortalidade com AST de $18,3 \pm 1,3$ dias ($n = 16$). Numa experiência independente, o vírus P51 provocou apenas 25 % de mortalidade com AST de $15,9 \pm 5,5$ dias ($n = 16$). Um segundo fenótipo consistiu dos vírus atenuados em ratinho PDK-53, W45R, V513, Pp, e do cognato V51 (P3), que não provocaram nenhuma mortalidade, e os vírus quase atenuados V1, cognato V13 (P5), P5p, e V5p que mataram apenas 1 dos 16 ratinhos (tabela 12). A presença dos dois loci 5'NC-57-T e NS1-53-Asp específicos do vírus PDK-53 dentro do fundo

genético de 16681 foi suficiente para resultar na atenuação ou manter a mesma em vírus cognatos V51 (P3). Exceto para o vírus P53, todos os vírus que contêm o locus NS1-53-Asp específico do vírus PDK-53 foram atenuados ou quase atenuados.

O terceiro fenótipo, o de virulência intermédia, caracterizou os pares de vírus cognatos V5 (P13) e V53 (P1), e o vírus P53 que causaram 18,75 % - 37,5 % de mortalidade do ratinho e perda de peso significativa ($p < 0,001$, teste t de Student, às 3 semanas após a infecção, em relação aos ratinhos de controlo inoculados com o diluente) em ratinhos que sobreviveram o desafio com vírus. Os vírus V5 (P13) e V53 (P1) continham o locus 5'NC-57-T, mas não o locus NS1-53-Asp, do vírus PDK-53. O vírus V1 (6,25 % de mortalidade) foi mais atenuado do que o vírus V5, que produziu 18,75 % de mortalidade e perda de peso significativa nos sobreviventes. Por outro lado, o locus 5'-NC-57-C específico do vírus 16681 provocou pouca reversão para virulência no vírus P5 (6,25 % de mortalidade), enquanto que a porção NS1-53-Gly no vírus P1 resultou num nível intermédio (37,5 %) de mortalidade e perda de peso significativa nos sobreviventes. Ao contrário do vírus V1 cognato quase atenuado, o vírus P53 teve um fenótipo de virulência intermédia. O locus NS1-53 teve um efeito mais significativo sobre o fenótipo de virulência do que o locus 5'-NC-57.

O locus prM-29 não mostrou nenhum efeito nos vírus P5p, Pp, e Vp, em relação aos vírus P5, PDK-53, e 16681, respetivamente. O vírus V5p, que continha ambos os loci 5'NC-57-T e prM-29-Val específicos do vírus PDK-53, foi quase atenuado (tabela 2). O locus NS3-250 não pareceu contribuir significativamente para o fenótipo de neurovirulência em ratinho nos vírus V3, P13, V53, V13, e P3, que exibiram fenótipos que foram equivalentes aos vírus 16681, P1, V5, V1, e PDK-53, respetivamente. A

diferença no nível de mortalidade provocada pelos vírus P53 e P5 sugeriu que o locus NS3-250-Glu específico do vírus 16681 pode contribuir de alguma forma para o fenótipo de virulência dentro de certos contextos genéticos.

Evolução de mutações no vírus de vacina de DEN-2 PDK-53

Foram analisadas as passagens intermédias PDK-5, -10, -14, -35, e -45 do vírus 16681 para determinar o acréscimo das nove mutações de nucleótidos na estirpe de vacina de PDK-53. Os amplicões foram amplificados diretamente a partir do ARNm genómico extraído da semente viral por RT/PCR. Foi realizada a sequenciação automática de pequenas regiões genómicas, que continham os nove loci relevantes, utilizando os oligonucleótidos iniciadores apropriados. Os resíduos de nucleótidos identificados em cada um dos nove loci para estes vírus são apresentados na tabela 13. A mutação Leu para Phe em NS2A-181 e as mutações silenciosas em E-37, NS3-342, e NS5-334 apareceram pela passagem PDK-5 e foram as porções predominantes pela passagem PDK-10 (NS2A-181), PDK-14 (E-373, NS3-342), ou PDK-35 (NS5-334). As mutações C para T em 5'-NC-57, Asp para Val em prM-29, Gly para Asp em NS1-53, e Gly para Ala em NS4A-75 ocorreram pela passagem PDK-35. O 5'-NC-57-T foi predominante na passagem PDK-35, enquanto que as outras mutações enumeradas tornaram-se predominantes pela passagem PDK-45. A mutação Glu para Val em NS3-250 apareceu pela passagem PDK-45 e não está completamente mutada para a Val específica do vírus no candidato atual a vacina de PDK-53 (tabela 13). Aproximadamente 29 % da população viral na vacina de PDK-53 contém NS3-250-Glu. O vírus PDK-45 foi geneticamente equivalente ao vírus de vacina de PDK-53. No presente estudo, não foi feita nenhuma tentativa para determinar as proporções relativas dos dois nucleótidos nos loci genéticos mistos mostrados na tabela 13.

Exemplo 6***Construção de clones e vírus DEN-2 / Nilo Ocidental quiméricos***

Foram construídos os clones de ADNc infeccioso de DEN-2/WN quimérico, de comprimento de genoma, que contêm os genes estruturais do vírus do WN no fundo genético do vírus DEN-2, utilizando a estratégia de ligação *in vitro* utilizada para derivar os vírus DEN-2/3 quiméricos descritos anteriormente.

Num primeiro exemplo, o ADNc que codifica para prM-E do subclone da extremidade 5' que foi utilizado para derivar um clone de vírus DEN-2 / 3-PP1 quimérico (veja-se o exemplo 2) foi substituído pela região do gene prM-E do vírus do WN, estirpe NY 99 (Nova Iorque, 1999). O fragmento de ADNc que contém a região do gene prM-E do vírus do WN foi amplificado pela reação em cadeia da polimerase por transcriptase reversa (RT-PCR) a partir do ARN genómico específico do vírus do WN com o oligonucleótido iniciador direto WN-M (5'-GGAGCAGTTACCCTCTCTACGCGTCAAGGGAAGGTGATG-3' (SEQ ID NO: 37); local MluI sublinhado seguido pela sequência do vírus do WN próxima da extremidade 5' do gene prM) e o oligonucleótido iniciador inverso cWN-E (5'-GAAGAGCAGAACTCCGCCGGCTGCGAGAAACGTGAGAGCTATGG-3' (SEQ ID NO: 38); local NgoMIV sublinhado seguido pela sequência que foi complementar à sequência genómica de WN próxima da extremidade 3' do gene E do vírus do WN). O fragmento de ADNc de prM-E amplificado foi clonado nos locais MluI - NgoMIV dos clones pD2I / D3-P1-AscI intermédios, exatamente como foi feito durante a derivação do vírus DEN-2/3 quimérico (veja-se o exemplo 2, construção dos clones infecciosos de DEN-2/3 quimérico), para substituir a região prM/E do vírus DEN-3 16562. Neste ADNc de DEN-2/WN quimérico, o sítio de "splicing" de prM codifica para a sequência de aminoácidos (mostrada como abreviaturas de

aminoácidos de uma só letra, e escrita na ordem amino para carboxilo) SAGMIIMLIPTVMA-FHLT-TRQGKVMMTV (SEQ ID NO: 39). O hífen nesta sequência indica o local de clivagem da poliproteína localizado entre os genes da cápside e prM; a sequência do terminal amino é específica do vírus DEN-2; a sequência do terminal carboxilo em letra a negrito, sublinhada, itálica indica a sequência específica do vírus do WN (QGKVMMTV) (SEQ ID NO: 40), perto do terminal amino da proteína prM codificada por esta construção quimérica. Este subclone intermédio de DEN-2/WN foi ligado *in vitro* com a extremidade 3' do subclone intermédio de DEN-2, pD2-Pm^b-Asc, pelo mesmo protocolo descrito no exemplo 2 (construção de clones infecciosos de DEN-2/3 quimérico) para produzir o ADNc viral do DEN-2/WN quimérico, de genoma de comprimento completo, o qual foi utilizado para transcrever o ARN genómico específico do vírus DEN-2/WN quimérico. Utilizando o mesmo protocolo descrito no exemplo 2, as células LLC-MK₂ de mamífero ou C6/36 de mosquito foram transfectadas com o ARN transcrito de DEN-2/WN quimérico. Esta estratégia, que envolveu utilizar o mesmo local de "splicing" da enzima de restrição MluI que foi utilizado para derivar os vírus DEN-2/3 e DEN-2/4 quiméricos, não conseguiu produzir vírus quiméricos viáveis após a transfecção do ARN transcrito de DEN-2/WN quimérico em células LLC-MK₂, e apenas resultou em títulos muito baixos (< 100 ufp/ml) em células C6/36. Esta falha é, provavelmente, devida à variação significativa da sequência do gene entre as extremidades do terminal carboxilo das proteínas da cápside viral e entre as extremidades do terminal amino da proteína prM dos vírus DEN-2 e WN. A região do terminal carboxilo da proteína da cápside de flavivírus serve como uma sequência peptídica de sinal para a inserção de prM em membranas intracelulares (retículo endoplasmático) durante a maturação e a clivagem da proteína prM. Esta construção

DEN-2/WN, que continha a sequência de sinal da cápside do terminal carboxilo, assim como os resíduos do terminal amino da proteína prM da estrutura de DEN-2, aparentemente, não permitiu a maturação adequada do vírus quimérico.

Num segundo exemplo, o clone de ADNc de DEN-2/WN quimérico foi modificado de modo a codificar para a região do terminal carboxilo da proteína da cápside, assim como para toda a proteína prM e a maior parte da proteína E do vírus do WN. Foi introduzido um único local de restrição SstII por mutagénese dirigida ao local perto do terminal 3' do gene da cápside específica do vírus DEN-2 para servir como um novo local de "splicing" 5' para a região do gene da cápside do terminal carboxilo / prM/E do WN. Este local SstII introduziu duas mutações silenciosas na sequência específica do vírus DEN-2 que codifica para a tríade de aminoácidos SAG (abreviaturas de uma só letra). A região do gene adequada foi amplificada a partir do ARN viral do WN, utilizando o oligonucleótido iniciador direto WN-452.SAG (5'-AATTCAACGCGTACATCCGCGGGCACCGGAATTGCAGTCA TGATTGGCCTGATGGC-3' (SEQ ID NO: 41); local SstII sublinhado seguido pela sequência específica do vírus do WN) e o mesmo oligonucleótido iniciador reverso cWN-E que foi utilizado na construção anterior (tal como descrito anteriormente). Este ADNc amplificado foi clonado para fazer o subclone intermédio pDEN-2 / WN-P-SA, que continha o ADNc que codifica para a região não codificante 5' e para a maior parte do gene da cápside do vírus DEN-2 16681 e a cápside do terminal carboxilo, toda a prM e a maior parte do gene E do vírus do WN, bem como um local AscI único a jusante do local NgoMIV. Neste ADNc de DEN-2/WN quimérico, o local de "splicing" de prM codifica para a sequência de aminoácidos (mostrada como abreviaturas de aminoácidos de uma só letra, e escrita na ordem amino para carboxilo) SAGTGI~~AV~~MI~~GL~~IASVGA-VT~~LS~~NFOGKV~~MM~~TV (SEQ ID NO:

42). O hífen nesta sequência indica o local de clivagem da poliproteína localizado entre os genes da cápside e da prM; a sequência do terminal amino é específica do vírus DEN-2; a sequência do terminal carboxilo em letra a negrito, sublinhada, itálica indica a sequência específica do vírus do WN (TGIAVMIGLIASVGA-VTLSNFQ GKVM MT V) (SEQ ID NO: 43) codificada por esta construção quimérica. Após o protocolo de ligação *in vitro* dos clones de DEN-2/3 quimérico descrito para o vírus DEN-2/3 quimérico no exemplo 2, o subclone intermédio pDEN-2 / WN-P-SA da extremidade 5' foi ligado ao subclone pD2-Pm^b-Asc da extremidade 3' para produzir o ADNc de genoma de comprimento completo do vírus DEN-2 / WN-PP1 quimérico para a transcrição do ARN quimérico. A construção de DEN-2 / WN-PP1 quimérico codifica para a região indicada do gene estrutural de WN no fundo genético do vírus DEN-2 16681 do tipo selvagem. Ambas as células LLC-MK₂ e C6/36 foram transfetadas com ARN transcrito a partir deste clone DEN-2 / WN-PP1 quimérico, ligado *in vitro*. O vírus DEN-2 / WN-PP1 quimérico, viável, em títulos de vírus de 10³ - 10⁶ ufp/ml de meio de cultura, foi recuperado com sucesso a partir de ambas as culturas de células transfetadas. A análise da sequência de nucleótidos de todo o genoma do vírus DEN-2/WN quimérico, derivado de clone, demonstrou a sequência genómica esperada. Estes resultados demonstram que os clones infecciosos de DEN-2 aqui descritos podem ser utilizados para construir vírus quiméricos que expressam os genes estruturais de flavivírus heterólogos para além dos vírus DEN-1, DEN-3, e DEN-4.

Tabela 1

Candidatos virais para quimeras⁺ de flavivírus

<i>Espécies primárias transmitidas por mosquitos</i>	<i>Outras espécies transmitidas por mosquitos</i>
Dengue-1 **	Alfuy Jugra

<i>Espécies primárias transmitidas por mosquitos</i>	<i>Outras espécies transmitidas por mosquitos</i>	
Dengue-3 ** *	Bagaza	Kokobera
Dengue-4 **	Banzi **	Ntaya
Febre amarela **	Bouboui	Rocio **
Encefalite japonesa **	Bussuquara **	Sepik **
Encefalite de Murray Valley **	Edge Hill **	Spondweni **
Encefalite de St. Louis **	Ilheus	Stratford
Nilo Ocidental **	Naranjal	Tembusu
Kunjin *	Meningite de Israel Turquia	Uganda S
		Usutu **
		Wesselsbron **
		Zika **

<i>Transmitido por carrapatos</i>	<i>Sem vetor artrópode demonstrado</i>
Absettarov **	Apoi *
Gadgets Gully	Aroa
Hanzalova **	Cacipacore
Hypr **	Ilha de Carey
Kadam	Cowbone Ridge
Karshi	Dakarbat **
Kumlinge **	morcego de Entebbe
Doença da floresta Kyasanur **	Jutiapa
Langat	Koutango *
Doente Louping **	Modoc *
Meaban	Leucemia de Montana
	Myotis
Febre hemorrágica de Omsk **	Negishi **
Powassan **	morcego de Phnom-Penh
Royal Farm	Rio Bravo **
Primavera Russa	Saboya
encefalite de verão **	Sal Vieja
Saumarez Reef	San Perlita
Tyuleniy	Sokuluk

**Transmitido por
carrapatos****Sem vetor artrópode demonstrado**

* = Infecção relatada em laboratório

** = Infecção natural e relatada em laboratório

+ = A lista inclui os membros atualmente classificados do género Flavivírus no International Catalogue of Arboviruses, Including Certain Other Viruses of Vertebrates, Nick Karabatsos, ed., The American Society of Tropical Medicine and Hygiene. 1985.

Tabela 2

**Resumo das diferenças de nucleótidos e de aminoácidos
entre os genomas do vírus DEN-1 16007 e a sua estirpe PDK-
13 derivada de vacina**

Posição do nucleótido no genoma	Nucleótido		Aminoácido		Posição da proteína	Posição da poliproteína
	16007	PDK- 13	16007	PDK- 13		
1323	T	C	Val	Ala	E-130	410
1541)	G	A	Glu	Lys	E-203	483
1543)	A	G				
1545	G	A	Arg	Lys	E-204	484
1567	A	G	Gln	Gln	E-211	491
1608	C	T	Ser	Leu	E-225	505
2363	A	G	Met	Val	E-477	757
2695	T	C	Asp	Asp	NS1-92	867
2782	C	T	Ala	Ala	NS1-121	896
5063	G	A	Glu	Lys	NS3-182	1657
6048	A	T	Tyr	Phe	NS3-510	1985
6806	A	G	Met	Val	NS4A-144	2238
7330	A	G	Gln	Gln	NS4B-168	2412
9445	C	T	Ser	Ser	NS5-624	3117

Tabela 3

Resumo das diferenças de nucleótidos e de aminoácidos entre os genomas do vírus DEN-2 16681 e a sua estirpe PDK-13 derivada de vacina

Posição do nucleótido no genoma	Nucleótido		Aminoácido		Posição da proteína	Posição da poliproteína
	16681	PDK-53	16681	PDK-53		
57	C	T ^a	--			
524	A	T	Asp	Val	prM-29	
2055	C	T	Phe	Phe	E-373	653
2579	G	A	Gly	Asp	NS1-53	828
4018	C	T	Leu	Phe	NS2A-181	1308
5270	A	T / A	Glu	Glu / Val ^b	NS3-250	1725
5547	T	C	Arg	Arg	NS3-342	1817
6599	G	C	Gly	Ala	NS4A-75	2168
8571	C	T	Val	Val	NS5-334	2825

^[a] região 5' não codificante.

^[b] A vacina de PDK-53 contém duas variantes genéticas no nt-5270.

Tabela 4

Resumo das diferenças de nucleótidos e de aminoácidos entre os genomas do vírus DEN-3 16562 e a sua estirpe PGMK-30 / Frhl-3 derivada de vacina

Posição do nucleótido no genoma	Nucleótido		Aminoácido		Posição da proteína	Posição da Poliproteína
	16562	PGMK-30 / FRhL-3	16562	PGMK-30 / FRhL-3		
550	C	T	Ala	Ala	prM-38	152
1521	C / T	C	Ser / Leu ^a	Ser	E-196	476
1813	G	A	Lys	Lys	E-293	573
1838	A	G	Ser	Gly	E-302	582
1913	G	A	Glu	Lys	E-327	617
2140	C	T	Ala	Ala	E-402	682
3725	T	C	Phe	Leu	NS2A-86	1211
4781	C	A	Gln	Lys	NS3-90	1563

^a Foram localizadas duas variantes genéticas significativas no nt-1521.

Tabela 5

Resumo das diferenças de nucleótidos e de aminoácidos entre os genomas do vírus DEN-4 1036 e a sua estirpe PDK-48 derivada de vacina

Posição do nucleótido no genoma	Nucleótido		Aminoácido		Posição da proteína	Posição da poliproteína
	1036	PDK-48	1036	PDK-48		
1211	T	C	Ile	Ile	E-91	370
1971	G	A	Glu	Lys	E-345	624
3182	G	C	Gln	His	NS1-253	1027
6660	C	T	Leu	Phe	NS4A-95	2187
6957	A	A / T	Ile	Ile / Phe	NS4B-44	2286
7162	T	C	Leu	Ser	NS4B-112	2354
7546	C	C / T	Ala	Ala / Val	NS4B-240	2366
7623	G	T / G	Asp	Tyr / Asp	NS5-21	2508

Tabela 6

Resumo das mutações não silenciosas entre os genomas das estirpes de vacina parental dos vírus DEN-1, DEN-2, DEN-3 e DEN-4

Posição do nucleótido no genoma	DEN-1	DEN-2	DEN-3	DEN-4
57		5'NC-57 c-t		
524		prM-29 D-V		
1323	E-130 V-A			
1521			E-196 S / L-S	
1541)	E-203 E-K			
1543)				
1545	E-204 R-K			
1608	E-225 S-L			
1838			E-302 S-G	
1913			E-327 E-K	
1971				E-345 E-K
2363	E-477 M-V			

Posição do nucleótido no genoma	DEN-1	DEN-2	DEN-3	DEN-4
2579		NS1-53 G-D		
3182				NS1-253 Q-H
3725			NS2A-86 F- L	
4018		NS2A-181 L-F		
4781			NS3-90 Q-K	
5063	NS3-182 E-K			
5270		NS3-250 E-V / E		
6048	NS3-510 Y-F			
6599		NS4A-75 G-A		
6660				NS4A-95 L-F
6806	NS4A-144 M-V			
6957				NS4B-44 I-I / F
7162				NS4B-112 L-S
7546				NS4B-240 A-A / V
7623				NS5-21 D-D / Y

Tabela 7

Imunogenicidade de vírus em ratinhos. Título^a de redução de placa - neutralização contra o vírus DEN-1 16007

	Experiência 1 ^b		Experiência 2 ^b	
	Primário	Reforço	Primário	Reforço
Conjunto de imunização	Vírus dos soros (gama)	Conjunto de soros (gama)	Conjunto de soros (gama)	Conjunto de soros (gama)
DEN-1 16007	80 (20 - 80)	2560 (80 - 20480)	80 (20 - 160)	2560 (160 - 5120)
D2 / 1-PP	80 °	5120 °	40 (10 - 160)	5120 (160 - 10240)
D2 / 1-EP	80 (20 - 320)	10240 (640 - 20480)	160 (20 - 320)	5120 (2560 - ≥ 10240)
D2 / 1-VP	40 (10 - 160)	2560 (160 - 5120)	80 (10 - 320)	5120 (40 - ≥ 10240)

	Experiência 1 ^b		Experiência 2 ^b	
	Primário	Reforço	Primário	Reforço
Conjunto de imunização	Vírus dos soros (gama)	Conjunto de soros (gama)	Conjunto de soros (gama)	Conjunto de soros (gama)
DEN-1 PDK-13	10 (10 - 40)	80 ^c	40 (20 - 80)	320 (20 - 640)
D2 / 1- PV	10 (< 10 ^d - 20)	80 ^c	40 (10 - 80)	2560 (20 - ≥ 10240)
D2 / 1-EV	20 (< 10 ^d - 20)	80 ^c	40 (< 10 ^d - 40)	160 (10 - 320)
D2 / 1-VV	10 (10 - 40)	80 ^c	40 (10 - 160)	160 (20 - 640)

^a Os títulos são a diluição recíproca produzindo pelo menos 50 % de redução de placa.

^b Ratinhos ICR não consanguíneos de 3 semanas foram imunizados por via intraperitoneal com 10⁴ ufp do vírus, e foram reforçados com a mesma dose de vírus 3 semanas mais tarde na experiência 1, ou 6 semanas mais tarde na experiência 2. Primário = soro recolhido 20 dias (experiência 1) ou 41 dias (experiência 2) após a imunização primária. Reforço = soro recolhido 21 dias após o reforço em ambas as experiências.

^c Não foram determinados os títulos individuais.

^d Apenas um título de soro de ratinho foi inferior a 10 nestes grupos.

^e Os títulos em negrito indicam que os títulos dos conjuntos de soros foram 4 vezes mais elevados do que os títulos calculados com 70 % de redução de placa. Todos os outros conjuntos de títulos eram ou não diferentes ou duas vezes superiores aos títulos de 70 % de redução de placa.

Tabela 8
Neutralização de vírus DEN quiméricos com anticorpos convencionais

Vírus	Anticorpo					
	D1-AF ^a	D2-AF ^a	D2-H5 ^b	D3-AF ^a	D3-8A1 ^b	D4-AF ^a
DEN-1 16007 ^c	1280 ^d	40	20	40	< 20	20
DEN-2/1-EP ^e	2560	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
DEN-216681 ^c	80	2560 >	40960	40	< 20	40
DEN-316562 ^c	80	160	40	1280	5120	40
DEN-2/3-EP1 ^e	<i>i</i>	<i>i</i>	<i>i</i>	2560	20480	<i>i</i>
DEN-4 1036 ^c	20	40	< 20	80	< 20	1280
DEN-2/4-EP1 ^e	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>	2560

Vírus	Anticorpo					
	D1-AF ^a	D2-AF ^a	D2-H5 ^b	D3-AF ^a	D3-8A1 ^b	D4-AF ^a
^a AF = fluidos ascíticos de vírus específico de ratinho de DEN-1 (D1), DEN-2 (D2), DEN-3 (D3) e DEN-4 (D4) (antissoros polivalentes).						
^b Anticorpos monoclonais D2-H5 e D3-8A1, específico para os vírus DEN-2 e DEN-3, específico para a glicoproteína de envólucro (específico-E), respectivamente.						
^c Vírus DEN do tipo selvagem.						
^d São mostrados os títulos recíprocos de neutralização de diluição de soro - redução de placa (parâmetro de 50 %).						
; = Não testado.						
^e Vírus DEN-2/1-EP, DEN-2/3-EP1, ou DEN-2/4-EP1 quimérico expressando os genes estruturais dos vírus DEN-1 16007, DEN-3 16562, ou DEN-4 1036, respectivamente.						

Tabela 9
Neurovirulência dos vírus DEN-2/3 e DEN-2/4 quiméricos em
ratinhos ICR brancos recém-nascidos

Vírus de desafio ^a	Percentagem de mortalidade ^b	AST (SD) (dias) ^c
DEN-3 16562 ^d	100,0	14,1 (2,1)
DEN-3 P30 / FRh1-3 ^e	15,0	16,2 (3,2)
DEN-2/3-PP1 ^f	32,5	19,0 (2,1)
DEN-2/3-EP1 ^f	0,0	—
DEN-2/3-VP1 ^f	0,0	—
DEN-41036 ^d	100,0	8,6 (0,6)
DEN-4 PDK-48 ^e	100,0	10,7 (1,5)
DEN-2/4-PP1 ^f	62,5	17,8 (2,8)
DEN-2/4-EP1 ^f	0,0	—
DEN-2/4-VP1 ^f	0,0	—
DEN-2 16681 ^d	87,5	15,2 (2,6)
DEN-2 PDK53-E48 ^e	0,0	—
DEN-2 PDK53-V48 ^e	0,0	—
Diluyente	0,0	—

o vírus DEN-2/3 ou DEN-2/4 quimérico. A região do gene prM/E do vírus DEN-3 16562 ou do vírus DEN-4 1036 foi expressa no fundo genético do vírus DEN-2 16681 do tipo selvagem (vírus PP1), na variante DEN-2 PDK-53-E candidata a vacina da Mahidol (vírus EP1), ou na variante DEN-2 PDK-53-V candidata a vacina da Mahidol (vírus VP1).

Eficácia protetora do vírus DEN-2/1 quimérico em ratinhos
AG-129 ^a

Vírus de imunização	4 semanas	6 semanas
DEN-3 16562 ^b	320 ^c	640
DEN-3 P30 / FRhL-3 ^d	160	320
DEN-2/3-EP1 ^e	80	320
DEN-4 1036 ^b	160	1280
DEN-4 PDK-48 ^d	80	320
DEN-2/4-EP1 ^e	20	40

Título recíproco de anticorpo neutralizante contra o vírus DEN-3 ou DEN-4 homólogo apropriado a 4 ou 6 semanas após a imunização

Vírus de imunização 4 semanas 6 semanas

^a Os ratinhos AG-129 são uma estirpe consanguínea que não possui recetores para o interferão alfa / beta e o interferão gama. Os ratinhos de 3,5 - 4,5 semanas de idade foram imunizados por via intraperitoneal com 10^5 ufp de vírus.

^b Vírus DEN do tipo selvagem.

^c Diluição recíproca de conjuntos de soro que neutralizou 70 % ou mais do vírus de entrada DEN-3 16562 do tipo selvagem que foi utilizado para testar os soros de ratinhos imunizados com o vírus DEN-3 ou com o vírus DEN-2/3-EP1 quimérico, ou com o vírus de entrada DEN-4 1036 do tipo selvagem que foi utilizado para testar os soros de ratinhos imunizados com o vírus DEN-4 ou com o vírus DEN-2/4-EP1 quimérico.

^d Vírus candidato a vacina da Mahidol, DEN-3 PGMK-30 / FRhL-3 (P30 / FRhL-3), DEN-4 PD-48.

^e DEN-2/3-EP1 ou DEN-2/4-EP1 quimérico expressando a região do gene prM/E dos vírus DEN-3 16562 ou DEN-4 1036, respetivamente.

Tabela 11

Genótipos de vírus DEN-2 16681 / PDK-53 recombinantes

Vírus derivado de clone (cognato) ^b	5'NC-57	Determinantes de dengue-2 fundo 16681 em PDK-53 ^a				
		prM-29	NS1-53	NS2A-181	NS3-250	NS4A-75
DEN-2 PDK-53 t		V	D	F	V	A
W45R (V513)	.	.	.	.	.	.
P5 (V13)	c	.	.	.	.	.
P1 (V53)	.	.	G	.	.	.
P3 (V51)	.	.	.	.	E	.
P51 (V3)	c	.	G	.	.	.
P53 (V1)	c	.	.	.	E	.
P13 (V5)	.	.	G	.	E	.
P513 (30P-A)	c	.	G	.	E	.

Determinantes de dengue-2 PDK-53 em fundo 16681 ^c						
	5'NC-57	prM-29	NS1-53	NS2A-181	NS3-250	NS4A-75
DEN-2 16681	c	D	G	L	E	G

	Determinantes de dengue-2 PDK-53 em fundo 16681 °					
	5'NC- 57	prM-29	NS1-53	NS2A-181	NS3-250	NS4A- 75
30P-A (P513)	.	.	.	.	.	.
V5 (P13)	t	.	.	.	.	.
V1 (P53)	.	.	D	.	.	.
V3 (P51)	.	.	.	.	V	.
V51 (P3)	t	.	D	.	.	.
V53 (P1)	t	.	.	.	V	.
V13 (P5)	.	.	D	.	V	.
V513 (VV45R)	t	.	D	.	V	.

Determinantes de dengue-2 PDK-53 em fundo 16681 ^c					
5'NC- 57	prM-29	NS1-53	NS2A-181	NS3-250	NS4A- 75
^a O genoma do vírus PDK-53 do dengue-2 candidato a vacina difere daquele do seu parental 16681 em nove loci de nucleótidos, incluindo três mutações silenciosas (não mostrado), uma mutação na posição do nucleótido 57 do genoma na região não codificante 5' (5'NC-57; letras minúsculas), e cinco nucleótidos que codificam para as mutações de aminoácidos (maiúsculas, abreviaturas de uma letra) nas posições de polipeptído viral da pré-membrana (prM)-29, proteína não estrutural 1 (NS1)-53, NS2A-181, NS3-250, e NS4A-75 (32). Os loci genéticos do vírus parental 16681 foram manipulados no fundo de ADNc do clone infeccioso específico do vírus PDK-53, pD2 / IC-W45R. Os pontos indicam a identidade de sequência com o vírus PDK-53 (variante NS3-250-Val). O candidato PDK-53 a vacina também contém uma variante genética que tem Glu em NS3-250 (32).					
^b São mostrados os genótipos do vírus DEN-2 16681 do tipo selvagem, do seu derivado de vacina atenuado, do vírus DEN-2 PDK-53, do vírus W45R derivado de clone infeccioso (equivalente geneticamente à variante PDK-53 NS3-250-Val) e do vírus 30P-A (equivalente ao 16681 do tipo selvagem), e do vírus 16681 / PDK-53 recombinante. As designações numéricas para os vírus recombinantes Px e Vx (em que x = loci 5'NC, NS1 e/ou NS3) indicam os loci específicos do vírus 16681 parental (P na designação de vírus) concebido no clone de ADNc infeccioso específico do vírus PDK-53 (série do topo) ou nos loci recíprocos específicos do vírus PDK-53 candidato a vacina (V na designação de vírus) concebido no clone 16681 (série da base), respetivamente. P5 e V13 são vírus cognatos, assumindo que o fenótipo específico do vírus PDK-53 é determinado predominantemente pelos loci 5'NC-57, NS1-53 e NS3-250. Ambos os vírus P5 e V13 contêm os loci 5'NC-57-c, NS1-53-Asp e NS3-250-Val dentro dos fundos genéticos dos vírus PDK-53 e 16681, respetivamente.					
^c Os loci genéticos do vírus PDK-53 foram concebidos no fundo de ADNc do clone infeccioso específico do vírus 16681, pD2 / IC-30P-A. Os pontos indicam a identidade de sequência com o vírus 16681.					

Tabela 12

**Neurovirulência dos vírus DEN-216681, PDK-53, e 16681 /
PDK-53 recombinante em ratinhos ICR brancos recém-nascidos**

Desafio em ratinho ^a			Genótipo do vírus ^b					
Vírus (cognato) ^c	Mortalidade (%)	AST (SD) (dias)	5'- NC 57	prM 29	NS1 53	NS2A 181	NS3 250	NS4A 75
DEN-2 16681	68,75	15,2 (1,2)	c	D	.	L	.	G
30P-A (P513)	81,25	14,6 (2,3)	.	.	.	.	.	.
P513 (16681)	100,0	13,2 (1,6)	.	V	.	F	.	A
P51 (V3)	50,0 ^d	15,9 (5,5)	.	V	.	F	V	A
P1 (V53)	37,5 ^d	19,0 (4,2)	t	V	.	F	V	A
P13 (V5)	37,5 ^d	13,5 (2,1)	t	V	.	F	.	A
P53 (V1)	20,0 ^d	17,0 (7,8)	.	V	D	F	.	A
P5p (V13)	6,25	15,0	.	.	D	F	V	A
P5 (V13)	6,25	27,0	.	V	D	F	V	A
Pp (PDK- 53)	0	- ^c	t	.	D	F	V	A
P3 (V51)	0	-	t	V	D	F	.	A
V3 (P51)	7,50 ^d	16,4 (3,2)	.	.	.	.	V	.
Vp (16681)	87,5	17,0 (0,9)	.	V	.	.	.	.
V53 (P1)	18,75 ^d	21,3 (6,1)	t	.	.	.	V	.
V5 (P13)	18,75 ^d	21,7 (4,2)	t	.	.	.	.	.
V5p (P13)	6,25	20,0	t	V	.	.	.	.
V13 (P5)	6,25	17,0	.	.	D	.	V	.
V1 (P53)	6,25	22,0	.	.	D	.	.	.
V51 (P3)	0	-	t	.	D	.	.	.
V513 (PDK-53)	0	-	t	.	D	.	V	.
VV45R (V513)	0	-	t	V	D	F	V	A
DEN-2 PDK-53	0	-	t	V	D	F	V	A

Desafio em ratinho ^a			Genótipo do vírus ^b					
Vírus (cognato) ^c	Mortalidade (%)	AST (SD) (dias)	5'- NC 57	prM 29	NS1 53	NS2A 181	NS3 250	NS4A 75

^a Percentagem de mortalidade e tempo médio de sobrevivência (AST) $\pm$ desvio padrão (SD) de ratinhos ICR brancos, não consanguíneos, recém-nascidos, desafiados por via intracranial com 10^4 ufp de vírus. Dezasseeis ratinhos por grupo, exceto para os grupos de p53 e Vp nos quais morreu um único ratinho no dia 1 após a infecção, provavelmente como resultado de trauma da inoculação. Estes dois ratinhos foram excluídos do estudo.

^b Veja-se o texto para a explicação dos genótipos do vírus. Os pontos sólidos indicam a identidade da sequência com o vírus 16681.

^c Veja-se o texto para a explicação das designações do vírus e do cognato do vírus.

^d O peso corporal médio dos ratinhos que sobreviveram foi significativamente menor ($p < 0,001$, test t de Student) do que o dos ratinhos controle inoculados com o diluente (não mostrados) às 3 semanas após a infecção.

^e O tempo médio de sobrevivência não é aplicável, porque não houve mortalidade neste grupo de ratinho.

Tabela 13

Evolução do vírus DEN-2, estirpe de vacina PDK-53, durante a passagem da estirpe parental 16681 em células primárias de rim de cão (PDK)

Vírus	Posição do nucleótido no genoma ^a /				Posição do polipeptido traduzido ^b /			Aminoácidos codificados ^e		
	57 ^a	524 prM- 29 ^b	2055 D E-373 - V ^c	2579 NS1- 53	4018 NS2A-181 - F	5270 L NS3- 250	5547 NS3- 342	6599 NS4A- 75	8571 NS5- 334	
				G - D		E - V	R - R	G - A	V - V	
16681	C ^d	A	C	G	C	A	T	G	C	
PDK-5	C	A	C / T ^c	G	T / C	A	T / C	G	C / T	
PDK-10	C	A	T / C	G	T	A	C / T	G	T / C	
PDK-14	C	A	T	G	T	A	C	G	T / C	
PDK-35	T	T / A	T	A / G	T	A	C	C / G	T	

	Posição do nucleótido no genoma ^a /				Posição do polipeptido traduzido ^b /			Aminoácidos codificados ^e		
	57 ^a	524 prM-	2055	2579	4018		5270	5547	6599	8571
		29 ^b D	E-373	NS1-	NS2A-181	L	NS3-	NS3-	NS4A-	NS5-
		- V ^c	F - F	53	- F		250	342	75	334
Vírus				G - D			E - V	R - R	G - A	V - V
PDK-45	T	T	T	A	T		A / T	C	C	T
PDK-53	T	T	T	A	T		A / T	C	C	T

^a Posições de nucleótidos no genoma das nove diferenças de sequência de nucleótidos entre os vírus 16681 e PDK-53. A posição do nucleótido 57 encontra-se dentro da região não codificante 5' do genoma viral.

^b As designações das proteínas são tal como se segue: prM = proteína de pré-membrana; E = glicoproteína do envoltório; NS = proteína não estrutural.

^c Os resíduos de aminoácidos específicos do vírus (16681 - PDK-53) são mostrados para cada posição de aminoácido.

^d São indicados os nucleótidos A, C, G e T (sentido direto de ADNc).

^e Foram identificadas duas populações genéticas para este locus do vírus. A ordem dos dois nucleótidos reflete as alturas relativas dos picos dos sinais de nucleótidos em cromatogramas de sequência.

Tabela 14

Conservação dos marcadores de atenuação fenotípica de DEN-2 PDK-53 em vírus DEN quiméricos.

Fenótipo de atenuação	Vírus DEN			
	DEN-2 ^a	DEN-2/1	DEN-2/3	DEN-2/4
Atenuação em ratinhos	+	+	+	+
Diminuição da replicação em células C6/36	+	+	+	+
Placas pequenas em células LLC-MK ₂	+	+	+	+
Sensibilidade à temperatura em células LLC-MK ₂	+	+	+	+

^a Vírus DEN-2 PDK-53 candidato a vacina da Mahidol. Os vírus DEN-2/1, DEN-2/3 e DEN-2/4 quiméricos foram construídos no fundo genético de DEN-2 PDK-53, variante NS3-250-Val ou -Glu. Os fenótipos mostrados para os vírus quiméricos são representativos de vírus quiméricos construídos no fundo de qualquer uma das variantes do vírus DEN-2 PDK-53.

LISTAGEM DE SEQUÊNCIAS

<110> O Governo dos Estados Unidos da América, tal como representado pelo Secretário, Departamento de Saúde e Serviços Humanos; Universidade de Mahidol

<120> QUIMERAS DE FLAVIVÍRUS IMUNOGÉNICAS, NÃO VIRULENTAS

<130> SMW/FP6720908

<140> EP

<141> 16-02-2001

<150> EP 01910882.8

<151> 16-02-2001

<150> PCT/US01/05142

<151> 16-02-2001

<150> 60/182.829

<151> 16-02-2000

<160> 148

<170> FastSEQ para Windows Versão 4.0

<210> SEQ ID NO 5

<211> COMPRIMENTO: 10723

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10272)

<400> SEQUÊNCIA 5

```

agttgttagt ctacgtggac cgacaaagac agattccttg agggagctaa gctcaatgta 60
gttctaacag tttttaatt agagagcaga tctctg atg atc aac caa cga aaa 114
Met Ile Asn Gln Arg Lys
1
aag acg ggt cga ccg tct ttc aat atg ctg aaa cgc gcg aga aac cgc 162
Lys Thr Gly Arg Pro Ser Phe Asn Met Leu Lys Arg Ala Arg Asn Arg
10
gtg tca act gtt tca cag ttg gcg aag aga ttc tca aaa gga ttg ctc 210
Val Ser Thr Val Ser Gln Leu Ala Lys Arg Phe Ser Lys Gly Leu Leu
25
tca ggc caa gga ccc atg aaa ttg gtg atg gct ttc ata gca ttc tta 258
Ser Gly Gln Gly Pro Met Lys Leu Val Met Ala Phe Ile Ala Phe Leu
40
aga ttt cta gcc ata ccc cca aca gca gga att ttg gcc aga tgg ggc 306
Arg Phe Leu Ala Ile Pro Pro Thr Ala Gly Ile Leu Ala Arg Trp Gly
55
tca ttc aag aag aat gga gcg att aaa gtg tta cgg ggt ttc aag aga 354
Ser Phe Lys Lys Asn Gly Ala Ile Lys Val Leu Arg Gly Phe Lys Arg
75
gaa atc tca aac atg cta aac ata atg aac agg agg aaa aga tcc gtg 402
Glu Ile Ser Asn Met Leu Asn Ile Met Asn Arg Arg Lys Arg Ser Val
90
acc atg ctc ctt atg ctg ctg ccc aca gcc ctg gcg ttc cat ctg acg 450
Thr Met Leu Leu Met Leu Leu Thr Ala Leu Ala Phe His Leu Thr
105
aca cga ggg gga gag ccg cat atg ata gtt agc aag cag gaa aga gga 498
Thr Arg Gly Gly Glu Pro His Met Ile Val Ser Lys Gln Glu Arg Gly
120
aag tca ctt ttg ttc aag acc tct gca ggt gtc aac atg tgc acc ctc 546
Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly Val Asn Met Cys Thr Leu
135
att gcg atg gat ttg gga gag ttg tgt gag gac acg atg acc tac aaa 594
Ile Ala Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Met Thr Tyr Lys
155
tgc ccc cgg atc act gag gcg gaa cca gat gac gtt gac tgt tgg tgc 642
Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp Asp Val Asp Cys Trp Cys
170
aat gcc acg gac aca tgg gtg acc tat gga acg tgc tct caa act ggc 690
Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly Thr Cys Ser Gln Thr Gly
185
gaa cac cga cga gac aaa cgt tcc gtc gca ttg gcc cca cac gtg ggg 738
Glu His Arg Arg Asp Lys Arg Ser Val Ala Leu Ala Pro His Val Gly
200
ctt ggc cta gaa aca aga gcc gaa acg tgg atg tcc tct gaa ggt gct 786
Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp Met Ser Ser Glu Gly Ala
215
tgg aaa cag ata caa aaa gta gag act tgg gct ctg aga cat cca gga 834
Trp Lys Gln Ile Gln Lys Val Glu Thr Trp Ala Leu Arg His Pro Gly
235
ttc acg gtg ata gcc ctt ttt cta gca cat gcc ata gga aca tcc atc 882
Phe Thr Val Ile Ala Leu Phe Leu Ala His Ala Ile Gly Thr Ser Ile

```

acc	cag	aaa	ggg	atc	att	ttc	att	ttg	ctg	atg	ctg	gta	aca	cca	tct	930
Thr	Gln	Lys	Gly	Ile	Ile	Phe	Ile	Leu	Leu	Met	Leu	Val	Thr	Pro	Ser	
atg	gcc	atg	cga	tgc	gtg	gga	ata	ggc	aac	aga	gac	ttc	gtg	gaa	gga	978
Met	Ala	Met	Arg	Cys	Val	Gly	Ile	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	
ctg	tca	gga	gca	aca	tgg	gtg	gat	gtg	gta	ctg	gag	cat	gga	agt	tgc	1026
Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val	Leu	Glu	His	Gly	Ser	Cys	
gtc	acc	acc	atg	gca	aaa	aac	aaa	cca	aca	ctg	gac	att	gaa	ctc	ttg	1074
Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Ile	Glu	Leu	Leu	
aag	acg	gag	gtc	aca	aac	cct	gca	gtt	ctg	cgt	aaa	ttg	tgc	att	gaa	1122
Lys	Thr	Glu	Val	Thr	Asn	Pro	Ala	Val	Leu	Arg	Lys	Leu	Cys	Ile	Glu	
gct	aaa	ata	tca	aac	acc	acc	acc	gat	tcg	aga	tgt	cca	aca	caa	gga	1170
Ala	Lys	Ile	Ser	Asn	Thr	Thr	Thr	Asp	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
gaa	gcc	aca	ctg	gtg	gaa	gaa	caa	gac	gcg	aac	ttt	gtg	tgc	cga	cga	1218
Glu	Ala	Thr	Leu	Val	Glu	Glu	Gln	Asp	Ala	Asn	Phe	Val	Cys	Arg	Arg	
acg	ttc	gtg	gac	aga	ggc	tgg	ggc	aat	ggc	tgt	ggg	cta	ttc	gga	aaa	1266
Thr	Phe	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
ggt	agt	cta	ata	acg	tgt	gcc	aag	ttt	aag	tgt	gtg	aca	aaa	cta	gaa	1314
Gly	Ser	Leu	Ile	Thr	Cys	Ala	Lys	Phe	Lys	Cys	Val	Thr	Lys	Leu	Glu	
gga	aag	ata	gtt	caa	tat	gaa	aac	cta	aaa	tat	tca	gtg	ata	gtc	acc	1362
Gly	Lys	Ile	Val	Gln	Tyr	Glu	Asn	Leu	Lys	Tyr	Ser	Val	Ile	Val	Thr	
gtc	cac	act	gga	gat	cag	cac	cag	gtg	gga	aat	gag	act	aca	gaa	cat	1410
Val	His	Thr	Gly	Asp	Gln	His	Gln	Val	Gly	Asn	Glu	Thr	Thr	Glu	His	
gga	aca	act	gca	acc	ata	aca	cct	caa	gct	cct	acg	tcg	gaa	ata	cag	1458
Gly	Thr	Thr	Ala	Thr	Ile	Thr	Pro	Gln	Ala	Pro	Thr	Ser	Glu	Ile	Gln	
ctg	acc	gac	tac	gga	acc	ctt	aca	tta	gat	tgt	tca	cct	agg	aca	ggg	1506
Leu	Thr	Asp	Tyr	Gly	Thr	Leu	Thr	Leu	Asp	Cys	Ser	Pro	Arg	Thr	Gly	
cta	gat	ttt	aac	gag	atg	gtg	ttg	ctg	aca	atg	aaa	gaa	aga	tca	tgg	1554
Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Thr	Met	Lys	Glu	Arg	Ser	Trp	
ctt	gtc	cac	aaa	caa	tgg	ttc	cta	gac	tta	cca	ctg	cct	tgg	acc	tct	1602
Leu	Val	His	Lys	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Trp	Thr	Ser	
ggg	gct	tca	aca	tcc	caa	gag	act	tgg	aac	aga	caa	gat	tta	ctg	gtc	1650
Gly	Ala	Ser	Thr	Ser	Gln	Glu	Thr	Trp	Asn	Arg	Gln	Asp	Leu	Leu	Val	
aca	ttt	aag	aca	gct	cat	gca	aag	aag	cag	gaa	gta	gtc	gta	cta	gga	1698
Thr	Phe	Lys	Thr	Ala	His	Ala	Lys	Lys	Gln	Glu	Val	Val	Val	Leu	Gly	
tca	caa	gaa	gga	gca	atg	cac	act	gcg	ctg	act	gga	gcg	aca	gaa	atc	1746
Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile	
caa	acg	tca	gga	acg	aca	aca	att	ttc	gca	gga	cac	cta	aaa	tgc	aga	1794
Gln	Thr	Ser	Gly	Thr	Thr	Thr	Ile	Phe	Ala	Gly	His	Leu	Lys	Cys	Arg	
cta	aaa	atg	gac	aaa	cta	act	tta	aaa	ggg	atg	tca	tat	gtg	atg	tgc	1842
Leu	Lys	Met	Asp	Lys	Leu	Thr	Leu	Lys	Gly	Met	Ser	Tyr	Val	Met	Cys	
aca	ggc	tca	ttc	aag	tta	gag	aaa	gaa	gtg	gct	gag	acc	cag	cat	gga	1890
Thr	Gly	Ser	Phe	Lys	Leu	Glu	Lys	Glu	Val	Ala	Glu	Thr	Gln	His	Gly	
act	gtt	ctg	gtg	cag	gtt	aaa	tat	gaa	gga	aca	gac	gca	cca	tgc	aag	1938
Thr	Val	Leu	Val	Gln	Val	Lys	Tyr	Glu	Gly	Thr	Asp	Ala	Pro	Cys	Lys	
att	ccc	ttt	tcg	acc	caa	gat	gag	aaa	gga	gca	acc	cag	aat	ggg	aga	1986
Ile	Pro	Phe	Ser	Thr	Gln	Asp	Glu	Lys	Gly	Ala	Thr	Gln	Asn	Gly	Arg	

615	tta	ata	aca	gcc	aac	620	ccc	ata	gtc	act	gac	625	aaa	gaa	aaa	cca	gtc	630		
Leu	Ile	Thr	Ala	Asn	Pro	Ile	Val	Thr	Asp	Lys	Glu	Lys	Pro	Val	Asn	2034				
				635																
att	gag	gca	gaa	cca	ccc	ttt	ggt	gag	agc	tac	atc	gtg	gta	gga	gca	2082				
Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Glu	Ser	Tyr	Ile	Val	Val	Gly	Ala					
			650					655					660							
ggt	gaa	aaa	gct	ttg	aaa	cta	agc	tgg	ttc	aag	aaa	gga	agc	agc	ata	2130				
Gly	Glu	Lys	Ala	Leu	Lys	Leu	Ser	Trp	Phe	Lys	Lys	Gly	Ser	Ser	Ile					
		665						670				675								
ggg	aaa	atg	ttt	gaa	gca	act	gcc	cga	gga	gca	cga	agg	atg	gcc	att	2178				
Gly	Lys	Met	Phe	Glu	Ala	Thr	Ala	Arg	Gly	Ala	Arg	Arg	Met	Ala	Ile					
	680					685					690									
ctg	gga	gac	acc	gca	tgg	gac	ttc	ggt	tct	ata	gga	gga	gtg	ttc	acg	2226				
Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Ile	Gly	Gly	Val	Phe	Thr					
	695				700					705					710					
tct	atg	gga	aaa	ctg	gta	cac	cag	ggt	ttt	gga	act	gca	tat	gga	ggt	2274				
Ser	Met	Gly	Lys	Leu	Val	His	Gln	Val	Phe	Gly	Thr	Ala	Tyr	Gly	Val					
				715				720						725						
ttg	ttt	agc	gga	ggt	tct	tgg	acc	atg	aaa	ata	gga	ata	ggg	att	ctg	2322				
Leu	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	Ile	Gly	Ile	Gly	Ile	Leu					
			730					735					740							
ctg	aca	tgg	cta	gga	tta	aat	tca	agg	aac	acg	tcc	ctt	tcg	gtg	atg	2370				
Leu	Thr	Trp	Leu	Gly	Leu	Asn	Ser	Arg	Asn	Thr	Ser	Leu	Ser	Val	Met					
		745					750					755								
tgc	atc	gca	gcc	ggc	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	2418				
Cys	Ile	Ala	Ala	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln					
	760				765					770										
gcc	gat	agt	ggt	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	2466				
Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys					
	775				780					785					790					
ggc	agt	ggg	att	ttc	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	2514				
Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln					
				795				800						805						
tac	aag	ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	2562				
Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ala	Ile	Gln					
			810					815					820							
aaa	gcc	cat	gaa	gag	gac	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	2610				
Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu					
		825					830					835								
gag	aat	ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	2658				
Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu					
		840				845					850									
tca	gaa	aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	2706				
Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly					
		855			860				865					870						
atc	atg	cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	2754				
Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu					
			875					880					885							
aag	tat	tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	2802				
Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu					
			890					895					900							
tct	cat	aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	2850				
Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys					
		905					910					915								
ccc	aac	aca	aat	aga	gct	tgg	aat	tcg	ttg	gaa	gtt	gaa	gac	tat	ggc	2898				
Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly					
		920				925					930									
ttt	gga	gta	ttc	acc	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	cag	2946				
Phe	Gly	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln					
					940					945					950					
gat	gta	ttc	tgc	gac	tca	aaa	ctc	atg	tca	gcg	gcc	ata	aaa	gac	aac	2994				
Asp	Val	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn					
			955					960					965							
aga	gcc	gtc	cat	gcc	gat	atg	ggt	tat	tgg	ata	gaa	agt	gca	ctc	aat	3042				
Arg	Ala	Val	His	Ala	Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn					
			970					975					980							
gac	aca	tgg	aag	ata	gag	aaa	gcc	tct	ttc	att	gaa	gtt	aaa	aac	tgc	3090				
Asp	Thr	Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys					

cta	aaa	aat	gat	att	ccc	atg	aca	gga	cca	tta	gtg	gct	gga	ggg	ctc	4242
Leu	Lys	Asn	Asp	Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	
ctc	act	gtg	tgc	tac	gtg	ctc	act	gga	cga	tcg	gcc	gat	ttg	gaa	ctg	4290
Leu	Thr	Val	Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu	
gag	aga	gca	gcc	gat	gtc	aaa	tgg	gaa	gac	cag	gca	gag	ata	tca	gga	4338
Glu	Arg	Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	
agc	agt	cca	atc	ctg	tca	ata	aca	ata	tca	gaa	gat	ggg	agc	atg	tcg	4386
Ser	Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser	
ata	aaa	aat	gaa	gag	gaa	gaa	caa	aca	ctg	acc	ata	ctc	att	aga	aca	4434
Ile	Lys	Asn	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg	Thr		
gga	ttg	ctg	gtg	atc	tca	gga	ctt	ttt	cct	gta	tca	ata	cca	atc	acg	4482
Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro	Ile	Thr	
gca	gca	gca	tgg	tac	ctg	tgg	gaa	gtg	aag	aaa	caa	cgg	gcc	gga	gta	4530
Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg	Ala	Gly	Val	
ttg	tgg	gat	gtt	cct	tca	ccc	cca	ccc	atg	gga	aag	gct	gaa	ctg	gaa	4578
Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys	Ala	Glu	Leu	Glu	
gat	gga	gcc	tat	aga	att	aag	caa	aaa	ggg	att	ctt	gga	tat	tcc	cag	4626
Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile	Leu	Gly	Tyr	Ser	Gln	
atc	gga	gcc	gga	gtt	tac	aaa	gaa	gga	aca	ttc	cat	aca	atg	tgg	cat	4674
Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr	Phe	His	Thr	Met	Trp	His	
gtc	aca	cgt	ggc	gct	gtt	cta	atg	cat	aaa	gga	aag	agg	att	gaa	cca	4722
Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	
tca	tgg	gcg	gac	gtc	aag	aaa	gac	cta	ata	tca	tat	gga	gga	ggc	tgg	4770
Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	
aag	tta	gaa	gga	gaa	tgg	aag	gaa	gaa	gaa	gtc	cag	gta	ttg	gca		4818
Lys	Leu	Glu	Gly	Glu	Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	
ctg	gag	cct	gga	aaa	aat	cca	aga	gcc	gtc	caa	acg	aaa	cct	ggg	ctt	4866
Leu	Glu	Pro	Gly	Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	
ttc	aaa	acc	aac	gcc	gga	aca	ata	ggg	gct	gta	tct	ctg	gac	ttt	tct	4914
Phe	Lys	Thr	Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	
cct	gga	acg	tca	gga	tct	cca	att	atc	gac	aaa	aaa	gga	aaa	gtt	gtg	4962
Pro	Gly	Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	
ggg	ctt	tat	ggt	aat	ggt	gtt	gtt	aca	agg	agt	gga	gca	tat	gtg	agt	5010
Gly	Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser	
gct	ata	gcc	cag	act	gaa	aaa	agc	att	gaa	gac	aac	cca	gag	atc	gaa	5058
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile	Glu	
gat	gac	att	ttc	cga	aag	aga	aga	ctg	acc	atc	atg	gac	ctc	cac	cca	5106
Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu	His	Pro	
gga	gcg	gga	aag	acg	aag	aga	tac	ctt	ccg	gcc	ata	gtc	aga	gaa	gct	5154
Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val	Arg	Glu	Ala	
ata	aaa	cgg	ggt	ttg	aga	aca	tta	atc	ttg	gcc	ccc	act	aga	gtt	gtg	5202
Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro	Thr	Arg	Val	Val	
gca	gct	gaa	atg	gag	gaa	gcc	ctt	aga	gga	ctt	cca	ata	aga	tac	cag	5250
Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu	Pro	Ile	Arg	Tyr	Gln	
acc	cca	gcc	atc	aga	gct	gtg	cac	acc	ggg	cgg	gag	att	gtg	gac	cta	5298
Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly	Arg	Glu	Ile	Val	Asp	Leu	

1720	atg tgt cat gcc aca ttt acc atg agg ctg cta tca cca gtt aga gtg	1725	Met Cys His Ala Thr Phe Thr Met Arg Leu Leu Ser Pro Val Arg Val	1730	5346
1735	cca aac tac aac ctg att atc atg gac gaa gcc cat ttc aca gac cca	1740	Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu Ala His Phe Thr Asp Pro	1745	5394
1755	gca agt ata gca gct aga gga tac atc tca act cga gtg gag atg ggt	1760	Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg Val Glu Met Gly	1765	5442
1770	gag gca gct ggg att ttt atg aca gcc act ccc ccg gga agc aga gac	1775	Glu Ala Ala Gly Ile Phe Met Thr Ala Thr Pro Pro Gly Ser Arg Asp	1780	5490
1785	cca ttt cct cag agc aat gca cca atc ata gat gaa gaa aga gaa atc	1790	Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile	1795	5538
1800	cct gaa cgc tcg tgg aat tcc gga cat gaa tgg gtc acg gat ttt aaa	1805	Pro Glu Arg Ser Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys	1810	5586
1815	ggg aag act gtt tgg ttc gtt cca agt ata aaa gca gga aat gat ata	1820	Gly Lys Thr Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile	1825	5634
1835	gca gct tgc ctg agg aaa aat gga aag aaa gtg ata caa ctc agt agg	1840	Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg	1845	5682
1850	aag acc ttt gat tct gag tat gtc aag act aga acc aat gat tgg gac	1855	Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp	1860	5730
1865	ttc gtg gtt aca act gac att tca gaa atg ggt gcc aat ttc aag gct	1870	Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala	1875	5778
1880	gag agg gtt ata gac ccc aga cgc tgc atg aaa cca gtc ata cta aca	1885	Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu Thr	1890	5826
1895	gat ggt gaa gag cgg gtg att ctg gca gga cct atg cca gtg acc cac	1900	Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr His	1905	5874
1915	tct agt gca gca caa aga aga ggg aga ata gga aga aat cca aaa aat	1920	Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro Lys Asn	1925	5922
1930	gag aat gac cag tac ata tac atg ggg gaa cct ctg gaa aat gat gaa	1935	Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu Asn Asp Glu	1940	5970
1945	gac tgt gca cac tgg aaa gaa gct aaa atg ctc cta gat aac atc aac	1950	Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu Asp Asn Ile Asn	1955	6018
1960	acg cca gaa gga atc att cct agc atg ttc gaa cca gag cgt gaa aag	1965	Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu Pro Glu Arg Glu Lys	1970	6066
1975	gtg gat gcc att gat ggc gaa tac cgc ttg aga gga gaa gca agg aaa	1980	Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu Arg Gly Glu Ala Arg Lys	1985	6114
1995	acc ttt gta gac tta atg aga aga gga gac cta cca gtc tgg ttg gcc	2000	Thr Phe Val Asp Leu Met Arg Arg Gly Asp Leu Pro Val Trp Leu Ala	2005	6162
2010	tac aga gtg gca gct gaa ggc atc aac tac gca gac aga agg tgg tgt	2015	Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys	2020	6210
2025	ttt gat gga gtc aag aac aac caa atc cta gaa gaa aac gtg gaa gtt	2030	Phe Asp Gly Val Lys Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val	2035	6258
2040	gaa atc tgg aca aaa gaa ggg gaa agg aag aaa ttg aaa ccc aga tgg	2045	Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp	2050	6306
2055	ttg gat gct agg atc tat tct gac cca ctg gcg cta aaa gaa ttt aag	2060	Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys	2065	6354
2075	gaa ttt gca gcc gga aga aag tct ctg acc ctg aac cta atc aca gaa	2080	Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu	2085	6402

atg ggt agg ctc cca acc ttc atg act cag aag gca aga gac gca ctg	2090	2095	2100	6450
Met Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu	2105	2110	2115	
gac aac tta gca gtg ctg cac acg gct gag gca ggt gga agg gcg tac	2120	2125	2130	6498
Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr	2135	2140	2145	
aac cat gct ctc agt gaa ctg ccg gag acc ctg gag aca ttg ctt tta	2155	2160	2165	6546
Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu	2170	2175	2180	
ctg aca ctt ctg gct aca gtc acg gga ggg atc ttt tta ttc ttg atg	2185	2190	2195	6594
Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu Met	2200	2205	2210	
agc gca agg ggc ata ggg aag atg acc ctg gga atg tgc tgc ata atc	2215	2220	2225	6642
Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Ile Ile	2235	2240	2245	
acg gct agc atc ctc cta tgg tac gca caa ata cag cca cac tgg ata	2250	2255	2260	6690
Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro His Trp Ile	2265	2270	2275	
gca gct tca ata ata ctg gag ttt ttt ctc ata gtt ttg ctt att cca	2280	2285	2290	6738
Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val Leu Leu Ile Pro	2295	2300	2305	
gaa cct gaa aaa cag aga aca ccc caa gac aac caa ctg acc tac gtt	2315	2320	2325	6786
Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn Gln Leu Thr Tyr Val	2330	2335	2340	
gtc ata gcc atc ctc aca gtg gtg gcc gca acc atg gca aac gag atg	2345	2350	2355	6834
Val Ile Ala Ile Leu Thr Val Val Ala Ala Thr Met Ala Asn Glu Met	2360	2365	2370	
ggt ttc cta gaa aaa acg aag aaa gat ctc gga ttg gga agc att gca	2375	2380	2385	6882
Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu Gly Leu Gly Ser Ile Ala	2395	2400	2405	
acc cag caa ccc gag agc aac atc ctg gac ata gat cta cgt cct gca	2410	2415	2420	6930
Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala	2425	2430	2435	
tca gca tgg acg ctg tat gcc gtg gcc aca aca ttt gtt aca cca atg	2440	2445	2450	6978
Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met	2455	2460	2465	
ttg aga cat agc att gaa aat tcc tca gtg aat gtg tcc cta aca gct	2470	2475	2480	7026
Leu Arg His Ser Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala	2485	2490	2495	
ata gcc aac caa gcc aca gtg tta atg ggt ctc ggg aaa gga tgg cca	2500	2505	2510	7074
Ile Ala Asn Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro	2515	2520	2525	
ttg tca aag atg gac atc gga gtt ccc ctt ctc gcc att gga tgc tac	2530	2535	2540	7122
Leu Ser Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr	2545	2550	2555	
tca caa gtc aac ccc ata act ctc aca gca gct ctt ttc tta ttg gta	2560	2565	2570	7170
Ser Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val	2575	2580	2585	
gca cat tat gcc atc ata ggg cca gga ctc caa gca aaa gca acc aga	2590	2595	2600	7218
Ala His Tyr Ala Ile Ile Glu Pro Gly Leu Gln Ala Lys Ala Thr Arg	2605	2610	2615	
gaa gct cag aaa aga gca gcg gcg ggc atc atg aaa aac cca act gtc	2620	2625	2630	7266
Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val	2635	2640	2645	
gat gga ata aca gtg att gac cta gat cca ata cct tat gat cca aag	2650	2655	2660	7314
Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro Lys	2665	2670	2675	
ttt gaa aag cag ttg gga caa gta atg ctc cta gtc ctc tgc gtg act	2680	2685	2690	7362
Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys Val Thr	2695	2700	2705	
caa gta ttg atg atg agg act aca tgg gct ctg tgt gag gct tta acc	2710	2715	2720	7410
Gln Val Leu Met Met Arg Thr Trp Ala Leu Cys Glu Ala Leu Thr	2725	2730	2735	
tta gct acc ggg ccc atc tcc aca ttg tgg gaa gga aat cca ggg agg	2740	2745	2750	7458
Leu Ala Thr Gly Pro Ile tcc Ser Thr Leu Trp Glu Gly Asn Pro Gly Arg	2755	2760	2765	
ttt tgg aac act acc att gcg gtg tca atg gct aac att ttt aga ggg	2770	2775	2780	7506
Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala Asn Ile Phe Arg Gly	2785	2790	2795	

2455	agt	tac	ttg	gcc	gga	gct	gga	ctt	ctc	ttt	tct	att	atg	aag	aac	aca	7554
	Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu	Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	
					2475					2480							
	acc	aac	aca	aga	agg	gga	act	ggc	aac	ata	gga	gag	acg	ctt	gga	gag	7602
	Thr	Asn	Thr	Arg	Arg	Gly	Thr	Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	
					2490					2495							
	aaa	tgg	aaa	agc	cga	ttg	aac	gca	ttg	gga	aaa	agt	gaa	ttc	cag	atc	7650
	Lys	Trp	Lys	Ser	Arg	Leu	Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	
					2505					2510							
	tac	aag	aaa	agt	gga	atc	cag	gaa	gtg	gat	aga	acc	tta	gca	aaa	gaa	7698
	Tyr	Lys	Lys	Ser	Gly	Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	
					2520					2525							
	ggc	att	aaa	aga	gga	gaa	acg	gac	cat	cac	gct	gtg	tgc	cga	ggc	tca	7746
	Gly	Ile	Lys	Arg	Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	
					2535					2540							
	gca	aaa	ctg	aga	tgg	ttc	gtt	gag	aga	aac	atg	gtc	aca	cca	gaa	ggg	7794
	Ala	Lys	Leu	Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	
					2555					2560							
	aaa	gta	gtg	gac	ctc	ggt	tgt	ggc	aga	gga	ggc	tgg	tca	tac	tat	tgt	7842
	Lys	Val	Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	
					2570					2575							
	gga	gga	cta	aag	aat	gta	aga	gaa	gtc	aaa	ggc	cta	aca	aaa	gga	gga	7890
	Gly	Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly	
					2585					2590							
	cca	gga	cac	gaa	gaa	ccc	atc	atg	tca	aca	tat	ggg	tgg	aat	cta		7938
	Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn	Leu	
					2600					2605							
	gtg	cgt	ctt	caa	agt	gga	gtt	gac	gtt	ttc	ttc	atc	ccg	cca	gaa	aag	7986
	Val	Arg	Leu	Gln	Ser	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro	Glu	Lys	
					2615					2620							
	tgt	gac	aca	tta	ttg	tgt	gac	ata	ggg	gag	tca	tca	cca	aat	ccc	aca	8034
	Cys	Asp	Thr	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro	Thr		
					2635					2640							
	gtg	gaa	gca	gga	cga	aca	ctc	aga	gtc	ctt	aac	tta	gta	gaa	aat	tgg	8082
	Val	Glu	Ala	Gly	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu	Asn	Trp	
					2650					2655							
	ttg	aac	aac	aac	act	caa	ttt	tgc	ata	aag	gtt	ctc	aac	cca	tat	atg	8130
	Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn	Pro	Tyr	Met	
					2665					2670							
	ccc	tca	gtc	ata	gaa	aaa	atg	gaa	gca	cta	caa	agg	aaa	tat	gga	gga	8178
	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg	Lys	Tyr	Gly	Gly	
					2680					2685							
	gcc	tta	gtg	agg	aat	cca	ctc	tca	cga	aac	tcc	aca	cat	gag	atg	tac	8226
	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser	Thr	His	Glu	Met	Tyr	
					2700					2705							
	tgg	gta	tcc	aat	gct	ggg	aac	ata	gtg	tca	tca	gtg	aac	atg	att		8274
	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val	Ser	Ser	Val	Asn	Met	Ile	
					2715					2720							
	tca	agg	atg	ttg	atc	aac	aga	ttt	aca	atg	aga	tac	aag	aaa	gcc	act	8322
	Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	
					2730					2735							
	tac	gag	ccg	gat	gtt	gac	ctc	gga	agc	gga	acc	cgt	aac	atc	ggg	att	8370
	Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	
					2745					2750							
	gaa	agt	gag	ata	cca	aac	cta	gat	ata	att	ggg	aaa	aga	ata	gaa	aaa	8418
	Glu	Ser	Glu	Ile	Pro	Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	
					2760					2765							
	ata	aag	caa	gag	cat	gaa	aca	tca	tgg	cac	tat	gac	caa	gac	cac	cca	8466
	Ile	Lys	Gln	Glu	His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	
					2775					2780							
	tac	aaa	acg	tgg	gca	tac	cat	ggt	agc	tat	gaa	aca	aaa	cag	act	gga	8514
	Tyr	Lys	Thr	Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	
					2795					2800							
	tca	gca	tca	tcc	atg	gtc	aac	gga	gtg	gtc	agg	ctg	ctg	aca	aaa	cct	8562
	Ser	Ala	Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	
					2810					2815							
	tgg	gac	gtc	gtc	ccc	atg	gtg	aca	cag	atg	gca	atg	aca	gac	acg	act	8610
	Trp	Asp	Val	Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr	

cca	ttt	gga	caa	cag	cgc	gtt	ttt	aaa	gag	aaa	gtg	gac	acg	aga	acc	8658
Pro	Phe	Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	
2840																
caa	gaa	ccg	aaa	gaa	ggc	acg	aag	aaa	cta	atg	aaa	ata	aca	gca	gag	8706
Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu	
2855																
tgg	ctt	tgg	aaa	gaa	tta	ggg	aag	aaa	aag	aca	ccc	agg	atg	tgc	acc	8754
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys	Thr	
aga	gaa	gaa	ttc	aca	aga	aag	gtg	aga	agc	aat	gca	gcc	ttg	ggg	gcc	8802
Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu	Gly	Ala	
ata	ttc	act	gat	gag	aac	aag	tgg	aag	tgc	gca	cgt	gag	gct	gtt	gaa	8850
Ile	Phe	Thr	Asp	Glu	Asn	Lys	Trp	Lys	Ser	Ala	Arg	Glu	Ala	Val	Glu	
gat	agt	agg	ttt	tgg	gag	ctg	gtt	gac	aag	gaa	agg	aat	ctc	cat	ctt	8898
Asp	Ser	Arg	Phe	Trp	Glu	Leu	Val	Asp	Lys	Glu	Arg	Asn	Leu	His	Leu	
gaa	gga	aag	tgt	gaa	aca	tgt	gtg	tac	aac	atg	atg	gga	aaa	aga	gag	8946
Glu	Gly	Lys	Cys	Glu	Thr	Cys	Val	Tyr	Asn	Met	Met	Gly	Lys	Arg	Glu	
aag	aag	cta	ggg	gaa	ttc	ggc	aag	gca	aaa	ggc	agc	aga	gcc	ata	tgg	8994
Lys	Lys	Leu	Gly	Glu	Phe	Gly	Lys	Ala	Lys	Gly	Ser	Arg	Ala	Ile	Trp	
tac	atg	tgg	ctt	gga	gca	cgc	ttc	tta	gag	ttt	gaa	gcc	cta	gga	ttc	9042
Tyr	Met	Trp	Leu	Gly	Ala	Arg	Phe	Leu	Glu	Phe	Glu	Ala	Leu	Gly	Phe	
tta	aat	gaa	gat	cac	tgg	ttc	tcc	aga	gag	aac	tcc	ctg	agt	gga	gtg	9090
Leu	Asn	Glu	Asp	His	Trp	Phe	Arg	Ser	Glu	Asn	Ser	Leu	Ser	Gly	Val	
gaa	gga	gaa	ggg	ctg	cac	aag	cta	ggt	tac	att	cta	aga	gac	gtg	agc	9138
Glu	Gly	Glu	Gly	Leu	His	Lys	Leu	Gly	Tyr	Ile	Leu	Arg	Asp	Val	Ser	
aag	aaa	gag	gga	gga	gca	atg	tat	gcc	gat	gac	acc	gca	gga	tgg	gat	9186
Lys	Lys	Glu	Gly	Gly	Ala	Met	Tyr	Ala	Asp	Asp	Thr	Ala	Gly	Trp	Asp	
aca	aga	atc	aca	cta	gaa	gac	cta	aaa	aat	gaa	gaa	atg	gta	aca	aac	9234
Thr	Arg	Ile	Thr	Leu	Glu	Asp	Leu	Lys	Asn	Glu	Glu	Met	Val	Thr	Asn	
cac	atg	gaa	gga	gaa	cac	aag	aaa	cta	gcc	gag	gcc	att	ttc	aaa	cta	9282
His	Met	Glu	Gly	Glu	His	Lys	Lys	Leu	Ala	Glu	Ala	Ile	Phe	Lys	Leu	
acg	tac	caa	aac	aag	gtg	gtg	cgt	gtg	caa	aga	cca	aca	cca	aga	ggc	9330
Thr	Tyr	Gln	Asn	Lys	Val	Val	Arg	Val	Gln	Arg	Pro	Thr	Pro	Arg	Gly	
aca	gta	atg	gac	atc	ata	tgc	aga	aga	gac	caa	aga	ggt	agt	gga	caa	9378
Thr	Val	Met	Asp	Ile	Ile	Ser	Arg	Arg	Asp	Gln	Arg	Gly	Ser	Gly	Gln	
gtt	ggc	acc	tat	gga	ctc	aat	act	ttc	acc	aat	atg	gaa	gcc	caa	cta	9426
Val	Gly	Thr	Tyr	Gly	Leu	Asn	Thr	Phe	Thr	Asn	Met	Glu	Ala	Gln	Leu	
atc	aga	cag	atg	gag	gga	gaa	gga	gtc	ttt	aaa	agc	att	cag	cac	cta	9474
Ile	Arg	Gln	Met	Glu	Gly	Glu	Gly	Val	Phe	Lys	Ser	Ile	Gln	His	Leu	
aca	atc	aca	gaa	gaa	atc	gct	gtg	caa	aac	tgg	tta	gca	aga	gtg	ggg	9522
Thr	Ile	Thr	Glu	Glu	Ile	Ala	Val	Gln	Asn	Trp	Leu	Ala	Arg	Val	Gly	
cgc	gaa	agg	tta	tca	aga	atg	gcc	atc	agt	gga	gat	gat	tgt	gtt	gtg	9570
Arg	Glu	Arg	Leu	Ser	Arg	Met	Ala	Ile	Ser	Gly	Asp	Asp	Cys	Val	Val	
aaa	cct	tta	gat	gac	agg	ttc	gca	agc	gct	tta	aca	gct	cta	aat	gac	9618
Lys	Pro	Leu	Asp	Asp	Arg	Phe	Ala	Ser	Ala	Leu	Thr	Ala	Leu	Asn	Asp	
atg	gga	aag	att	agg	aaa	gac	ata	caa	caa	tgg	gaa	cct	tca	aga	gga	9666
Met	Gly	Lys	Ile	Arg	Lys	Asp	Ile	Gln	Gln	Trp	Glu	Pro	Ser	Arg	Gly	
tgg	aat	gat	tgg	aca	caa	gtg	ccc	ttc	tgt	tca	cac	cat	ttc	cat	gag	9714
Trp	Asn	Asp	Trp	Thr	Gln	Val	Pro	Phe	Cys	Ser	His	His	Phe	His	Glu	

```

          3195          3200          3205
tta atc atg aaa gac ggt cgc gta ctc gtt gtt cca tgt aga aac caa 9762
Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln
          3210          3215          3220
gat gaa ctg att ggc aga gcc cga atc tcc caa gga gca ggg tgg tct 9810
Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser
          3225          3230          3235
ttg cgg gag acg gcc tgt ttg ggg aag tct tac gcc caa atg tgg agc 9858
Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser
          3240          3245          3250
ttg atg tac ttc cac aga cgc gac ctc agg ctg gcg gca aat gct att 9906
Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile
          3255          3260          3265          3270
tgc tcg gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg 9954
Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp
          3275          3280          3285
tcc ata cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca 10002
Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr
          3290          3295          3300
gtc tgg aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa 10050
Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
          3305          3310          3315
act cca gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa 10098
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu
          3320          3325          3330
gac caa tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg 10146
Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
          3335          3340          3345          3350
gca aag aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc 10194
Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly
          3355          3360          3365
aat gaa gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga 10242
Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg
          3370          3375          3380
gaa gag gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa 10292
Glu Glu Glu Glu Ala Gly Val Leu Trp
          3385          3390
caaggctaga agtcaggtcg gattaagcca tagtacggaa aaaactatgc tacctgtgag 10352
ccccgtccaa ggacgttaaa agaagtcagg ccatacataaa tgccatagct tgagtaaaact 10412
atgcagcctg tagctccacc tgagaagggtg taaaaaatcc gggaggccac aaaccatgga 10472
agctgtacgc atggcgtagt ggactagcgg ttagaggaga cccctccctt acaaatcgca 10532
gcaacaatgg gggcccaagg cgagatgaag ctgtagtctc gctggaagga ctagagggtta 10592
gaggagaccc ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg 10652
ctgtctcttc agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat 10712
caacaggttc t 10723

```

<210> SEQ ID NO 6

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 6

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

			115					120					125			
Ser	Lys	Gln	Glu	Arg	Gly	Lys	Ser	Leu	Leu	Phe	Lys	Thr	Ser	Ala	Gly	
Val	130	Asn	Met	Cys	Thr	Leu	135	Ala	Met	Asp	Leu	140	Glu	Leu	Cys	Glu
145	Asp	Thr	Met	Thr	Tyr	150	Cys	Pro	Arg	155	Thr	Glu	Ala	Glu	Pro	Asp
Asp	Val	Asp	Cys	165	Trp	Cys	Asn	Ala	Thr	170	Asp	Thr	Trp	Val	Thr	Gly
Thr	Cys	Ser	Gln	180	Thr	Gly	Glu	His	Arg	185	Arg	Asp	Lys	Arg	190	Val
Leu	Ala	Pro	His	195	Val	Gly	Leu	Gly	Leu	200	Glu	Thr	Arg	205	Glu	Thr
Met	210	Ser	Ser	Glu	Gly	Ala	215	Trp	Lys	Gln	Ile	Gln	220	Lys	Val	Thr
225	Ala	Leu	Arg	His	Pro	Gly	Phe	Thr	Val	230	Ile	Ala	Leu	Phe	Leu	Ala
Ala	Ile	Gly	Thr	245	Ser	Ile	Thr	Gln	Lys	250	Gly	Ile	Ile	Phe	Ile	Leu
Met	Leu	Val	Thr	260	Pro	Ser	Met	Ala	Met	265	Arg	Cys	Val	Gly	Ile	Gly
Arg	Asp	Phe	Val	Glu	Gly	Leu	280	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val
Leu	305	Glu	His	Gly	Ser	Cys	Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr
Leu	Asp	Ile	Glu	Leu	310	Leu	Lys	Thr	Glu	Val	Thr	Asn	Pro	Ala	Val	Leu
Arg	Lys	Leu	Cys	Ile	Glu	Ala	Lys	Ile	Ser	330	Asn	Thr	Thr	Thr	Asp	Ser
Arg	Cys	Pro	Thr	Gln	Gly	Glu	Ala	Thr	Leu	345	Val	Glu	Glu	Gln	Asp	Ala
Asn	Phe	Val	Cys	Arg	Arg	Thr	Phe	Val	Asp	360	Arg	Gly	Trp	Gly	Asn	Gly
Cys	Gly	Leu	Phe	Gly	Lys	Gly	Ser	Leu	Ile	375	Thr	Cys	Ala	Lys	Phe	Lys
Cys	Val	Thr	Lys	Leu	Glu	Gly	Lys	Ile	Val	390	Gln	Tyr	Glu	Asn	Leu	Lys
Tyr	Ser	Val	Ile	Val	Thr	Val	His	Thr	Gly	405	Asp	Gln	His	Gln	Val	Gly
Asn	Glu	Thr	Thr	Glu	His	Gly	Thr	Thr	Ala	420	Thr	Ile	Thr	Pro	Gln	Ala
Pro	Thr	Ser	Glu	Ile	Gln	Leu	Thr	Asp	Tyr	440	Gly	Thr	Leu	Thr	Leu	Asp
Cys	450	Ser	Pro	Arg	Thr	Gly	Leu	Asp	Phe	455	Asn	Glu	Val	Leu	Leu	Thr
Met	Lys	Glu	Arg	Ser	Trp	Leu	Val	His	Lys	470	Gln	Trp	Phe	Leu	Asp	Leu
Pro	Leu	Pro	Trp	Thr	Ser	Gly	Ala	Ser	Thr	485	Ser	Gln	Glu	Thr	Trp	Asn
Arg	Gln	Asp	Leu	Leu	Val	Thr	Phe	Lys	Thr	500	Ala	His	Ala	Lys	Lys	Gln
Glu	Val	Val	Val	Leu	Gly	Ser	Gln	Glu	Gly	515	Met	His	Thr	Ala	Leu	
Thr	530	Gly	Ala	Thr	Glu	Ile	Gln	Thr	Ser	535	Thr	Gly	Thr	Ile	Phe	Ala
545	Gly	His	Leu	Lys	Cys	Arg	Leu	Lys	Met	550	Asp	Lys	Leu	Thr	Lys	Gly
Met	Ser	Tyr	Val	Met	Cys	Thr	Gly	Ser	Phe	565	Lys	Leu	Glu	Lys	Glu	Val
Ala	Glu	Thr	Gln	His	Gly	Thr	Val	Leu	Val	580	Gln	Val	Lys	Tyr	Glu	Gly
Thr	Asp	Ala	Pro	Cys	Lys	Ile	Pro	Phe	Ser	595	Gln	Val	Lys	Asp	Lys	Gly
Ala	610	Thr	Gln	Asn	Gly	Arg	Leu	Ile	Thr	600	Ala	Asn	Pro	Ile	Val	Thr
625	Lys	Glu	Lys	Pro	Val	Asn	Ile	Glu	Ala	615	Glu	Pro	Pro	Phe	Gly	Asp
Tyr	Ile	Val	Val	Gly	Ala	Gly	Glu	Lys	Ala	630	Ala	Leu	Lys	Leu	Trp	Phe
										645						
										660						

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
 675 680 685
 Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700
 Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
 705 710 715 720
 Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
 725 730 735
 Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
 740 745 750
 Thr Ser Leu Ser Val Met Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
 755 760 765
 Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
 770 775 780
 Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 785 790 795 800
 His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 805 810 815
 Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
 820 825 830
 Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
 835 840 845
 Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860
 Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875 880
 Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895
 Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910
 Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925
 Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940
 Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960
 Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975
 Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990
 Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005
 Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly
 1010 1015 1020
 Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035 1040
 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 1050 1055
 Gly Thr Thr Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 1065 1070
 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 1080 1085
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys
 1090 1095 1100
 Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu
 1105 1110 1115 1120
 Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
 1125 1130 1135
 Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
 1140 1145 1150
 Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
 1155 1160 1165
 Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
 1170 1175 1180
 Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
 1185 1190 1195 1200
 Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
 1205 1210 1215
 Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr

Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
 1220 1225 1230
 Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
 1235 1240 1245
 Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
 1250 1255 1260
 Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
 1265 1270 1275 1280
 Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser
 1285 1290 1295
 Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
 1300 1305 1310
 Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
 1315 1320 1325
 Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
 1330 1335 1340
 Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
 1345 1350 1355 1360
 Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
 1365 1370 1375
 Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
 1380 1385 1390
 Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
 1395 1400 1405
 Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Gln Thr Leu
 1410 1415 1420
 Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
 1425 1430 1435 1440
 Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
 1445 1450 1455
 Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Met
 1460 1465 1470
 Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
 1475 1480 1485
 Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
 1490 1495 1500
 Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
 1505 1510 1515 1520
 Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
 1525 1530 1535
 Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Gly Glu
 1540 1545 1550
 Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val
 1555 1560 1565
 Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
 1570 1575 1580
 Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
 1585 1590 1595 1600
 Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
 1605 1610 1615
 Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
 1620 1625 1630
 Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr
 1635 1640 1645
 Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1650 1655 1660
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1665 1670 1675 1680
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1685 1690 1695
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
 1700 1705 1710
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1715 1720 1725
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1730 1735 1740
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775

Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr
 1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu
 2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125
 Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205
 Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp
 2210 2215 2220
 Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 2230 2235 2240
 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu
 2245 2250 2255
 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp
 2260 2265 2270
 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr
 2275 2280 2285
 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val
 2290 2295 2300
 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly
 2305 2310 2315 2320
 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu

Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala
 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu
 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile
 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro
 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu
 Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala
 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met
 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile
 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly
 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp
 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His
 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn
 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly
 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys
 Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser
 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe
 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu
 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu
 Asn Leu Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys
 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn
 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met
 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly
 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile
 Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2325 2330 2335
 2340 2345 2350
 2355 2360 2365
 2370 2375 2380
 2385 2390 2395 2400
 2405 2410 2415
 2420 2425 2430
 2435 2440 2445
 2450 2455 2460
 2465 2470 2475 2480
 2485 2490 2495
 2500 2505 2510
 2515 2520 2525
 2530 2535 2540
 2545 2550 2555 2560
 2565 2570 2575
 2580 2585 2590
 2595 2600 2605
 2610 2615 2620
 2625 2630 2635 2640
 2645 2650 2655
 2660 2665 2670
 2675 2680 2685
 2690 2695 2700
 2705 2710 2715 2720
 2725 2730 2735
 2740 2745 2750
 2755 2760 2765
 2770 2775 2780
 2785 2790 2795 2800
 2805 2810 2815
 2820 2825 2830
 2835 2840 2845
 2850 2855 2860
 2865 2870 2875 2880

Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu
 2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe
 3105 3110 3115 3120
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser
 3140 3145 3150
 Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
 3170 3175 3180
 Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195 3200
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
 3205 3210 3215
 Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
 3220 3225 3230
 Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
 3235 3240 3245
 Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
 3250 3255 3260
 Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
 3265 3270 3275 3280
 Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
 3285 3290 3295
 Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
 3300 3305 3310
 Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro
 3315 3320 3325
 Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
 3330 3335 3340
 Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
 3345 3350 3355 3360
 Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
 3365 3370 3375
 Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

<210> SEQ ID NO 7
<211> COMPRIMENTO: 10723
<212> TIPO: ADN
<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:
<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:
<221> NOME/CHAVE: CDS
<222> LOCALIZAÇÃO: (97)...(10272)

<400> SEQUÊNCIA: 7

agttgttagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaatgta	60										
gttctaacag	ttttttaatt	agagagcaga	tctctg	atg	atc aac caa cga aaa	114										
			Met	Ile Asn Gln Arg Lys												
aag	acg	ggt	cga	ccg	tct	ttc	aat	atg	ctg	aaa	cgc	gcg	aga	aac	cgc	162
Lys	Thr	Gly	Arg	Pro	Ser	Phe	Asn	Met	Leu	Lys	Arg	Ala	Arg	Asn	Arg	
			10					15					20			
gtg	tca	act	ggt	tca	cag	ttg	gcg	aag	aga	ttc	tca	aaa	gga	ttg	ctc	210
Val	Ser	Thr	Val	Ser	Gln	Leu	Ala	Lys	Arg	Phe	Ser	Lys	Gly	Leu	Leu	
			25				30					35				
tca	ggc	caa	gga	ccc	atg	aaa	ttg	gtg	atg	gct	ttc	ata	gca	ttc	tta	258
Ser	Gly	Gln	Gly	Pro	Met	Lys	Leu	Val	Met	Ala	Phe	Ile	Ala	Phe	Leu	
			40			45					50					
aga	ttt	cta	gcc	ata	ccc	cca	aca	gca	gga	att	ttg	gct	aga	tgg	ggc	306
Arg	Phe	Leu	Ala	Ile	Pro	Pro	Thr	Ala	Gly	Ile	Leu	Ala	Arg	Trp	Gly	
					60					65					70	
tca	ttc	aag	aag	aat	gga	gcg	att	aaa	gtg	tta	cgg	ggt	ttc	aag	aga	354
Ser	Phe	Lys	Lys	Asn	Gly	Ala	Ile	Lys	Val	Leu	Arg	Gly	Phe	Lys	Arg	
				75					80					85		
gaa	atc	tca	aac	atg	cta	aac	ata	atg	aac	agg	agg	aaa	aga	tcc	gtg	402
Glu	Ile	Ser	Asn	Met	Leu	Asn	Ile	Met	Asn	Arg	Arg	Lys	Arg	Ser	Val	
			90					95					100			
acc	atg	ctc	ctt	atg	ctg	ctg	ccc	aca	gcc	ctg	gcg	ttc	cat	ctg	acg	450
Thr	Met	Leu	Leu	Met	Leu	Leu	Pro	Thr	Ala	Leu	Ala	Phe	His	Leu	Thr	
			105				110					115				
aca	cga	ggg	gga	gag	ccg	cat	atg	ata	ggt	agc	aag	cag	gaa	aga	gga	498
Thr	Arg	Gly	Gly	Glu	Pro	His	Met	Ile	Val	Ser	Lys	Gln	Glu	Arg	Gly	
			120			125					130					
aag	tca	ctt	ttg	ttc	aag	acc	tct	gca	ggt	gtc	aac	atg	tgc	acc	ctc	546
Lys	Ser	Leu	Leu	Phe	Lys	Thr	Ser	Ala	Gly	Val	Asn	Met	Cys	Thr	Leu	
			135		140					145					150	
att	gcg	atg	gat	ttg	gga	gag	ttg	tgt	gag	gac	acg	atg	acc	tac	aaa	594
Ile	Ala	Met	Asp	Leu	Gly	Glu	Leu	Cys	Glu	Asp	Thr	Met	Thr	Tyr	Lys	
				155					160					165		
tgc	ccc	cgg	atc	act	gag	gcg	gaa	cca	gat	gac	ggt	gac	tgt	tgg	tgc	642
Cys	Pro	Arg	Ile	Thr	Glu	Ala	Glu	Pro	Asp	Asp	Val	Asp	Cys	Trp	Cys	
			170					175					180			
aat	gcc	acg	gac	aca	tgg	gtg	acc	tat	gga	acg	tgc	tct	caa	act	ggc	690
Asn	Ala	Thr	Asp	Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Ser	Gln	Thr	Gly	
			185				190					195				
gaa	cac	cga	cga	gac	aaa	cgt	tcc	gtc	gca	ttg	gcc	cca	cac	gtg	ggg	738
Glu	His	Arg	Arg	Asp	Lys	Arg	Ser	Val	Ala	Leu	Ala	Pro	His	Val	Gly	
			200			205					210					
ctt	ggc	cta	gaa	aca	aga	gcc	gaa	acg	tgg	atg	tcc	tct	gaa	ggt	gct	786
Leu	Gly	Leu	Glu	Thr	Arg	Ala	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	
			215		220					225					230	
tgg	aaa	cag	ata	caa	aaa	gta	gag	act	tgg	gct	ctg	aga	cat	cca	gga	834
Trp	Lys	Gln	Ile	Gln	Lys	Val	Glu	Thr	Trp	Ala	Leu	Arg	His	Pro	Gly	
				235					240					245		
ttc	acg	gtg	ata	gcc	ctt	ttt	cta	gca	cat	gcc	ata	gga	aca	tcc	atc	882
Phe	Thr	Val	Ile	Ala	Leu	Phe	Leu	Ala	His	Ala	Ile	Gly	Thr	Ser	Ile	
			250					255				260				
acc	cag	aaa	ggg	atc	att	ttc	att	ttg	ctg	atg	ctg	gta	aca	cca	tct	930
Thr	Gln	Lys	Gly	Ile	Ile	Phe	Ile	Leu	Leu	Met	Leu	Val	Thr	Pro	Ser	
			265				270					275				
atg	gcc	atg	cga	tgc	gtg	gga	ata	ggc	aac	aga	gac	ttc	gtg	gaa	gga	978
Met	Ala	Met	Arg	Cys	Val	Gly	Ile	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	
			280			285					290					
ctg	tca	gga	gca	aca	tgg	gtg	gat	gtg	gta	ctg	gag	cat	gga	agt	tgc	1026
Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val	Leu	Glu	His	Gly	Ser	Cys	
					300					305					310	
gtc	acc	acc	atg	gca	aaa	aac	aaa	cca	aca	ctg	gac	att	gaa	ctc	ttg	1074

Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Ile	Glu	Leu	Leu	
aag	acg	gag	gtc	aca	aac	cct	gca	ggt	ctg	cgt	aaa	ttg	tgc	att	gaa	1122
Lys	Thr	Glu	Val	Thr	Asn	Pro	Ala	Val	Leu	Arg	Lys	Leu	Cys	Ile	Glu	
gct	aaa	ata	tca	aac	acc	acc	acc	gat	tcg	aga	tgt	cca	aca	caa	gga	1170
Ala	Lys	Ile	Ser	Asn	Thr	Thr	Thr	Asp	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
gaa	gcc	aca	ctg	gtg	gaa	gaa	caa	gac	gcg	aac	ttt	gtg	tgc	cga	cga	1218
Glu	Ala	Thr	Leu	Val	Glu	Glu	Gln	Asp	Ala	Asn	Phe	Val	Cys	Arg	Arg	
acg	ttc	gtg	gac	aga	ggc	tgg	ggc	aat	ggc	tgt	ggg	cta	ttc	gga	aaa	1266
Thr	Phe	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
ggt	agt	cta	ata	acg	gcc	aag	ttt	aag	tgt	gtg	aca	aaa	cta	gaa	gaa	1314
Gly	Ser	Leu	Ile	Thr	Cys	Ala	Lys	Phe	Lys	Cys	Val	Thr	Lys	Leu	Glu	
gga	aag	ata	gct	caa	tat	gaa	aac	cta	aaa	tat	tca	gtg	ata	gtc	acc	1362
Gly	Lys	Ile	Ala	Gln	Tyr	Glu	Asn	Leu	Lys	Tyr	Ser	Val	Ile	Val	Thr	
gtc	cac	act	gga	gat	cag	cac	cag	gtg	gga	aat	gag	act	aca	gaa	cat	1410
Val	His	Thr	Gly	Asp	Gln	His	Gln	Val	Gly	Asn	Glu	Thr	Thr	Glu	His	
gga	aca	act	gca	acc	ata	aca	cct	caa	gct	cct	acg	tcg	gaa	ata	cag	1458
Gly	Thr	Thr	Ala	Thr	Ile	Thr	Pro	Gln	Ala	Pro	Thr	Ser	Glu	Ile	Gln	
ctg	acc	gac	tac	gga	acc	ctt	aca	tta	gat	tgt	tca	cct	agg	aca	ggg	1506
Leu	Thr	Asp	Tyr	Gly	Thr	Leu	Thr	Leu	Asp	Cys	Ser	Pro	Arg	Thr	Gly	
cta	gat	ttt	aac	gag	atg	gtg	ttg	ctg	aca	atg	aaa	aag	aaa	tca	tgg	1554
Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Thr	Met	Lys	Lys	Lys	Ser	Trp	
ctt	gtc	cac	aaa	cag	tgg	ttc	cta	gac	tta	cca	ctg	cct	tgg	acc	tct	1602
Leu	Val	His	Lys	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Trp	Thr	Ser	
ggg	gct	tta	aca	tcc	caa	gag	act	tgg	aac	aga	caa	gat	tta	ctg	gtc	1650
Gly	Ala	Leu	Thr	Ser	Gln	Glu	Thr	Trp	Asn	Arg	Gln	Asp	Leu	Leu	Val	
aca	ttt	aag	aca	gct	cat	gca	aag	aag	cag	gaa	gta	gtc	gta	cta	gga	1698
Thr	Phe	Lys	Thr	Ala	His	Ala	Lys	Lys	Gln	Glu	Val	Val	Val	Leu	Gly	
tca	caa	gaa	gga	gca	atg	cac	act	gcg	ctg	act	gga	gcg	aca	gaa	atc	1746
Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile	
caa	acg	tca	gga	acg	aca	aca	att	ttc	gca	gga	cac	cta	aaa	tgc	aga	1794
Gln	Thr	Ser	Gly	Thr	Thr	Thr	Ile	Phe	Ala	Gly	His	Leu	Lys	Cys	Arg	
cta	aaa	atg	gac	aaa	cta	act	tta	aaa	ggg	atg	tca	tat	gtg	atg	tgc	1842
Leu	Lys	Met	Asp	Lys	Leu	Thr	Leu	Lys	Gly	Met	Ser	Tyr	Val	Met	Cys	
aca	ggc	tca	ttc	aag	tta	gag	aaa	gaa	gtg	gct	gag	acc	cag	cat	gga	1890
Thr	Gly	Ser	Phe	Lys	Leu	Glu	Lys	Glu	Val	Ala	Glu	Thr	Gln	His	Gly	
act	gtt	ctg	gtg	cag	gtt	aaa	tat	gaa	gga	aca	gac	gca	cca	tgc	aag	1938
Thr	Val	Leu	Val	Gln	Val	Lys	Tyr	Glu	Gly	Thr	Asp	Ala	Pro	Cys	Lys	
att	ccc	ttt	tcg	acc	caa	gat	gag	aaa	gga	gca	acc	cag	aat	ggg	aga	1986
Ile	Pro	Phe	Ser	Thr	Gln	Asp	Glu	Lys	Gly	Ala	Thr	Gln	Asn	Gly	Arg	
tta	ata	aca	gcc	aac	ccc	ata	gtc	act	gac	aaa	gaa	aaa	cca	gtc	aat	2034
Leu	Ile	Thr	Ala	Asn	Pro	Ile	Val	Thr	Asp	Lys	Glu	Lys	Pro	Val	Asn	
att	gag	gca	gaa	cca	ccc	ttt	ggt	gag	agc	tac	atc	gtg	gta	gga	gca	2082
Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Glu	Ser	Tyr	Ile	Val	Val	Gly	Ala	
ggt	gaa	aaa	gct	ttg	aaa	cta	agc	tgg	ttc	aag	aaa	gga	agc	agc	ata	2130
Gly	Glu	Lys	Ala	Leu	Lys	Leu	Ser	Trp	Phe	Lys	Lys	Gly	Ser	Ser	Ile	
ggg	aaa	atg	ttt	gaa	gca	act	gcc	cga	gga	gca	cga	agg	atg	gcc	att	2178

Gly	Lys	Met	Phe	Glu	Ala	Thr	Ala	Arg	Gly	Ala	Arg	Arg	Met	Ala	Ile	
680	680					685				690						
ctg	gga	gac	acc	gca	tgg	gac	ttc	ggt	tct	ata	gga	gga	gtg	ttc	acg	2226
Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Ile	Gly	Gly	Val	Phe	Thr	
695					700					705					710	
tct	atg	gga	aaa	ctg	gta	cac	cag	ggt	ttt	gga	act	gca	tat	gga	ggt	2274
Ser	Met	Gly	Lys	Leu	Val	His	Gln	Val	Phe	Gly	Thr	Ala	Tyr	Gly	Val	
				715					720					725		
ttg	ttt	agc	gga	ggt	tct	tgg	acc	atg	aaa	ata	gga	ata	ggg	att	ctg	2322
Leu	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	Ile	Gly	Ile	Gly	Ile	Leu	
			730					735					740			
ctg	aca	tgg	cta	gga	tta	aat	tca	agg	aac	acg	tcc	ctt	tcg	gtg	atg	2370
Leu	Thr	Trp	Leu	Gly	Leu	Asn	Ser	Arg	Asn	Thr	Ser	Leu	Ser	Val	Met	
		745					750				755					
tgc	atc	gca	gcc	ggc	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	2418
Cys	Ile	Ala	Ala	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln	
	760				765					770						
gcc	gat	agt	ggt	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	2466
Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys	
	775			780						785				790		
ggc	agt	ggg	att	ttc	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	2514
Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln	
				795				800						805		
tac	aag	ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	2562
Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser		Ile	Gln	
		810						815					820			
aaa	gcc	cat	gaa	gag	gac	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	2610
Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu	
		825					830					835				
gag	aat	ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	2658
Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu	
	840				845						850					
tca	gaa	aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	2706
Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	
				860						865					870	
atc	atg	cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	2754
Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu	
			875					880						885		
aag	tat	tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	2802
Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu	
			890					895					900			
tct	cat	aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	2850
Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys	
		905					910					915				
ccc	aac	aca	aat	aga	gct	tgg	aat	tcg	ttg	gaa	gtt	gaa	gac	tat	ggc	2898
Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly	
		920				925					930					
ttt	gga	gta	ttc	acc	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	cag	2946
Phe	Gly	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln	
		935				940				945					950	
gat	gta	ttc	tgc	gac	tca	aaa	ctc	atg	tca	gcg	gcc	ata	aaa	gac	aac	2994
Asp	Val	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn	
				955					960					965		
aga	gcc	gtc	cat	gcc	gat	atg	ggt	tat	tgg	ata	gaa	agt	gca	ctc	aat	3042
Arg	Ala	Val	His	Ala	Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn	
			970				975						980			
gac	aca	tgg	aag	ata	gag	aaa	gcc	tct	ttc	att	gaa	gtt	aaa	aac	tgc	3090
Asp	Thr	Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys	
			985				990						995			
cac	tgg	cca	aaa	tca	cac	acc	ctc	tgg	agc	aat	gga	gtg	cta	gaa	agt	3138
His	Trp	Pro	Lys	Ser	His	Thr	Leu	Trp	Ser	Asn	Gly	Val	Leu	Glu	Ser	
		1000				1005					1010					
gag	atg	ata	att	cca	aag	aat	ctc	gct	gga	cca	gtg	tct	caa	cac	aac	3186
Glu	Met	Ile	Ile	Pro	Lys	Asn	Leu	Ala	Gly	Pro	Val	Ser	Gln	His	Asn	
		1015				1020				1025					1030	
tat	aga	cca	ggc	tac	cat	aca	caa	ata	aca	gga	cca	tgg	cat	cta	ggt	3234
Tyr	Arg	Pro	Gly	His	His	Thr	Gln	Ile	Thr	Gly	Pro	Trp	His	Leu	Gly	
				1035					1040					1045		
aag	ctt	gag	atg	gac	ttt	gat	ttc	tgt	gat	gga	aca	aca	gtg	gta	gtg	3282

Lys	Leu	Glu	Met	Asp	Phe	Asp	Phe	Cys	Asp	Gly	Thr	Thr	Val	Val	Val	
act	gag	gac	tgc	gga	aat	aga	gga	ccc	tct	ttg	aga	aca	acc	act	gcc	3330
Thr	Glu	Asp	Cys	Gly	Asn	Arg	Gly	Pro	Ser	Leu	Arg	Thr	Thr	Thr	Ala	
tct	gga	aaa	ctc	ata	aca	gaa	tgg	tgc	tgc	cga	tct	tgc	aca	tta	cca	3378
Ser	Gly	Lys	Leu	Ile	Thr	Glu	Trp	Cys	Cys	Arg	Ser	Cys	Thr	Leu	Pro	
ccg	cta	aga	tac	aga	ggt	gag	gat	ggg	tgc	tgg	tac	ggg	atg	gaa	atc	3426
Pro	Leu	Arg	Tyr	Arg	Gly	Glu	Asp	Gly	Cys	Trp	Tyr	Gly	Met	Glu	Ile	
aga	cca	ttg	aag	gag	aaa	gaa	gag	aat	ttg	gtc	aac	tcc	ttg	gtc	aca	3474
Arg	Pro	Leu	Lys	Glu	Lys	Glu	Glu	Asn	Leu	Val	Asn	Ser	Leu	Val	Thr	
gct	gga	cat	ggg	cag	gtc	gac	aac	ttt	tca	cta	gga	gtc	ttg	gga	atg	3522
Ala	Gly	His	Gly	Gln	Val	Asp	Asn	Phe	Ser	Leu	Gly	Val	Leu	Gly	Met	
gca	ttg	ttc	ctg	gag	gaa	atg	ctt	agg	acc	cga	gta	gga	acg	aaa	cat	3570
Ala	Leu	Phe	Leu	Glu	Glu	Met	Leu	Arg	Thr	Arg	Val	Gly	Thr	Lys	His	
gca	ata	cta	cta	gtt	gca	gtt	tct	ttt	gtg	aca	ttg	atc	aca	ggg	aac	3618
Ala	Ile	Leu	Leu	Val	Ala	Val	Ser	Phe	Val	Thr	Leu	Ile	Thr	Gly	Asn	
atg	tcc	ttt	aga	gac	ctg	gga	aga	gtg	atg	gtt	atg	gta	ggc	gcc	act	3666
Met	Ser	Phe	Arg	Asp	Leu	Gly	Arg	Val	Met	Val	Met	Val	Gly	Ala	Thr	
atg	acg	gat	gac	ata	ggt	atg	ggc	gtg	act	tat	ctt	gcc	cta	cta	gca	3714
Met	Thr	Asp	Asp	Ile	Gly	Met	Gly	Val	Thr	Tyr	Leu	Ala	Leu	Leu	Ala	
gcc	ttc	aaa	gtc	aga	cca	act	ttt	gca	gct	gga	cta	ctc	ttg	aga	aag	3762
Ala	Phe	Lys	Val	Arg	Pro	Thr	Phe	Ala	Ala	Gly	Leu	Leu	Leu	Arg	Lys	
ctg	acc	tcc	aag	gaa	ttg	atg	atg	act	act	ata	gga	att	gta	ctc	ctc	3810
Leu	Thr	Ser	Lys	Glu	Leu	Met	Met	Thr	Thr	Ile	Gly	Ile	Val	Leu	Leu	
tcc	cag	agc	acc	ata	cca	gag	acc	att	ctt	gag	ttg	act	gat	ggc	tta	3858
Ser	Gln	Ser	Thr	Ile	Pro	Glu	Thr	Ile	Leu	Glu	Leu	Thr	Asp	Ala	Leu	
gcc	tta	ggc	atg	atg	gtc	ctc	aaa	atg	gtg	aga	aat	atg	gaa	aag	tat	3906
Ala	Leu	Gly	Met	Met	Val	Leu	Lys	Met	Val	Arg	Asn	Met	Glu	Lys	Tyr	
caa	ttg	gca	gtg	act	atc	atg	gct	atc	ttg	tgc	gtc	cca	aac	gca	gtg	3954
Gln	Leu	Ala	Val	Thr	Ile	Met	Ala	Ile	Leu	Cys	Val	Pro	Asn	Ala	Val	
ata	tta	caa	aac	gca	tgg	aaa	gtg	agt	tgc	aca	ata	ttg	gca	gtg	gtg	4002
Ile	Leu	Gln	Asn	Ala	Trp	Lys	Val	Ser	Cys	Thr	Ile	Leu	Ala	Val	Val	
tcc	gtt	tcc	cca	ctg	ttc	tta	aca	tcc	tca	cag	caa	aaa	aca	gat	tgg	4050
Ser	Val	Ser	Pro	Leu	Phe	Leu	Thr	Ser	Ser	Gln	Gln	Lys	Thr	Asp	Trp	
ata	cca	tta	gca	ttg	acg	atc	aaa	ggt	ctc	aat	cca	aca	gct	att	ttt	4098
Ile	Pro	Leu	Ala	Leu	Thr	Ile	Lys	Gly	Leu	Asn	Pro	Thr	Ala	Ile	Phe	
cta	aca	acc	ctc	tca	aga	acc	agc	aag	aaa	agg	agc	tgg	cca	tta	aat	4146
Leu	Thr	Thr	Leu	Ser	Arg	Thr	Ser	Lys	Lys	Arg	Ser	Trp	Pro	Leu	Asn	
gag	gct	atc	atg	gca	gtc	ggg	atg	gtg	agc	att	tta	gcc	agt	tct	ctc	4194
Glu	Ala	Ile	Met	Ala	Val	Gly	Met	Val	Ser	Ile	Leu	Ala	Ser	Ser	Leu	
cta	aaa	aat	gat	att	ccc	atg	aca	gga	cca	tta	gtg	gct	gga	ggg	ctc	4242
Leu	Lys	Asn	Asp	Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	
ctc	act	gtg	tgc	tac	gtg	ctc	act	gga	cga	tgc	gcc	gat	ttg	gaa	ctg	4290
Leu	Thr	Val	Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu	
gag	aga	gca	gcc	gat	gtc	aaa	tgg	gaa	gac	cag	gca	gag	ata	tca	gga	4338
Glu	Arg	Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	
agc	agt	cca	atc	ctg	tca	ata	aca	ata	tca	gaa	gat	ggg	agc	atg	tgc	4386

Ser 1415	Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser 1430	
ata	aaa	aat	gaa	gag	gaa	gaa	caa	aca	ctg	acc	ata	ctc	att	aga	aca	4434
Ile	Lys	Asn	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg	Thr		
gga	ttg	ctg	gtg	atc	tca	gga	ctt	ttt	cct	gta	tca	ata	cca	atc	acg	4482
Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro	Ile	Thr	
gca	gca	gca	tgg	tac	ctg	tgg	gaa	gtg	aag	aaa	caa	cgg	gcc	gga	gta	4530
Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg	Ala	Gly	Val	
ttg	tgg	gat	gtt	cct	tca	ccc	cca	ccc	atg	gga	aag	gct	gaa	ctg	gaa	4578
Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys	Ala	Glu	Leu	Glu	
gat	gga	gcc	tat	aga	att	aag	caa	aaa	ggg	att	ctt	gga	tat	tcc	cag	4626
Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile	Leu	Gly	Tyr	Ser	Gln	
atc	gga	gcc	gga	gtt	tac	aaa	gaa	gga	aca	ttc	cat	aca	atg	tgg	cat	4674
Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr	Phe	His	Thr	Met	Trp	His	
gtc	aca	cgt	ggc	gct	gtt	cta	atg	cat	aaa	gga	aag	agg	att	gaa	cca	4722
Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	
tca	tgg	gcg	gac	gtc	aag	aaa	gac	cta	ata	tca	tat	gga	gga	ggc	tgg	4770
Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	
aag	tta	gaa	gga	gaa	tgg	aag	gaa	gga	gaa	gaa	gtc	cag	gta	ttg	gca	4818
Lys	Leu	Glu	Gly	Glu	Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	
ctg	gag	cct	gga	aaa	aat	cca	aga	gcc	gtc	caa	acg	aaa	cct	ggt	ctt	4866
Leu	Glu	Pro	Gly	Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	
ttc	aaa	acc	aac	gcc	gga	aca	ata	ggt	gct	gta	tct	ctg	gac	ttt	tct	4914
Phe	Lys	Thr	Asn	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser		
cct	gga	acg	tca	gga	tct	cca	att	atc	gac	aaa	aaa	gga	aaa	gtt	gtg	4962
Pro	Gly	Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	
ggt	ctt	tat	ggt	aat	ggt	gtt	gtt	aca	agg	agt	gga	gca	tat	gtg	agt	5010
Gly	Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser	
gct	ata	gcc	cag	act	gaa	aaa	agc	att	gaa	gac	aac	cca	gag	atc	gaa	5058
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile	Glu	
gat	gac	att	ttc	cga	aag	aga	aga	ctg	acc	atc	atg	gac	ctc	cac	cca	5106
Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu	His	Pro	
gga	gcg	gga	aag	acg	aag	aga	tac	ctt	ccg	gcc	ata	gtc	aga	gaa	gct	5154
Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val	Arg	Glu	Ala	
ata	aaa	cgg	ggt	ttg	aga	aca	tta	atc	ttg	gcc	ccc	act	aga	gtt	gtg	5202
Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro	Thr	Arg	Val	Val	
gca	gct	gaa	atg	gag	gaa	gcc	ctt	aga	gga	ctt	cca	ata	aga	tac	cag	5250
Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu	Pro	Ile	Arg	Tyr	Gln	
acc	cca	gcc	atc	aga	gct	gtg	cac	acc	ggg	cgg	gag	att	gtg	gac	cta	5298
Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly	Arg	Glu	Ile	Val	Asp	Leu	
atg	tgt	cat	gcc	aca	ttt	acc	atg	agg	ctg	cta	tca	cca	gtt	aga	gtg	5346
Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg	Leu	Leu	Ser	Pro	Val	Arg	Val	
cca	aac	tac	aac	ctg	att	atc	atg	gac	gaa	gcc	cat	ttc	aca	gac	cca	5394
Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met	Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	
gca	agt	ata	gca	gct	aga	gga	tac	atc	tca	act	cga	gtg	gag	atg	ggt	5442
Ala	Ser	Ile	Ala	Ala	Arg	Gly	Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	
gag	gca	gct	ggg	att	ttt	atg	aca	gcc	act	ccc	ccg	gga	agc	aga	gac	5490

Glu	Ala	Ala	Gly	Ile	Phe	Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	
cca	ttt	cct	cag	agc	aat	gca	cca	atc	ata	gat	gaa	gaa	aga	gaa	atc	5538
Pro	Phe	Pro	Gln	Ser	Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	
1800						1805					1810					
cct	gaa	cgc	tcg	tgg	aat	tcc	gga	cat	gaa	tgg	gtc	acg	gat	ttt	aaa	5586
Pro	Glu	Arg	Ser	Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	
1815					1820					1825					1830	
ggg	aag	act	gtt	tgg	ttc	gtt	cca	agt	ata	aaa	gca	gga	aat	gat	ata	5634
Gly	Lys	Thr	Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	
					1835					1840					1845	
gca	gct	tgc	ctg	agg	aaa	aat	gga	aag	aaa	gtg	ata	caa	ctc	agt	agg	5682
Ala	Ala	Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	
			1850							1855					1860	
aag	acc	ttt	gat	tct	gag	tat	gtc	aag	act	aga	acc	aat	gat	tgg	gac	5730
Lys	Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	
					1865					1870					1875	
ttc	gtg	gtt	aca	act	gac	att	tca	gaa	atg	ggt	gcc	aat	ttc	aag	gct	5778
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	
					1880					1885					1890	
gag	agg	gtt	ata	gac	ccc	aga	cgc	tgc	atg	aaa	cca	gtc	ata	cta	aca	5826
Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr	
1895					1900					1905					1910	
gat	ggt	gaa	gag	cgg	gtg	att	ctg	gca	gga	cct	atg	cca	gtg	acc	cac	5874
Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr	His	
					1915					1920					1925	
tct	agt	gca	gca	caa	aga	aga	ggg	aga	ata	gga	aga	aat	cca	aaa	aat	5922
Ser	Ser	Ala		Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro	Lys	Asn	
					1930					1935					1940	
gag	aat	gac	cag	tac	ata	tac	atg	ggg	gaa	cct	ctg	gaa	aat	gat	gaa	5970
Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu	Asn	Asp	Glu	
					1945					1950					1955	
gac	tgt	gca	cac	tgg	aaa	gaa	gct	aaa	atg	ctc	cta	gat	aac	atc	aac	6018
Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu	Asp	Asn	Ile	Asn	
					1960					1970						
acg	cca	gaa	gga	atc	att	cct	agc	atg	ttc	gaa	cca	gag	cgt	gaa	aag	6066
Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu	Pro	Glu	Arg	Glu	Lys	
1975					1980					1985					1990	
gtg	gat	gcc	att	gat	gaa	tac	cgc	ttg	aga	gga	gaa	gca	agg	aaa		6114
Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr	Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	
					1995					2000					2005	
acc	ttt	gta	gac	tta	atg	aga	aga	gga	gac	cta	cca	gtc	tgg	ttg	gcc	6162
Thr	Phe	Val	Asp	Leu	Met	Arg	Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	
					2010					2015					2020	
tac	aga	gtg	gca	gct	gaa	ggc	atc	aac	tac	gca	gac	aga	agg	tgg	tgt	6210
Tyr	Arg	Val	Ala	Ala	Glu	Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	
					2025					2030					2035	
ttt	gat	gga	gtc	aag	aac	aac	caa	atc	cta	gaa	gaa	aac	gtg	gaa	gtt	6258
Phe	Asp	Gly	Val	Lys	Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	
					2040					2050						
gaa	atc	tgg	aca	aaa	gaa	ggg	gaa	agg	aag	aaa	ttg	aaa	ccc	aga	tgg	6306
Glu	Ile	Trp	Thr	Lys	Gly	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	
2055					2060					2065					2070	
ttg	gat	gct	agg	atc	tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	6354
Leu	Asp	Ala	Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	
					2075					2080					2085	
gaa	ttt	gca	gcc	gga	aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	6402
Glu	Phe	Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	
					2090					2095					2100	
atg	ggt	agg	ctc	cca	acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	6450
Met	Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu	
					2105					2110					2115	
gac	aac	tta	gca	gtg	ctg	cac	acg	gct	gag	gca	ggt	gga	agg	gcg	tac	6498
Asp	Asn	Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr	
					2120					2130						
aac	cat	gct	ctc	agt	gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	6546
Asn	His	Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu	
					2135					2145					2150	
ctg	aca	ctt	ctg	gct	aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	6594

Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met	
agc	gca	agg	ggc	ata	ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	6642
Ser	Ala	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile	
			2170						2175					2180		
acg	gct	agc	atc	ctc	cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	6690
Thr	Ala	Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile	
			2185						2190					2195		
gca	gct	tca	ata	ata	ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	6738
Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro	
			2200						2205					2210		
gaa	cct	gaa	aaa	cag	aga	aca	ccc	caa	gac	aac	caa	ctg	acc	tac	gtt	6786
Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln	Asp	Asn	Gln	Leu	Thr	Tyr	Val	
			2215						2225						2230	
gtc	ata	gcc	atc	ctc	aca	gtg	gtg	gcc	gca	acc	atg	gca	aac	gag	atg	6834
Val	Ile	Ala	Ile	Leu	Thr	Val	Val	Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	
									2240					2245		
ggt	ttc	cta	gaa	aaa	acg	aag	aaa	gat	ctc	gga	ttg	gga	agc	att	gca	6882
Gly	Phe	Leu	Glu	Lys	Thr	Lys	Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	
									2255					2260		
acc	cag	caa	ccc	gag	agc	aac	atc	ctg	gac	ata	gat	cta	cgt	cct	gca	6930
Thr	Gln	Gln	Pro	Glu	Ser	Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	
			2265						2270					2275		
tca	gca	tgg	acg	ctg	tat	gcc	gtg	gcc	aca	aca	ttt	gtt	aca	cca	atg	6978
Ser	Ala	Trp	Thr	Leu	Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	
			2280						2285					2290		
ttg	aga	cat	agc	att	gaa	aat	tcc	tca	gtg	aat	gtg	tcc	cta	aca	gct	7026
Leu	Arg	His	Ser	Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	
			2295						2305						2310	
ata	gcc	aac	caa	gcc	aca	gtg	tta	atg	ggt	ctc	ggg	aaa	gga	tgg	cca	7074
Ile	Ala	Asn	Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	
									2320					2325		
ttg	tca	aag	atg	gac	atc	gga	gtt	ccc	ctt	ctc	gcc	att	gga	tgc	tac	7122
Leu	Ser	Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	
									2335					2340		
tca	caa	gtc	aac	ccc	ata	act	ctc	aca	gca	gct	ctt	ttc	tta	ttg	gta	7170
Ser	Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	
			2345						2350					2355		
gca	cat	tat	gcc	atc	ata	ggg	cca	gga	ctc	caa	gca	aaa	gca	acc	aga	7218
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	
			2360						2365					2370		
gaa	gct	cag	aaa	aga	gca	gcg	ggc	atc	atg	aaa	aac	cca	act	gtc		7266
Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val	
									2385					2390		
gat	gga	ata	aca	gtg	att	gac	cta	gat	cca	ata	cct	tat	gat	cca	aag	7314
Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr	Asp	Pro	Lys	
									2400					2405		
ttt	gaa	aag	cag	ttg	gga	caa	gta	atg	ctc	cta	gtc	ctc	tgc	gtg	act	7362
Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val	Leu	Cys	Val	Thr	
									2415					2420		
caa	gta	ttg	atg	atg	agg	act	aca	tgg	gct	ctg	tgt	gag	gct	tta	acc	7410
Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu	Cys	Glu	Ala	Leu	Thr	
			2425						2430					2435		
tta	gct	acc	ggg	ccc	atc	tcc	aca	ttg	tgg	gaa	gga	aac	cca	ggg	agg	7458
Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp	Glu	Gly	Asn	Pro	Gly	Arg	
			2440						2445					2450		
ttt	tgg	aac	act	acc	att	gcg	gtg	tca	atg	gct	aac	att	ttt	aga	ggg	7506
Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser	Met	Ala	Asn	Ile	Phe	Arg	Gly	
									2465					2470		
agt	tac	ttg	gcc	gga	gct	gga	ctt	ctc	ttt	tct	att	atg	aag	aac	aca	7554
Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu	Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	
									2480					2485		
acc	aac	aca	aga	agg	gga	act	ggc	aac	ata	gga	gag	acg	ctt	gga	gag	7602
Thr	Asn	Thr	Arg	Arg	Gly	Thr	Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	
									2495					2500		
aaa	tgg	aaa	agc	cga	ttg	aac	gca	ttg	gga	aaa	agt	gaa	ttc	cag	atc	7650
Lys	Trp	Lys	Ser	Arg	Leu	Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	
									2510					2515		
tac	aag	aaa	agt	gga	atc	cag	gaa	gtg	gat	aga	acc	tta	gca	aaa	gaa	7698

Tyr	Lys	Lys	Ser	Gly	Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	
2520	2520	2520	2520	2520	2525	2525	2525	2530	2530	2530	2530	2530	2530	2530	2530	
ggc	att	aaa	aga	gga	gaa	acg	gac	cat	cac	gct	gtg	tcg	cga	ggc	tca	7746
Gly	Ile	Lys	Arg	Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	
2535	2535	2535	2535	2540	2540	2540	2540	2545	2545	2545	2545	2545	2545	2550	2550	
gca	aaa	ctg	aga	tgg	ttc	gtt	gag	aga	aac	atg	gtc	aca	cca	gaa	ggg	7794
Ala	Lys	Leu	Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	
2555	2555	2555	2555	2555	2555	2555	2555	2560	2560	2560	2560	2560	2560	2565	2565	
aaa	gta	gtg	gac	ctc	ggt	tgt	ggc	aga	gga	ggc	tgg	tca	tac	tat	tgt	7842
Lys	Val	Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	
2570	2570	2570	2570	2570	2570	2570	2570	2575	2575	2575	2575	2575	2580	2580	2580	
gga	gga	cta	aag	aat	gta	aga	gaa	gtc	aaa	ggc	cta	aca	aaa	gga	gga	7890
Gly	Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly	
2585	2585	2585	2585	2585	2585	2585	2585	2590	2590	2590	2590	2590	2595	2595	2595	
cca	gga	cac	gaa	gaa	ccc	atc	ccc	atg	tca	aca	tat	ggg	tgg	aat	cta	7938
Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn	Leu	
2600	2600	2600	2600	2600	2605	2605	2605	2610	2610	2610	2610	2610	2610	2610	2610	
gtg	cgt	ctt	caa	agt	gga	gtt	gac	gtt	ttc	ttc	atc	ccg	cca	gaa	aag	7986
Val	Arg	Leu	Gln	Ser	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro	Glu	Lys	
2615	2615	2615	2615	2620	2620	2620	2620	2625	2625	2625	2625	2625	2630	2630	2630	
tgt	gac	aca	tta	ttg	tgt	gac	ata	ggg	gag	tca	tca	cca	aat	ccc	aca	8034
Cys	Asp	Thr	Leu	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro	Thr	
2635	2635	2635	2635	2635	2635	2635	2635	2640	2640	2640	2640	2640	2645	2645	2645	
gtg	gaa	gca	gga	cga	aca	ctc	aga	gtc	ctt	aac	tta	gta	gaa	aat	tgg	8082
Val	Glu	Ala	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu	Asn	Trp		
2650	2650	2650	2650	2650	2650	2650	2650	2655	2655	2655	2655	2660	2660	2660	2660	
ttg	aac	aac	aac	act	caa	ttt	tgc	ata	aag	gtt	ctc	aac	cca	tat	atg	8130
Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn	Pro	Tyr	Met	
2665	2665	2665	2665	2665	2665	2665	2670	2670	2670	2670	2670	2675	2675	2675	2675	
ccc	tca	gtc	ata	gaa	aaa	atg	gaa	gca	cta	caa	agg	aaa	tat	gga	gga	8178
Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg	Lys	Tyr	Gly	Gly	
2680	2680	2680	2680	2680	2685	2685	2685	2685	2685	2685	2685	2690	2690	2690	2690	
gcc	tta	gtg	agg	aat	cca	ctc	tca	cga	aac	tcc	aca	cat	gag	atg	tac	8226
Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser	Thr	His	Glu	Met	Tyr	
2695	2695	2695	2695	2695	2700	2700	2700	2700	2700	2700	2700	2705	2705	2710	2710	
tgg	gta	tcc	aat	gct	tcc	ggg	aac	ata	gtg	tca	tca	gtg	aac	atg	att	8274
Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val	Ser	Ser	Val	Asn	Met	Ile	
2715	2715	2715	2715	2715	2715	2715	2715	2720	2720	2720	2720	2720	2725	2725	2725	
tca	agg	atg	ttg	atc	aac	aga	ttt	aca	atg	aga	tac	aag	aaa	gcc	act	8322
Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	
2730	2730	2730	2730	2730	2730	2730	2735	2735	2735	2735	2735	2740	2740	2740	2740	
tac	gag	ccg	gat	gtt	gac	ctc	gga	agc	gga	acc	cgt	aac	atc	ggg	att	8370
Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	
2745	2745	2745	2745	2745	2745	2745	2750	2750	2750	2750	2750	2755	2755	2755	2755	
gaa	agt	gag	ata	cca	aac	cta	gat	ata	att	ggg	aaa	aga	ata	gaa	aaa	8418
Glu	Ser	Glu	Ile	Pro	Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	
2760	2760	2760	2760	2760	2765	2765	2765	2765	2765	2765	2765	2770	2770	2770	2770	
ata	aag	caa	gag	cat	gaa	aca	tca	tgg	cac	tat	gac	caa	gac	cac	cca	8466
Ile	Lys	Gln	Glu	His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	
2775	2775	2775	2775	2780	2780	2780	2780	2780	2785	2785	2785	2785	2790	2790	2790	
tac	aaa	acg	tgg	gca	tac	cat	ggt	agc	tat	gaa	aca	aaa	cag	act	gga	8514
Tyr	Lys	Thr	Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	
2795	2795	2795	2795	2795	2795	2795	2800	2800	2800	2800	2800	2805	2805	2805	2805	
tca	gca	tca	tcc	atg	gtc	aac	gga	gtg	gtc	agg	ctg	ctg	aca	aaa	cct	8562
Ser	Ala	Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	
2810	2810	2810	2810	2810	2810	2810	2815	2815	2815	2815	2815	2820	2820	2820	2820	
tgg	gac	gtc	gtc	ccc	atg	gtg	aca	cag	atg	gca	atg	aca	gac	acg	act	8610
Trp	Asp	Val	Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr	
2825	2825	2825	2825	2825	2825	2825	2830	2830	2830	2830	2830	2835	2835	2835	2835	
cca	ttt	gga	caa	cag	cgc	gtt	ttt	aaa	gag	aaa	gtg	gac	acg	aga	acc	8658
Pro	Phe	Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	
2840	2840	2840	2840	2840	2845	2845	2845	2845	2845	2845	2845	2850	2850	2850	2850	
caa	gaa	ccg	aaa	gaa	ggc	acg	aag	aaa	cta	atg	aaa	ata	aca	gca	gag	8706
Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu	
2855	2855	2855	2855	2855	2860	2860	2860	2860	2865	2865	2865	2865	2865	2865	2870	
tgg	ctt	tgg	aaa	gaa	tta	ggg	aag	aaa	aag	aca	ccc	agg	atg	tgc	acc	8754
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys	Thr	
2875	2875	2875	2875	2875	2875	2875	2880	2880	2880	2880	2880	2885	2885	2885	2885	
aga	gaa	gaa	ttc	aca	aga	aag	gtg	aga	agc	aat	gca	gcc	ttg	ggg	gcc	8802

Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu	Gly	Ala		
			2890					2895					2900				
ata	ttc	act	gat	gag	aac	aag	tgg	aag	tcg	gca	cgt	gag	gct	gtt	gaa	8850	
Ile	Phe	Thr	Asp	Glu	Asn	Lys	Trp	Lys	Ser	Ala	Arg	Glu	Ala	Val	Glu		
			2905				2910					2915					
gat	agt	agg	ttt	tgg	gag	ctg	gtt	gac	aag	gaa	agg	aat	ctc	cat	ctt	8898	
Asp	Ser	Arg	Phe	Trp	Glu	Leu	Val	Asp	Lys	Glu	Arg	Asn	Leu	His	Leu		
			2920				2925				2930						
gaa	gga	aag	tgt	gaa	aca	tgt	gtg	tac	aac	atg	atg	gga	aaa	aga	gag	8946	
Glu	Gly	Lys	Cys	Glu	Thr	Cys	Val	Tyr	Asn	Met	Met	Gly	Lys	Arg	Glu		
			2935			2940				2945					2950		
aag	aag	cta	ggg	gaa	ttc	ggc	aag	gca	aaa	ggc	agc	aga	gcc	ata	tgg	8994	
Lys	Lys	Leu	Gly	Glu	Phe	Gly	Lys	Ala	Lys	Gly	Ser	Arg	Ala	Ile	Trp		
			2955					2960						2965			
tac	atg	tgg	ctt	gga	gca	cgc	ttc	tta	gag	ttt	gaa	gcc	cta	gga	ttc	9042	
Tyr	Met	Trp	Leu	Gly	Ala	Arg	Phe	Leu	Glu	Phe	Glu	Ala	Leu	Gly	Phe		
			2970					2975						2980			
tta	aat	gaa	gat	cac	tgg	ttc	tcc	aga	gag	aac	tcc	ctg	agt	gga	gtg	9090	
Leu	Asn	Glu	Asp	His	Trp	Phe	Ser	Arg	Glu	Asn	Ser	Leu	Ser	Gly	Val		
			2985				2990					2995					
gaa	gga	gaa	ggg	ctg	cac	aag	cta	ggt	tac	att	cta	aga	gac	gtg	agc	9138	
Glu	Gly	Glu	Gly	Leu	His	Lys	Leu	Gly	Tyr	Ile	Leu	Arg	Asp	Val	Ser		
			3000			3005					3010						
aag	aaa	gag	gga	gga	gca	atg	tat	gcc	gat	gac	acc	gca	gga	tgg	gat	9186	
Lys	Lys	Glu	Gly	Gly	Ala	Met	Tyr	Ala	Asp	Asp	Thr	Ala	Gly	Trp	Asp		
			3015			3020			3025					3030			
aca	aga	atc	aca	cta	gaa	gac	cta	aaa	aat	gaa	gaa	atg	gta	aca	aac	9234	
Thr	Arg	Ile	Thr	Leu	Glu	Asp	Leu	Lys	Asn	Glu	Glu	Met	Val	Thr	Asn		
			3035					3040						3045			
cac	atg	gaa	gga	cac	aag	aaa	cta	gcc	gag	gcc	att	ttc	aaa	cta		9282	
His	Met	Glu	Gly	Glu	His	Lys	Lys	Leu	Ala	Glu	Ala	Ile	Phe	Lys	Leu		
			3050				3055					3060					
acg	tac	caa	aac	aag	gtg	gtg	cgt	gtg	caa	aga	cca	aca	cca	aga	ggc	9330	
Thr	Tyr	Gln	Asn	Lys	Val	Val	Arg	Val	Gln	Arg	Pro	Thr	Pro	Arg	Gly		
			3065				3070					3075					
aca	gta	atg	gac	atc	ata	tcg	aga	aga	gac	caa	aga	ggt	agt	gga	caa	9378	
Thr	Val	Met	Asp	Ile	Ile	Ser	Arg	Arg	Asp	Gln	Arg	Gly	Ser	Gly	Gln		
			3080			3085					3090						
gtt	ggc	acc	tat	gga	ctc	aat	act	ttc	acc	aat	atg	gaa	gcc	caa	cta	9426	
Val	Gly	Thr	Tyr	Gly	Leu	Asn	Thr	Phe	Thr	Asn	Met	Glu	Ala	Gln	Leu		
			3095			3100				3105				3110			
atc	aga	cag	atg	gag	gga	gaa	gga	gtc	ttt	aaa	agc	att	cag	cac	cta	9474	
Ile	Arg	Gln	Met	Glu	Gly	Glu	Gly	Val	Phe	Lys	Ser	Ile	Gln	His	Leu		
			3115					3120						3125			
aca	atc	aca	gaa	atc	gct	gtg	caa	aac	tgg	tta	gca	aga	gtg	ggg		9522	
Thr	Ile	Thr	Glu	Glu	Ile	Ala	Val	Gln	Asn	Trp	Leu	Ala	Arg	Val	Gly		
			3130				3135					3140					
cgc	gaa	agg	tta	tca	aga	atg	gcc	atc	agt	gga	gat	gat	tgt	gtt	gtg	9570	
Arg	Glu	Arg	Leu	Ser	Arg	Met	Ala	Ile	Ser	Gly	Asp	Asp	Cys	Val	Val		
			3145			3150					3155						
aaa	cct	tta	gat	gac	agg	ttc	gca	agc	gct	tta	aca	gct	cta	aat	gac	9618	
Lys	Pro	Leu	Asp	Asp	Arg	Phe	Ala	Ser	Ala	Leu	Thr	Ala	Leu	Asn	Asp		
			3160			3165					3170						
atg	gga	aag	att	agg	aaa	gac	ata	caa	caa	tgg	gaa	cct	tca	aga	gga	9666	
Met	Gly	Lys	Ile	Arg	Lys	Asp	Ile	Gln	Gln	Trp	Glu	Pro	Ser	Arg	Gly		
			3175			3180				3185					3190		
tgg	aat	gat	tgg	aca	caa	gtg	ccc	ttc	tgt	tca	cac	cat	ttc	cat	gag	9714	
Trp	Asn	Asp	Trp	Thr	Gln	Val	Pro	Phe	Cys	Ser	His	His	Phe	His	Glu		
			3195					3200						3205			
tta	atc	atg	aaa	gac	ggt	cgc	gta	ctc	gtt	gtt	cca	tgt	aga	aac	caa	9762	
Leu	Ile	Met	Lys	Asp	Gly	Arg	Val	Leu	Val	Val	Pro	Cys	Arg	Asn	Gln		
			3210				3215					3220					
gat	gaa	ctg	att	ggc	aga	gcc	cga	atc	tcc	caa	gga	gca	ggg	tgg	tct	9810	
Asp	Glu	Leu	Ile	Gly	Arg	Ala	Arg	Ile	Ser	Gln	Gly	Ala	Gly	Trp	Ser		
			3225				3230				3235						
ttg	cgg	gag	acg	gcc	tgt	ttg	ggg	aag	tct	tac	gcc	caa	atg	tgg	agc	9858	
Leu	Arg	Glu	Thr	Ala	Cys	Leu	Gly	Lys	Ser	Tyr	Ala	Gln	Met	Trp	Ser		
			3240			3245					3250						
ttg	atg	tac	ttc	cac	aga	cgc	gac	ctc	agg	ctg	gcg	gca	aat	gct	att	9906	

Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile	
3255 3260 3265 3270	
tgc tgc gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg	9954
Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp	
3275 3280 3285	
tcc ata cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca	10002
Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr	
3290 3300	
gtc tgg aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa	10050
Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys	
3305 3310 3315	
act cca gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa	10098
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu	
3320 3325 3330	
gac caa tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg	10146
Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp	
3335 3340 3345 3350	
gca aag aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc	10194
Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly	
3355 3360 3365	
aat gaa gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga	10242
Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg	
3370 3375 3380	
gaa gag gaa gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa	10292
Glu Glu Glu Glu Ala Gly Val Leu Trp *	
3385 3390	
caaggctaga agtcaggtcg gattaagcca tagtacggaa aaaactatgc tacctgtgag	10352
ccccgtccaa ggacgttaaa agaagtcagg ccatcataaa tgccatagct tgagttaaact	10412
atgcagcctg tagctccacc tgagaagggtg taaaaaatcc gggaggccac aaaccatgga	10472
agctgtacgc atggcgtagt ggactagcgg ttagaggaga cccctccctt acaaatcgca	10532
gcaacaatgg gggcccaagg cgagatgaag ctgtagtctc gctggaagga ctagaggtta	10592
gaggagaccc ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg	10652
ctgtctcctc agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat	10712
caacagggttc t	10723

<210> SEQ ID NO 8

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 8

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

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
 210 215 220
 Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
 225 230 235 240
 Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
 245 250 255
 Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
 260 265 270
 Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
 275 280 285
 Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
 290 295 300
 Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
 305 310 315 320
 Leu Asp Ile Glu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
 325 330 335
 Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
 340 345 350
 Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
 355 360 365
 Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
 370 375 380
 Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
 385 390 395 400
 Cys Val Thr Lys Leu Glu Gly Lys Ile Ala Gln Tyr Glu Asn Leu Lys
 405 410 415
 Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
 420 425 430
 Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
 435 440 445
 Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
 450 455 460
 Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
 465 470 475 480
 Met Lys Lys Lys Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
 485 490 495
 Pro Leu Pro Trp Thr Ser Gly Ala Leu Thr Ser Gln Glu Thr Trp Asn
 500 505 510
 Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
 515 520 525
 Glu Val Val Val Leu Gly Ser Gln Glu Gly Ala Met Thr Ala Leu
 530 535 540
 Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
 545 550 555 560
 Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
 565 570 575
 Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
 580 585 590
 Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
 595 600 605
 Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
 610 615 620
 Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
 625 630 635 640
 Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
 645 650 655
 Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
 660 665 670
 Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
 675 680 685
 Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700
 Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
 705 710 715 720
 Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
 725 730 735
 Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
 740 745 750
 Thr Ser Leu Ser Val Met Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
 755 770 775 780 785
 Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 790 795 800
 His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 805 810 815
 Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
 820 825 830
 Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
 835 840 845
 Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860
 Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875
 Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895
 Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910
 Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925
 Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940
 Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955
 Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975
 Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990
 Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005
 Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly
 1010 1015 1020
 Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035
 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 1050 1055
 Gly Thr Thr Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 1065 1070
 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 1080 1085
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys
 1090 1095 1100
 Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu
 1105 1110 1115
 Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
 1125 1130 1135
 Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
 1140 1145 1150
 Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
 1155 1160 1165
 Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
 1170 1175 1180
 Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
 1185 1190 1195
 Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
 1205 1210 1215
 Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr
 1220 1225 1230
 Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
 1235 1240 1245
 Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
 1250 1255 1260
 Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
 1265 1270 1275
 Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
 1285 1290 1295
 Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser
 1300 1305 1310

Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
 1315 1320 1325
 Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
 1330 1335 1340
 Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
 1345 1350 1355 1360
 Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
 1365 1370 1375
 Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
 1380 1385 1390
 Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
 1395 1400 1405
 Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
 1410 1415 1420
 Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Glu Thr Leu
 1425 1430 1435 1440
 Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
 1445 1450 1455
 Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
 1460 1465 1470
 Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Met
 1475 1480 1485
 Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
 1490 1495 1500
 Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
 1505 1510 1515 1520
 Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
 1525 1530 1535
 Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
 1540 1545 1550
 Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu
 1555 1560 1565
 Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val
 1570 1575 1580
 Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
 1585 1590 1595 1600
 Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
 1605 1610 1615
 Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
 1620 1625 1630
 Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
 1635 1640 1645
 Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr
 1650 1655 1660
 Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1665 1670 1675 1680
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1685 1690 1695
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1700 1705 1710
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
 1715 1720 1725
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1730 1735 1740
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775
 Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr

1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu
 2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125
 Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205
 Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp
 2210 2215 2220
 Asn Gln Leu Thr Tyr Val Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 2230 2235 2240
 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu
 2245 2250 2255
 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp
 2260 2265 2270
 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr
 2275 2280 2285
 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val
 2290 2295 2300
 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly
 2305 2310 2315 2320
 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu
 2325 2330 2335
 Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala
 2340 2345 2350
 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu
 2355 2360 2365
 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Gly Ile
 2370 2375 2380
 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro
 2385 2390 2395 2400
 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu
 2405 2410 2415

Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala
 2420 2425 2430
 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
 2435 2440 2445
 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met
 2450 2455 2460
 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
 2465 2470 2475 2480
 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile
 2485 2490 2495
 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly
 2500 2505 2510
 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp
 2515 2520 2525
 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His
 2530 2535 2540
 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn
 2545 2550 2555 2560
 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly
 2565 2570 2575
 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys
 2580 2585 2590
 Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser
 2595 2600 2605
 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe
 2610 2615 2620
 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu
 2625 2630 2635 2640
 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu
 2645 2650 2655
 Asn Leu Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys
 2660 2665 2670
 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
 2675 2680 2685
 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn
 2690 2695 2700
 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
 2705 2710 2715 2720
 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met
 2725 2730 2735
 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly
 2740 2745 2750
 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile
 2755 2760 2765
 Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 2770 2775 2780
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 2785 2790 2795 2800
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 2805 2810 2815
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 2820 2825 2830
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 2835 2840 2845
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 2850 2855 2860
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2865 2870 2875 2880
 Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu

2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Val Phe
 3105 3110 3115 3120
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser
 3140 3145 3150
 Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
 3170 3175 3180
 Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195 3200
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
 3205 3210 3215
 Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
 3220 3225 3230
 Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
 3235 3240 3245
 Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
 3250 3255 3260
 Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
 3265 3270 3275 3280
 Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
 3285 3290 3295
 Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
 3300 3305 3310
 Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Ile Pro
 3315 3320 3325
 Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
 3330 3335 3340
 Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
 3345 3350 3355 3360
 Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
 3365 3370 3375
 Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

<210> SEQ ID NO 9

<211> COMPRIMENTO: 10717

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

EP2278012B1

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10266)

<400> SEQUÊNCIA 9

agttgtagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaatgta	60				
gttctaacag	ttttttaatt	agagagcaga	tctctg	atg	aat	aac	caa	cgg	aaa	114
				Met	Asn	Asn	Gln	Arg	Lys	
				1				5		

aag	gcg	aaa	aac	acg	cct	ttc	aat	atg	ctg	aaa	cgc	gag	aga	aac	cgc	162
Lys	Ala	Lys	Asn	Thr	Pro	Phe	Asn	Met	Leu	Lys	Arg	Glu	Arg	Asn	Arg	
			10					15					20			
gtg	tcg	act	gtg	caa	cag	ctg	aca	aag	aga	ttc	tca	ctt	gga	atg	ctg	210
Val	Ser	Thr	Val	Gln	Gln	Leu	Thr	Lys	Arg	Phe	Ser	Leu	Gly	Met	Leu	
		25				30						35				
cag	gga	cga	gga	cca	tta	aaa	ctg	ttc	atg	gcc	ctg	gtg	gag	ttc	ctt	258
Gln	Gly	Arg	Gly	Pro	Leu	Lys	Leu	Phe	Met	Ala	Leu	Val	Ala	Phe	Leu	
	40					45				50						
cgt	ttc	cta	aca	atc	cca	cca	aca	gca	ggg	ata	ttg	aag	aga	tgg	gga	306
Arg	Phe	Leu	Thr	Ile	Pro	Pro	Thr	Ala	Gly	Ile	Leu	Lys	Arg	Trp	Gly	
	55				60					65				70		
aca	att	aaa	aaa	tca	aaa	gct	att	aat	gtt	ttg	aga	ggg	ttc	agg	aaa	354
Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val	Leu	Arg	Gly	Phe	Arg	Lys	
				75					80					85		
gag	att	gga	agg	atg	ctg	aac	atc	ttg	aat	agg	aga	cgc	aga	tct	gca	402
Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn	Arg	Arg	Arg	Arg	Ser	Ala	
			90					95					100			
ggc	atg	atc	att	atg	ctg	att	cca	aca	gtg	atg	gcg	ttc	cat	tta	acc	450
Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val	Met	Ala	Phe	His	Leu	Thr	
	105					110						115				
acg	cgt	gat	gga	gag	ccg	cgc	atg	att	gtg	ggg	aag	aat	gaa	aga	gga	498
Thr	Arg	Asp	Gly	Glu	Pro	Arg	Met	Ile	Val	Gly	Lys	Asn	Glu	Arg	Gly	
	120				125					130						
aaa	tcc	cta	ctt	ttc	aag	aca	gcc	tct	gga	atc	aac	atg	tgc	aca	ctc	546
Lys	Ser	Leu	Leu	Phe	Lys	Thr	Ala	Ser	Gly	Ile	Asn	Met	Cys	Thr	Leu	
	135			140					145					150		
ata	gct	atg	gat	ctg	gga	gag	atg	tgt	gat	gac	acg	gtc	act	tac	aaa	594
Ile	Ala	Met	Asp	Leu	Gly	Glu	Met	Cys	Asp	Asp	Thr	Val	Thr	Tyr	Lys	
			155					160					165			
tgc	ccc	cac	att	acc	gaa	gtg	gag	cct	gaa	gac	att	gac	tgc	tgg	tgc	642
Cys	Pro	His	Ile	Thr	Glu	Val	Glu	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	
			170					175					180			
aac	ctt	aca	tcg	aca	tgg	gtg	act	tat	gga	aca	tgc	aat	caa	gct	gga	690
Asn	Leu	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Asn	Gln	Ala	Gly	
	185					190						195				
gag	cat	aga	cgc	gat	aag	aga	tca	gtg	gcg	tta	gct	ccc	cat	gtt	ggc	738
Glu	His	Arg	Arg	Asp	Lys	Arg	Ser	Val	Ala	Leu	Ala	Pro	His	Val	Gly	
	200			205						210						
atg	gga	ctg	gac	aca	cgc	act	caa	acc	tgg	atg	tcg	gct	gaa	gga	gct	786
Met	Gly	Leu	Asp	Thr	Arg	Thr	Gln	Thr	Trp	Met	Ser	Ala	Glu	Gly	Ala	
	215			220				225						230		
tgg	aga	caa	gtc	gag	aag	gta	gag	aca	tgg	gcc	ctt	agg	cac	cca	ggg	834
Trp	Arg	Gln	Val	Glu	Lys	Val	Glu	Thr	Trp	Ala	Leu	Arg	His	Pro	Gly	
			235					240					245			
ttt	acc	ata	cta	gcc	cta	ttt	ctt	gcc	cat	tac	ata	ggc	act	tcc	ttg	882
Phe	Thr	Ile	Leu	Ala	Leu	Phe	Leu	Ala	His	Tyr	Ile	Gly	Thr	Ser	Leu	
		250				255						260				
acc	cag	aaa	gtg	gtt	att	ttt	ata	cta	tta	atg	ctg	gtt	acc	cca	tcc	930
Thr	Gln	Lys	Val	Val	Ile	Phe	Ile	Leu	Leu	Met	Leu	Val	Thr	Pro	Ser	
	265					270						275				
atg	aca	atg	aga	tgt	gta	gga	gta	gga	aac	aga	gat	ttt	gtg	gaa	ggc	978
Met	Thr	Met	Arg	Cys	Val	Gly	Val	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	
	280				285					290						
cta	tcg	gga	gct	acg	tgg	gtt	gac	gtg	gtg	ctc	gag	cac	ggt	ggg	tgt	1026
Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val	Leu	Glu	His	Gly	Gly	Cys	
	295			300				305						310		
gtg	act	acc	atg	gct	aag	aac	aag	ccc	acg	ctg	gac	ata	gag	ctt	cag	1074
Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Ile	Glu	Leu	Gln	
			315					320						325		
aag	acc	gag	gcc	acc	caa	ctg	gcg	acc	cta	agg	aag	cta	tgc	att	gag	1122
Lys	Thr	Glu	Ala	Thr	Gln	Leu	Ala	Thr	Leu	Arg	Lys	Leu	Cys	Ile	Glu	
			330				335						340			
gga	aaa	att	acc	aac	ata	aca	acc	gac	tca	aga	tgt	ccc	acc	caa	ggg	1170
Gly	Lys	Ile	Thr	Asn	Ile	Thr	Thr	Asp	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
		345				350						355				
gaa	gcg	att	tta	cct	gag	gag	cag	cag	aac	tac	gtg	tgt	aag	cat		1218
Glu	Ala	Ile	Leu	Pro	Glu	Glu	Gln	Asp	Gln	Asn	Val	Cys	Lys	His		
	360				365					370						

aca Thr 375	tac Tyr	gtg Val	gac Asp	aga Arg	ggc Gly 380	tgg Trp	gga Gly	aac Asn	ggt Gly	tgt Cys 385	ggt Gly	ttg Leu	ttt Phe	ggc Gly	aag Lys 390	1266
gga Gly 400	agc Ser	ttg Leu	gtg Val	aca Thr 395	tgc Cys	gcg Ala	aaa Lys	ttt Phe	caa Gln 400	tgt Cys	tta Leu	gaa Glu	tca Ser	ata Ile 405	gag Glu	1314
gga Gly 410	aaa Lys	gtg Val	gtg Val	caa Gln	cat His	gag Glu	aac Asn	ctc Leu 415	aaa Lys	tac Tyr	acc Thr	gtc Val	atc Ile 420	atc Ile	aca Thr	1362
gtg Val 425	cac His	aca Thr	gga Gly	gac Asp	caa Gln	cac His	cag Gln	gtg Val 430	gga Gly	aat Asn	gaa Glu	acg Thr	cag Gln	gga Gly	gtc Val	1410
acg Thr 440	gct Glu	gag Ile	ata Ile	aca Thr	ccc Pro	cag Gln	gca Ala	tca Ser	acc Thr	gct Ala	gaa Glu	gcc Ala	att Ile	tta Leu	cct Pro	1458
gaa Glu 455	tat Tyr	gga Gly	acc Thr	ctc Leu	ggg Leu 460	cta Leu	gaa Glu	tgc Cys	tca Ser	cca Pro	cgg Arg	aca Thr	ggt Gly	ttg Leu	gat Asp 470	1506
ttc Phe	aat Asn	gaa Glu	atg Met	atc Ile 475	tca Ser	ttg Leu	aca Thr	atg Met	aag Lys 480	aac Asn	aaa Lys	gca Ala	tgg Trp	atg Met	gta Val	1554
cat His	aga Arg	caa Gln	tgg Trp 490	ttc Phe	ttt Phe	gac Asp	tta Leu	ccc Pro 495	cta Leu	cca Pro	tgg Trp	aca Thr	tca Ser 500	gga Gly	gct Ala	1602
aca Thr 505	gca Ala	gaa Glu	aca Thr	cca Pro	act Thr	tgg Trp	aac Asn 510	agg Arg	aaa Lys	gag Glu	ctt Leu	ctt Leu	gtg Val	aca Thr	ttt Phe	1650
aaa Lys 520	aat Asn	gca Ala	cat His	gca Ala	aaa Lys	aag Lys	caa Gln	gaa Glu	gta Val	gtt Val	gtt Val	ctt Leu	gga Gly	tca Ser	caa Gln	1698
gag Glu 535	gga Gly	gca Ala	atg Met	cat His	aca Thr 540	gca Ala	ctg Leu	aca Thr	gga Gly	gct Ala	aca Thr	gag Glu	atc Ile	caa Gln	acc Thr 550	1746
tca Ser	gga Gly	ggc Gly	aca Thr	agt Ser 555	atc Ile	ttt Phe	gcg Ala	ggg Gly	cac His 560	tta Leu	aaa Lys	tgt Cys	aga Arg	ctc Leu	aag Lys	1794
atg Met	gac Asp	aaa Lys	ttg Leu 570	gaa Glu	ctc Leu	aag Lys	ggg Gly	atg Met	agc Ser	tat Tyr	gca Ala	atg Met	tgc Cys 580	ttg Leu	agt Ser	1842
agc Ser	ttt Phe	gtg Val	ttg Leu	aag Lys	aaa Lys	gaa Glu	gtc Val 590	tcc Ser	gaa Glu	acg Thr	cag Gln	cat His 595	ggg Gly	aca Thr	ata Ile	1890
ctc Leu	att Ile 600	aag Lys	gtt Val	gag Glu	tac Tyr	aaa Lys 605	ggg Gly	gaa Glu	gat Asp	gca Ala	ccc Pro	tgc Cys	aag Lys	att Ile	cct Pro	1938
ttc Phe 615	tcc Ser	acg Thr	gag Glu	gat Asp	gga Gly	caa Gln	gga Gly	aaa Lys	gct Ala	ctc Leu	aat Asn	ggc Gly	aga Arg	ctg Leu	atc Ile 630	1986
aca Thr	gcc Ala	aat Asn	cca Pro	gtg Val 635	gtg Val	acc Thr	aag Lys	aag Lys	gag Glu	gag Glu	cct Pro	gtc Val	aac Asn	att Ile 645	gag Glu	2034
gct Ala	gaa Glu	cct Pro	cct Pro	ttt Phe 650	gga Gly	gaa Glu	agt Ser	aac Asn	ata Ile	gta Val	att Ile	gga Gly	att Ile	gga Gly	gac Asp	2082
aaa Lys	gcc Ala	ctg Leu	aaa Lys	atc Ile	aac Asn	tgg Trp	tac Tyr 670	aag Lys	gga Lys	agc Ser	tcg Ser	att Ile	ggg Gly	aag Lys		2130
atg Met	ttc Phe 680	gag Glu	gcc Ala	act Thr	gcc Ala	aga Arg 685	ggt Gly	gca Ala	agg Arg	cgc Arg	atg Met	gcc Ala	atc Ile	ttg Leu	gga Gly	2178
gac Asp 695	aca Thr	gcc Ala	tgg Trp	gac Asp	ttt Phe 700	gga Gly	tca Ser	gtg Val	ggt Gly	ggt Gly 705	gtt Val	ttg Leu	aat Asn	tca Ser	tta Leu 710	2226
ggg Gly	aaa Lys	atg Met	gtc Val	cac His 715	caa Gln	ata Ile	ttt Phe	ggg Gly	agt Ser 720	gct Ala	tac Tyr	aca Thr	gcc Ala	cta Leu	ttt Phe 725	2274
ggt Gly	gga Gly	gtc Val	tcc Ser 730	tgg Trp	atg Met	atg Met	aaa Lys	att Ile 735	gga Gly	ata Ile	ggt Gly	gtc Val	ctc Leu	tta Leu	acc Thr	2322

tgg	ata	ggg	ttg	aac	tca	aaa	aat	act	tct	atg	tca	ttt	tca	tgc	atc	2370
Trp	Ile	Gly	Leu	Asn	Ser	Lys	Asn	Thr	Ser	Met	Ser	Phe	Ser	Cys	Ile	
		745					750					755				
gcg	gcc	ggc	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	gcc	gat	2418
Ala	Ala	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln	Ala	Asp	
	760					765					770					
agt	ggt	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	ggc	agt	2466
Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys	Gly	Ser	
	775				780					785					790	
ggg	att	ttc	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	tac	aag	2514
Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln	Tyr	Lys	
				795				800						805		
ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	aaa	gcc	2562
Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ala	Ile	Gln	Lys	Ala	
			810					815					820			
cat	gaa	gag	gac	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	gag	aat	2610
His	Glu	Glu	Asp	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu	Glu	Asn	
	825						830					835				
ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	tca	gaa	2658
Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu	Ser	Glu	
	840					845					850					
aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	atc	atg	2706
Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	Ile	Met	
	855				860					865					870	
cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	aag	tat	2754
Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu	Lys	Tyr	
				875				880						885		
tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	tct	cat	2802
Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu	Ser	His	
			890					895					900			
aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	ccc	aac	2850
Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys	Pro	Asn	
			905				910					915				
aca	aat	aga	gct	tgg	aat	tcg	ttg	gaa	gtt	gaa	gac	tat	ggc	ttt	gga	2898
Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly	Phe	Gly	
	920					925						930				
gta	ttc	acc	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	cag	gat	gta	2946
Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln	Asp	Val	
	935				940					945					950	
ttc	tgc	gac	tca	aaa	ctc	atg	tca	gcg	gcc	ata	aaa	gac	aac	aga	gcc	2994
Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn	Arg	Ala	
				955				960						965		
gtc	cat	gcc	gat	atg	ggc	tat	tgg	ata	gaa	agt	gca	ctc	aat	gac	aca	3042
Val	His	Ala	Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn	Asp	Thr	
			970					975					980			
tgg	aag	ata	gag	aaa	gcc	tct	ttc	att	gaa	gtt	aaa	aac	tgc	cac	tgg	3090
Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys	His	Trp	
		985					990					995				
cca	aaa	tca	cac	acc	ctc	tgg	agc	aat	gga	gtg	cta	gaa	agt	gag	atg	3138
Pro	Lys	Ser	His	Thr	Leu	Trp	Ser	Asn	Gly	Val	Leu	Glu	Ser	Glu	Met	
	1000					1005					1010					
ata	att	cca	aag	aat	ctc	gct	gga	cca	gtg	tct	caa	cac	aac	tat	aga	3186
Ile	Ile	Pro	Lys	Asn	Leu	Ala	Gly	Pro	Val	Ser	Gln	His	Asn	Tyr	Arg	
	1015				1020					1025					1030	
cca	ggc	tac	cat	aca	caa	ata	aca	gga	cca	tgg	cat	cta	ggt	aag	ctt	3234
Pro	Gly	Tyr	His	Thr	Gln	Ile	Thr	Gly	Pro	Trp	His	Leu	Gly	Lys	Leu	
				1035				1040						1045		
gag	atg	gac	ttt	gat	ttc	tgt	gat	gga	aca	aca	gtg	gta	gtg	act	gag	3282
Glu	Met	Asp	Phe	Asp	Phe	Cys	Asp	Gly	Thr	Thr	Val	Val	Val	Thr	Glu	
			1050					1055					1060			
gac	tgc	gga	aat	aga	gga	ccc	tct	ttg	aga	aca	acc	act	gcc	tct	gga	3330
Asp	Cys	Gly	Asn	Arg	Gly	Pro	Ser	Leu	Arg	Thr	Thr	Thr	Ala	Ser	Gly	
		1065					1070					1075				
aaa	ctc	ata	aca	gaa	tgg	tgc	cga	tct	tgc	aca	tta	cca	ccg	cta		3378
Lys	Leu	Ile	Thr	Glu	Trp	Cys	Cys	Arg	Ser	Cys	Thr	Leu	Pro	Pro	Leu	
	1080					1085						1090				
aga	tac	aga	ggt	gag	gat	ggg	tgc	tgg	tac	ggg	atg	gaa	atc	aga	cca	3426
Arg	Tyr	Arg	Gly	Glu	Asp	Gly	Cys	Trp	Tyr	Gly	Met	Glu	Ile	Arg	Pro	
	1095				1100					1105					1110	

ttg aag gag aaa gaa gag aat ttg gtc aac tcc ttg gtc aca gct gga	3474
Leu Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly	
1115	
cat ggg cag gtc gac aac ttt tca cta gga gtc ttg gga atg gca ttg	3522
His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala Leu	
1130	
ttc ctg gag gaa atg ctt agg acc cga gta gga acg aaa cat gca ata	3570
Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His Ala Ile	
1145	
cta cta gtt gca gtt tct ttt gtg aca ttg atc aca ggg aac atg tcc	3618
Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly Asn Met Ser	
1160	
ttt aga gac ctg gga aga gtg atg gtt atg gta ggc gcc act atg acg	3666
Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly Ala Thr Met Thr	
1175	
gat gac ata ggt atg ggc gtg act tat ctt gcc cta cta gca gcc ttc	3714
Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala Leu Leu Ala Ala Phe	
1195	
aaa gtc aga cca act ttt gca gct gga cta ctc ttg aga aag ctg acc	3762
Lys Val Arg Pro Thr Phe Ala Ala Gly Leu Leu Leu Arg Lys Leu Thr	
1210	
tcc aag gaa ttg atg atg act act ata gga att gta ctc ctc tcc cag	3810
Ser Lys Glu Leu Met Met Thr Thr Ile Gly Ile Val Leu Leu Ser Gln	
1225	
agc acc ata cca gag acc att ctt gag ttg act gat gcg tta gcc tta	3858
Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu Thr Asp Ala Leu Ala Leu	
1240	
ggc atg atg gtc ctc aaa atg gtg aga aat atg gaa aag tat caa ttg	3906
Gly Met Met Val Leu Lys Met Val Arg Asn Met Glu Lys Tyr Gln Leu	
1255	
gca gtg act atc atg gct atc ttg tgc gtc cca aac gca gtg ata tta	3954
Ala Val Thr Ile Met Ala Ile Leu Cys Val Pro Asn Ala Val Ile Leu	
1275	
caa aac gca tgg aaa gtg agt tgc aca ata ttg gca gtg gtg tcc gtt	4002
Gln Asn Ala Trp Lys Val Ser Cys Thr Ile Leu Ala Val Val Ser Val	
1290	
tcc cca ctg ttc tta aca tcc tca cag caa aaa aca gat tgg ata cca	4050
Ser Pro Leu Phe Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro	
1305	
tta gca ttg acg atc aaa ggt ctc aat cca aca gct att ttt cta aca	4098
Leu Ala Leu Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr	
1320	
acc ctc tca aga acc agc aag aaa agg agc tgg cca tta aat gag gct	4146
Thr Leu Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala	
1335	
atc atg gca gtc ggg atg gtg agc att tta gcc agt tct ctc cta aaa	4194
Ile Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys	
1355	
aat gat att ccc atg aca gga cca tta gtg gct gga ggg ctc ctc act	4242
Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu Thr	
1370	
gtg tgc tac gtg ctc act gga cga tgc gcc gat ttg gaa ctg gag aga	4290
Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu Glu Arg	
1385	
gca gcc gat gtc aaa tgg gaa gac cag gca gag ata tca gga agc agt	4338
Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly Ser Ser	
1400	
cca atc ctg tca ata aca ata tca gaa gat ggt agc atg tcg ata aaa	4386
Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser Ile Lys	
1415	
aat gaa gag gaa gaa caa aca ctg acc ata ctc att aga aca gga ttg	4434
Asn Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg Thr Gly Leu	
1435	
ctg gtg atc tca gga ctt ttt cct gta tca ata cca atc acg gca gca	4482
Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro Ile Thr Ala Ala	
1450	
gca tgg tac ctg tgg gaa gtg aag aaa caa cgg gcc gga gta ttg tgg	4530
Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg Ala Gly Val Leu Trp	
1465	
1470	
1475	

gat gtt cct tca ccc cca ccc atg gga aag gct gaa ctg gaa gat gga	4578
Asp Val Pro Ser Pro Pro Pro Met Gly Lys Ala Glu Leu Glu Asp Gly	
1480 1485 1490	
gcc tat aga att aag caa aaa ggg att ctt gga tat tcc cag atc gga	4626
Ala Tyr Arg Ile Lys Gln Lys Gly Ile Leu Gly Tyr Ser Gln Ile Gly	
1495 1500 1505 1510	
gcc gga gtt tac aaa gaa gga aca ttc cat aca atg tgg cat gtc aca	4674
Ala Gly Val Tyr Lys Glu Gly Thr Phe His Thr Met Trp His Val Thr	
1515 1520 1525	
cgt ggc gct gtt cta atg cat aaa gga aag agg att gaa cca tca tgg	4722
Arg Gly Ala Val Leu Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp	
1530 1535 1540	
gcg gac gtc aag aaa gac cta ata tca tat gga gga ggc tgg aag tta	4770
Ala Asp Val Lys Lys Asp Leu Ile Ser Tyr Gly Gly Trp Lys Leu	
1545 1550 1555	
gaa gga gaa tgg aag gaa gga gaa gaa gtc cag gta ttg gca ctg gag	4818
Glu Gly Glu Trp Lys Glu Gly Glu Val Gln Val Leu Ala Leu Glu	
1560 1565 1570	
cct gga aaa aat cca aga gcc gtc caa acg aaa cct ggt ctt ttc aaa	4866
Pro Gly Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys	
1575 1580 1585 1590	
acc aac gcc gga aca ata ggt gct gta tct ctg gac ttt tct cct gga	4914
Thr Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly	
1595 1600 1605	
acg tca gga tct cca att atc gac aaa aaa gga aaa gtt gtg ggt ctt	4962
Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly Leu	
1610 1615 1620	
tat ggt aat ggt gtt gtt aca agg agt gga gca tat gtg agt gct ata	5010
Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser Ala Ile	
1625 1630 1635	
gcc cag act gaa aaa agc att gaa gac aac cca gag atc gaa gat gac	5058
Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile Glu Asp Asp	
1640 1645 1650	
att ttc cga aag aga aga ctg acc atc atg gac ctc cac cca gga gcg	5106
Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro Gly Ala	
1655 1660 1665 1670	
gga aag acg aag aga tac ctt ccg gcc ata gtc aga gaa gct ata aaa	5154
Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu Ala Ile Lys	
1675 1680 1685	
cgg ggt ttg aga aca tta atc ttg gcc ccc act aga gtt gtg gca gct	5202
Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg Val Val Ala Ala	
1690 1695 1700	
gaa atg gag gaa gcc ctt aga gga ctt cca ata aga tac cag acc cca	5250
Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile Arg Tyr Gln Thr Pro	
1705 1710 1715	
gcc atc aga gct gtg cac acc ggg cgg gag att gtg gac cta atg tgt	5298
Ala Ile Arg Ala Val His Thr Gly Arg Glu Ile Val Asp Leu Met Cys	
1720 1725 1730	
cat gcc aca ttt acc atg agg ctg cta tca cca gtt aga gtg cca aac	5346
His Ala Thr Phe Thr Met Arg Leu Leu Ser Pro Val Arg Val Pro Asn	
1735 1740 1745 1750	
tac aac ctg att atc atg gac gaa gcc cat ttc aca gac cca gca agt	5394
Tyr Asn Leu Ile Ile Met Asp Glu Ala His Phe Thr Asp Pro Ala Ser	
1755 1760 1765	
ata gca gct aga gga tac atc tca act cga gtg gag atg ggt gag gca	5442
Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala	
1770 1775 1780	
gct ggg att ttt atg aca gcc act ccc ccg gga agc aga gac cca ttt	5490
Ala Gly Ile Phe Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe	
1785 1790 1795	
cct cag agc aat gca cca atc ata gat gaa gaa aga gaa atc cct gaa	5538
Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu Glu Glu Ile Pro Glu	
1800 1805 1810	
cgc tcg tgg aat tcc gga cat gaa tgg gtc acg gat ttt aaa ggg aag	5586
Arg Ser Trp Asn Ser Gln His Glu Trp Val Thr Asp Phe Lys Gly Lys	
1815 1820 1825 1830	
act gtt tgg ttc gtt cca agt ata aaa gca gga aat gat ata gca gct	5634
Thr Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala	
1835 1840 1845	

tgc	ctg	agg	aaa	aat	gga	aag	aaa	gtg	ata	caa	ctc	agt	agg	aag	acc	5682
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys	Thr	
1850	1850	1850	1850	1850	1850	1850	1850	1855	1855	1860	1860	1860	1860	1860	1860	
ttt	gat	tct	gag	tat	gtc	aag	act	aga	acc	aat	gat	tgg	gac	ttc	gtg	5730
Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	Phe	Val	
1865	1865	1865	1865	1865	1865	1865	1870	1870	1870	1875	1875	1875	1875	1875	1875	
gtt	aca	act	gac	att	tca	gaa	atg	ggt	gcc	aat	ttc	aag	gct	gag	agg	5778
Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	Glu	Arg	
1880	1880	1880	1880	1880	1885	1885	1885	1885	1885	1890	1890	1890	1890	1890	1890	
gtt	ata	gac	ccc	aga	cgc	tgc	atg	aaa	cca	gtc	ata	cta	aca	gat	ggt	5826
Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr	Asp	Gly	
1895	1895	1895	1895	1895	1900	1900	1900	1900	1905	1905	1905	1905	1905	1910	1910	
gaa	gag	cgg	gtg	att	ctg	gca	gga	cct	atg	cca	gtg	acc	cac	tct	agt	5874
Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr	His	Ser	Ser	
1915	1915	1915	1915	1915	1915	1915	1915	1920	1920	1920	1920	1925	1925	1925	1925	
gca	gca	caa	aga	agg	aga	ata	gga	aga	aat	cca	aaa	aat	gag	aat		5922
Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro	Lys	Asn	Glu	Asn	
1930	1930	1930	1930	1930	1935	1935	1935	1935	1940	1940	1940	1940	1940	1940	1940	
gac	cag	tac	ata	tac	atg	ggg	gaa	cct	ctg	gaa	aat	gat	gaa	gac	tgt	5970
Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu	Asn	Asp	Glu	Asp	Cys	
1945	1945	1945	1945	1945	1950	1950	1950	1950	1955	1955	1955	1955	1955	1955	1955	
gca	cac	tgg	aaa	gaa	gct	aaa	atg	ctc	cta	gat	aac	atc	aac	acg	cca	6018
Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	
1960	1960	1960	1960	1960	1965	1965	1965	1965	1970	1970	1970	1970	1970	1970	1970	
gaa	gga	atc	att	cct	agc	atg	ttc	gaa	cca	gag	cgt	gaa	aag	gtg	gat	6066
Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu	Pro	Glu	Arg	Glu	Lys	Val	Asp	
1975	1975	1975	1975	1980	1980	1980	1980	1985	1985	1985	1985	1985	1985	1990	1990	
gcc	att	gat	ggc	gaa	tac	cgc	ttg	aga	gga	gaa	gca	agg	aaa	acc	ttt	6114
Ala	Ile	Asp	Gly	Glu	Tyr	Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	
1995	1995	1995	1995	1995	1995	1995	2000	2000	2000	2000	2000	2000	2005	2005	2005	
gta	gac	tta	atg	aga	aga	gga	gac	cta	cca	gtc	tgg	ttg	gcc	tac	aga	6162
Val	Asp	Leu	Met	Arg	Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	
2010	2010	2010	2010	2010	2010	2010	2015	2015	2015	2015	2020	2020	2020	2020	2020	
gtg	gca	gct	gaa	ggc	atc	aac	tac	gca	gac	aga	agg	tgg	tgt	ttt	gat	6210
Val	Ala	Ala	Glu	Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	Phe	Asp	
2025	2025	2025	2025	2030	2030	2030	2030	2030	2035	2035	2035	2035	2035	2035	2035	
gga	gtc	aag	aac	aac	caa	atc	cta	gaa	gaa	aac	gtg	gaa	gtt	gaa	atc	6258
Gly	Val	Lys	Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	Glu	Ile	
2040	2040	2040	2040	2045	2045	2045	2045	2050	2050	2050	2050	2050	2050	2050	2050	
tgg	aca	aaa	gaa	ggg	gaa	agg	aag	aaa	ttg	aaa	ccc	aga	tgg	ttg	gat	6306
Trp	Thr	Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	Leu	Asp	
2055	2055	2055	2055	2060	2060	2060	2060	2065	2065	2065	2065	2065	2065	2070	2070	
gct	agg	atc	tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	gaa	ttt	6354
Ala	Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	Glu	Phe	
2075	2075	2075	2075	2075	2075	2075	2080	2080	2080	2080	2080	2085	2085	2085	2085	
gca	gcc	gga	aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	atg	ggt	6402
Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	Met	Gly	
2090	2090	2090	2090	2095	2095	2095	2095	2095	2095	2100	2100	2100	2100	2100	2100	
agg	ctc	cca	acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	gac	aac	6450
Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu	Asp	Asn	
2105	2105	2105	2105	2110	2110	2110	2110	2110	2115	2115	2115	2115	2115	2115	2115	
tta	gca	gtg	ctg	cac	acg	gct	gag	gca	ggt	gga	agg	gcg	tac	aac	cat	6498
Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr	Asn	His	
2120	2120	2120	2120	2125	2125	2125	2125	2130	2130	2130	2130	2130	2130	2130	2130	
gct	ctc	agt	gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	ctg	aca	6546
Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu	Leu	Thr	
2135	2135	2135	2135	2140	2140	2140	2140	2145	2145	2145	2145	2145	2145	2150	2150	
ctt	ctg	gct	aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	agc	gca	6594
Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met	Ser	Ala	
2155	2155	2155	2155	2155	2155	2155	2160	2160	2160	2160	2160	2160	2160	2165	2165	
agg	ggc	ata	ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	acg	gct	6642
Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile	Thr	Ala	
2170	2170	2170	2170	2175	2175	2175	2175	2175	2175	2180	2180	2180	2180	2180	2180	
agc	atc	ctc	cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	gca	gct	6690
Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile	Ala	Ala	
2185	2185	2185	2185	2190	2190	2190	2190	2195	2195	2195	2195	2195	2195	2195	2195	
tca	ata	ata	ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	gaa	cct	6738
Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro	Glu	Pro	
2200	2200	2200	2200	2205	2205	2205	2205	2210	2210	2210	2210	2210	2210	2210	2210	

gaa aaa cag aga aca ccc caa gac aac caa ctg acc tac gtt gtc ata Glu Lys Gln Arg Thr Pro Gln Asp Asn Gln Leu Thr Tyr Val Val Ile 2215 2220 2225 2230	6786
gcc atc ctc aca gtg gtg gcc gca acc atg gca aac gag atg ggt ttc Ala Ile Leu Thr Val Val Ala Ala Thr Met Ala Asn Glu Met Gly Phe 2235 2240 2245	6834
cta gaa aaa acg aag aaa gat ctc gga ttg gga agc att gca acc cag Leu Glu Lys Thr Lys Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln 2250 2255 2260	6882
caa ccc gag agc aac atc ctg gac ata gat cta cgt cct gca tca gca Gln Pro Glu Ser Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala 2265 2270 2275	6930
tgg acg ctg tat gcc gtg gcc aca aca ttt gtt aca cca atg ttg aga Trp Thr Leu Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg 2280 2285 2290	6978
cat agc att gaa aat tcc tca gtg aat gtg tcc cta aca gct ata gcc His Ser Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala 2295 2300 2305 2310	7026
aac caa gcc aca gtg tta atg ggt ctc ggg aaa gga tgg cca ttg tca Asn Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser 2315 2320 2325	7074
aag atg gac atc gga gtt ccc ctt ctc gcc att gga tgc tac tca caa Lys Met Asp Ile Gly Val Pro Leu Ala Ile Gly Cys Tyr Ser Gln 2330 2335 2340	7122
gtc aac ccc ata act ctc aca gca gct ctt ttc tta ttg gta gca cat Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Val Ala His 2345 2350 2355	7170
tat gcc atc ata ggg cca gga ctc caa gca aaa gca acc aga gaa gct Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr Arg Glu Ala 2360 2365 2370	7218
cag aaa aga gca gcg gcg gcc atc atg aaa aac cca act gtc gat gga Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val Asp Gly 2375 2380 2385 2390	7266
ata aca gtg att gac cta gat cca ata cct tat gat cca aag ttt gaa Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro Lys Phe Glu 2395 2400 2405	7314
aag cag ttg gga caa gta atg ctc cta gtc ctc tgc gtg act caa gta Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys Val Thr Gln Val 2410 2415 2420	7362
ttg atg atg agg act aca tgg gct ctg tgt gag gct tta acc tta gct Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu Ala Leu Thr Leu Ala 2425 2430 2435	7410
acc ggg ccc atc tcc aca ttg tgg gaa gga aat cca ggg agg ttt tgg Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly Asn Pro Gly Arg Phe Trp 2440 2445 2450	7458
aac act acc att gcg gtg tca atg gct aac att ttt aga ggg agt tac Asn Thr Thr Ile Ala Val Ser Met Ala Asn Ile Phe Arg Gly Ser Tyr 2455 2460 2465 2470	7506
ttg gcc gga gct gga ctt ctc ttt tct att atg aag aac aca acc aac Leu Ala Gly Ala Gly Leu Leu Phe Ser Ile Met Lys Asn Thr Thr Asn 2475 2480 2485	7554
aca aga agg gga act gcc aac ata gga gag acg ctt gga gag aaa tgg Thr Arg Arg Gly Thr Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp 2490 2495 2500	7602
aaa agc cga ttg aac gca ttg gga aaa agt gaa ttc cag atc tac aag Lys Ser Arg Leu Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys 2505 2510 2515	7650
aaa agt gga atc cag gaa gtg gat aga acc tta gca aaa gaa gcc att Lys Ser Gly Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile 2520 2525 2530	7698
aaa aga gga gaa acg gac cat cac gct gtg tgc cga gcc tca gca aaa Lys Arg Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys 2535 2540 2545 2550	7746
ctg aga tgg ttc gtt gag aga aac atg gtc aca cca gaa ggg aaa gta Leu Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val 2555 2560 2565	7794
gtg gac ctc ggt tgt gcc aga gga gcc tca tac tat tgt gga gga Val Asp Leu Gly Cys Gly Arg Gly Trp Ser Tyr Tyr Cys Gly Gly 2570 2575 2580	7842

cta aag aat gta aga gaa gtc aaa ggc cta aca aaa gga gga cca gga	7890
Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly	
2585	
cac gaa gaa ccc atc ccc atg tca aca tat ggg tgg aat cta gtg cgt	7938
His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val Arg	
2600	
ctt caa agt gga gtt gac gtt ttc ttc atc ccg cca gaa aag tgt gac	7986
Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys Cys Asp	
2615	
aca tta ttg tgt gac ata ggg gag tca tca cca aat ccc aca gtg gaa	8034
Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro Thr Val Glu	
2635	
gca gga cga aca ctc aga gtc ctt aac tta gta gaa aat tgg ttg aac	8082
Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu Asn Trp Leu Asn	
2650	
aac aac act caa ttt tgc ata aag gtt ctc aac cca tat atg ccc tca	8130
Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn Pro Tyr Met Pro Ser	
2665	
gtc ata gaa aaa atg gaa gca cta caa agg aaa tat gga gga gcc tta	8178
Val Ile Glu Lys Met Glu Ala Leu Gln Arg Lys Tyr Gly Gly Ala Leu	
2680	
gtg agg aat cca ctc tca cga aac tcc aca cat gag atg tac tgg gta	8226
Val Arg Asn Pro Leu Ser Arg Asn Ser Thr His Glu Met Tyr Trp Val	
2695	
tcc aat gct tcc ggg aac ata gtg tca tca gtg aac atg att tca agg	8274
Ser Asn Ala Ser Gly Asn Ile Val Ser Ser Val Asn Met Ile Ser Arg	
2715	
atg ttg atc aac aga ttt aca atg aga tac aag aaa gcc act tac gag	8322
Met Leu Ile Asn Arg Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu	
2730	
ccg gat gtt gac ctc gga agc gga acc cgt aac atc ggg att gaa agt	8370
Pro Asp Val Asp Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser	
2745	
gag ata cca aac cta gat ata att ggg aaa aga ata gaa aaa ata aag	8418
Glu Ile Pro Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys	
2760	
caa gag cat gaa aca tca tgg cac tat gac caa gac cac cca tac aaa	8466
Gln Glu His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys	
2775	
acg tgg gca tac cat ggt agc tat gaa aca aaa cag act gga tca gca	8514
Thr Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala	
2795	
tca tcc atg gtc aac gga gtg gtc agg ctg ctg aca aaa cct tgg gac	8562
Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp Asp	
2810	
gtc gtc ccc atg gtg aca cag atg gca atg aca gac acg act cca ttt	8610
Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr Pro Phe	
2825	
gga caa cag cgc gtt ttt aaa gag aaa gtg gac acg aga acc caa gaa	8658
Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg Thr Gln Glu	
2840	
ccg aaa gaa ggc acg aag aaa cta atg aaa ata aca gca gag tgg ctt	8706
Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu Trp Leu	
2855	
tgg aaa gaa tta ggg aag aaa aag aca ccc agg atg tgc acc aga gaa	8754
Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys Thr Arg Glu	
2875	
gaa ttc aca aga aag gtg aga agc aat gca gcc ttg ggg gcc ata ttc	8802
Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Leu Gly Ala Ile Phe	
2890	
act gat gag aac aag tgg aag tgc gca cgt gag gct gtt gaa gat agt	8850
Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu Ala Val Glu Asp Ser	
2905	
agg ttt tgg gag ctg gtt gac aag gaa agg aat ctc cat ctt gaa gga	8898
Arg Phe Trp Glu Leu Val Asp Lys Glu Arg Asn Leu His Leu Glu Gly	
2920	
aag tgt gaa aca tgt gtg tac aac atg atg gga aaa aga gag aag aag	8946
Lys Cys Glu Thr Cys Val Tyr Asn Met Met Gly Lys Arg Glu Lys Lys	
2935	
2940	
2945	
2950	

cta ggg gaa ttc ggc aag gca aaa ggc agc aga gcc ata tgg tac atg	8994
Leu Gly Glu Phe Gly Lys Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met	
2955	
tgg ctt gga gca cgc ttc tta gag ttt gaa gcc cta gga ttc tta aat	9042
Trp Leu Gly Ala Arg Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn	
2970	
gaa gat cac tgg ttc tcc aga gag aac tcc ctg agt gga gtg gaa gga	9090
Glu Asp His Trp Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly	
2985	
gaa ggg ctg cac aag cta ggt tac att cta aga gac gtg agc aag aaa	9138
Glu Gly Leu His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys	
3000	
gag gga gga gca atg tat gcc gat gac acc gca gga tgg gat aca aga	9186
Glu Gly Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg	
3015	
atc aca cta gaa gac cta aaa aat gaa gaa atg gta aca aac cac atg	9234
Ile Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met	
3035	
gaa gga gaa cac aag aaa cta gcc gag gcc att ttc aaa cta acg tac	9282
Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr Tyr	
3050	
caa aac aag gtg gtg cgt gtg caa aga cca aca cca aga ggc aca gta	9330
Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly Thr Val	
3065	
atg gac atc ata tcg aga aga gac caa aga ggt agt gga caa gtt ggc	9378
Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln Val Gly	
3080	
acc tat gga ctc aat act ttc acc aat atg gaa gcc caa cta atc aga	9426
Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu Ile Arg	
3095	
cag atg gag gga gaa gga gtc ttt aaa agc att cag cac cta aca atc	9474
Gln Met Glu Gly Val Phe Lys Ser Ile Gln His Leu Thr Ile	
3115	
aca gaa gaa atc gct gtg caa aac tgg tta gca aga gtg ggg cgc gaa	9522
Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg Val Gly Arg Glu	
3130	
agg tta tca aga atg gcc atc agt gga gat gat tgt gtt gtg aaa cct	9570
Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp Cys Val Val Lys Pro	
3145	
tta gat gac agg ttc gca agc gct tta aca gct cta aat gac atg gga	9618
Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr Ala Leu Asn Asp Met Gly	
3160	
aag att agg aaa gac ata caa caa tgg gaa cct tca aga gga tgg aat	9666
Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg Gly Trp Asn	
3175	
gat tgg aca caa gtg ccc ttc tgt tca cac cat ttc cat gag tta atc	9714
Asp Trp Thr Gln Val Pro Phe Cys Ser His His Phe His Glu Leu Ile	
3195	
atg aaa gac ggt cgc gta ctc gtt gtt cca tgt aga aac caa gat gaa	9762
Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu	
3210	
ctg att ggc aga gcc cga atc tcc caa gga gca ggg tgg tct ttg cgg	9810
Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg	
3225	
gag acg gcc tgt ttg ggg aag tct tac gcc caa atg tgg agc ttg atg	9858
Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met	
3240	
tac ttc cac aga cgc gac ctc agg ctg gcg gca aat gct att tgc tcg	9906
Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Asn Ala Ile Cys Ser	
3255	
gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg tcc ata	9954
Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile	
3275	
cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca gtc tgg	10002
His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp	
3290	
aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa act cca	10050
Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro	
3305	
3310	
3315	

```

gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa gac caa 10098
Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp Gln
3320 3325 3330
tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg gca aag 10146
Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp Ala Lys
3335 3340 3345 3350
aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc aat gaa 10194
Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly Asn Glu
3355 3360 3365
gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga gaa gag 10242
Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg Glu Glu
3370 3375 3380
gaa gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa caaggctaga 10296
Glu Glu Ala Gly Val Leu Trp *
3385
agtcagggtcg gattaagcca tagtacggaa aaaactatgc tacctgtgag ccccggtccaa 10356
ggacgttaaaa agaagtcagg ccatcataaa tgccatagct tgagtaaact atgcagcctg 10416
tagctccacc tgagaagggtg taataaatcc gggaggccac aaaccatgga agctgtacgc 10476
atggcgtagt ggactagcgg ttagaggaga cccctccctt acaaatcgca gcaacaatgg 10536
gggccaagg cgagatgaag ctgtagtctc gctggaagga ctagagggtta gaggagaccc 10596
ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg ctgtctcctc 10656
agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat caacagggtc 10716
t
10717

```

<210> SEQ ID NO 10

<211> COMPRIMENTO: 3389

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência
artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 10

Met 1	Asn	Asn	Gln	Arg	Lys	Lys	Ala	Lys	Asn	Thr	Pro	Phe	Asn	Met	Leu
Lys	Arg	Glu	Arg	Asn	Arg	Val	Ser	Thr	Val	Gln	Gln	Leu	Thr	Lys	Arg
Phe	Ser	Leu	Gly	Met	Leu	Gln	Gly	Arg	Gly	Pro	Leu	Lys	Leu	Phe	Met
Ala	Leu	Val	Ala	Phe	Leu	Arg	Phe	Leu	Thr	Ile	Pro	Pro	Thr	Ala	Gly
Ile	Leu	Lys	Arg	Trp	Gly	Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val
Leu	Arg	Gly	Phe	Arg	Lys	Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn
Arg	Arg	Arg	Arg	Ser	Ala	Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val
Met	Ala	Phe	His	Leu	Thr	Thr	Arg	Asp	Gly	Glu	Pro	Arg	Met	Ile	Val
Gly	Lys	Asn	Glu	Arg	Gly	Lys	Ser	Leu	Leu	Phe	Lys	Thr	Ala	Ser	Gly
Ile	Asn	Met	Cys	Thr	Leu	Ile	Ala	Met	Asp	Leu	Gly	Glu	Met	Cys	Asp
Asp	Thr	Val	Thr	Tyr	Lys	Cys	Pro	His	Ile	Thr	Glu	Val	Glu	Pro	Glu
Asp	Ile	Asp	Cys	Trp	Cys	Asn	Leu	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly
Thr	Cys	Asn	Gln	Ala	Gly	Glu	His	Arg	Arg	Asp	Lys	Arg	Ser	Val	Ala
Leu	Ala	Pro	His	Val	Gly	Met	Gly	Leu	Asp	Thr	Arg	Thr	Gln	Thr	Trp
Met	Ser	Ala	Glu	Gly	Ala	Trp	Arg	Gln	Val	Glu	Lys	Val	Glu	Thr	Trp
Ala	Leu	Arg	His	Pro	Gly	Phe	Thr	Ile	Leu	Ala	Leu	Phe	Leu	Ala	His
Tyr	Ile	Gly	Thr	Ser	Leu	Thr	Gln	Lys	Val	Val	Ile	Phe	Ile	Leu	Leu
Met	Leu	Val	Thr	Pro	Ser	Met	Thr	Met	Arg	Cys	Val	Gly	Val	Gly	Asn
Arg	Asp	Phe	Val	Glu	Gly	Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val

—

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
 850 855 860
 Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
 865 870 875 880
 Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
 885 890 895
 Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
 900 905 910
 Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
 915 920 925
 Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
 930 935 940
 Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
 945 950 955 960
 Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
 965 970 975
 Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
 980 985 990
 Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
 995 1000 1005
 Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro Val
 1010 1015 1020
 Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr Gly Pro
 1025 1030 1035 1040
 Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp Gly Thr
 1045 1050 1055
 Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser Leu Arg
 1060 1065 1070
 Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys Arg Ser
 1075 1080 1085
 Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys Trp Tyr
 1090 1095 1100
 Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu Val Asn
 1105 1110 1115 1120
 Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser Leu Gly
 1125 1130 1135
 Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val
 1140 1145 1150
 Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu
 1155 1160 1165
 Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met
 1170 1175 1180
 Val Gly Ala Thr Met Thr Asp Ile Gly Met Gly Val Thr Tyr Leu
 1185 1190 1195 1200
 Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu
 1205 1210 1215
 Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile Gly
 1220 1225 1230
 Ile Val Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu
 1235 1240 1245
 Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val Arg Asn
 1250 1255 1260
 Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu Cys Val
 1265 1270 1275 1280
 Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys Thr Ile
 1285 1290 1295
 Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser Gln Gln
 1300 1305 1310
 Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu Asn Pro
 1315 1320 1325
 Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys Arg Ser
 1330 1335 1340
 Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser Ile Leu
 1345 1350 1355 1360
 Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro Leu Val
 1365 1370 1375
 Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala
 1380 1385 1390
 Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala

1395 1400 1405
 Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp
 1410 1415 1420
 Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile
 1425 1430 1435 1440
 Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
 1445 1450 1455
 Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln
 1460 1465 1470
 Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys
 1475 1480 1485
 Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile Leu
 1490 1495 1500
 Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr Phe His
 1505 1510 1515 1520
 Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys Gly Lys
 1525 1530 1535
 Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile Ser Tyr
 1540 1545 1550
 Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu Val
 1555 1560 1565
 Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val Gln Thr
 1570 1575 1580
 Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala Val Ser
 1585 1590 1595 1600
 Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys
 1605 1610 1615
 Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly
 1620 1625 1630
 Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn
 1635 1640 1645
 Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met
 1650 1655 1660
 Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile
 1665 1670 1675 1680
 Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro
 1685 1690 1695
 Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro
 1700 1705 1710
 Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu
 1715 1720 1725
 Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu Leu Ser
 1730 1735 1740
 Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu Ala His
 1745 1750 1755 1760
 Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg
 1765 1770 1775
 Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr Pro Pro
 1780 1785 1790
 Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu
 1795 1800 1805
 Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu Trp Val
 1810 1815 1820
 Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile Lys Ala
 1825 1830 1835 1840
 Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys Val Ile
 1845 1850 1855
 Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr
 1860 1865 1870
 Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala
 1875 1880 1885
 Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro
 1890 1895 1900
 Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met
 1905 1910 1915 1920
 Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
 1925 1930 1935
 Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu
 1940 1945 1950

Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu
 1955 1960 1965
 Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu Pro
 1970 1975 1980
 Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu Arg Gly
 1985 1990 1995 2000
 Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp Leu Pro
 2005 2010 2015
 Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr Ala Asp
 2020 2025 2030
 Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu Glu Glu
 2035 2040 2045
 Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys Lys Leu
 2050 2055 2060
 Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu Ala Leu
 2065 2070 2075 2080
 Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn
 2085 2090 2095
 Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala
 2100 2105 2110
 Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly
 2115 2120 2125
 Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu
 2130 2135 2140
 Thr Leu Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe
 2145 2150 2155 2160
 Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
 2165 2170 2175
 Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln
 2180 2185 2190
 Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val
 2195 2200 2205
 Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn Gln
 2210 2215 2220
 Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala Thr Met
 2225 2230 2235 2240
 Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu Gly Leu
 2245 2250 2255
 Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp Ile Asp
 2260 2265 2270
 Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr Thr Phe
 2275 2280 2285
 Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val Asn Val
 2290 2295 2300
 Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly Leu Gly
 2305 2310 2315 2320
 Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu Leu Ala
 2325 2330 2335
 Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu
 2340 2345 2350
 Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala
 2355 2360 2365
 Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys
 2370 2375 2380
 Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro
 2385 2390 2395 2400
 Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
 2405 2410 2415
 Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys
 2420 2425 2430
 Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly
 2435 2440 2445
 Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala Asn
 2450 2455 2460
 Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe Ser Ile
 2465 2470 2475 2480
 Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile Gly Glu
 2485 2490 2495
 Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly Lys Ser

Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp Arg Thr
 2500 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055

Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His Ala Val
 Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn Met Val
 Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly Gly Trp
 Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu
 Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr
 Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile
 Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser
 Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
 Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu
 Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg
 Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser Thr
 His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val Ser Ser
 Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met Arg Tyr
 Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly Thr Arg
 Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile Gly Lys
 Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His Tyr Asp
 Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr Glu Thr
 Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val Arg Leu
 Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met Ala Met
 Thr Asp Thr Thr Pro Phe Gly Gln Arg Val Phe Lys Glu Lys Val
 Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys
 Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro
 Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
 Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg
 Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg
 Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met Met
 Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys Gly Ser
 Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu Phe Glu
 Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu Asn Ser
 Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr Ile Leu
 Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp Asp Thr
 Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn Glu Glu
 Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala Glu Ala

Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro
 3060 3065 3070
 Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg
 3075 3080 3085
 Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met
 3090 3095 3100
 Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser
 3105 3110 3115 3120
 Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
 3125 3130 3135
 Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp
 3140 3145 3150
 Asp Cys Val Val Lys Pro Leu Asp Arg Phe Ala Ser Ala Leu Thr
 3155 3160 3165
 Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu
 3170 3175 3180
 Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys Ser His
 3185 3190 3195 3200
 His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro
 3205 3210 3215
 Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly
 3220 3225 3230
 Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala
 3235 3240 3245
 Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala
 3250 3255 3260
 Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser
 3265 3270 3275 3280
 Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu
 3285 3290 3295
 Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp
 3300 3305 3310
 Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu
 3315 3320 3325
 Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser
 3330 3335 3340
 Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg
 3345 3350 3355 3360
 Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
 3365 3370 3375
 Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385

<210> SEQ ID NO 11

<211> COMPRIMENTO: 10723

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<400> SEQUÊNCIA 11

148

55	aca	att	aaa	aaa	tca	aaa	gct	att	aat	gtt	ttg	aga	ggg	ttc	agg	aaa	354
	Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val	Leu	Arg	Gly	Phe	Arg	Lys	
					75					80					85		
	gag	att	gga	agg	atg	ctg	aac	atc	ttg	aat	agg	aga	cgc	agc	tct	gca	402
	Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn	Arg	Arg	Arg	Ser	Ser	Ala	
				90					95					100			
	ggc	atg	atc	att	atg	ctg	att	cca	aca	gtg	atg	gcg	ttc	cat	tta	acc	450
	Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val	Met	Ala	Phe	His	Leu	Thr	
				105				110					115				
	acg	cgt	gat	ggc	gaa	ccc	ctc	atg	ata	gtg	gca	aaa	cat	gaa	agg	ggg	498
	Thr	Arg	Asp	Gly	Glu	Pro	Leu	Met	Ile	Val	Ala	Lys	His	Glu	Arg	Gly	
		120					125					130					
	aga	cct	ctc	ttg	ttt	aag	aca	gag	ggg	atc	aac	aaa	tgc	act	ctc		546
	Arg	Pro	Leu	Leu	Phe	Lys	Thr	Thr	Glu	Gly	Ile	Asn	Lys	Cys	Thr	Leu	
	135					140				145					150		
	att	gcc	atg	gac	ttg	ggt	gaa	atg	tgt	gag	gac	act	gtc	acg	tat	aaa	594
	Ile	Ala	Met	Asp	Leu	Gly	Glu	Met	Cys	Glu	Asp	Thr	Val	Thr	Tyr	Lys	
					155					160					165		
	tgc	ccc	tta	ctg	gtc	aat	acc	gaa	cct	gaa	gac	att	gat	tgc	tgg	tgc	642
	Cys	Pro	Leu	Val	Asn	Thr	Glu	Glu	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	
				170				175						180			
	aat	ctc	acg	tct	acc	tgg	gtc	atg	tat	ggg	aca	tgc	acc	cag	agc	gga	690
	Asn	Leu	Thr	Ser	Thr	Trp	Val	Met	Tyr	Gly	Thr	Cys	Thr	Gln	Ser	Gly	
				185				190					195				
	gaa	cgg	aga	cga	gag	aag	cgc	tca	gta	gct	tta	aca	cca	cat	tca	gga	738
	Glu	Arg	Arg	Arg	Glu	Lys	Arg	Ser	Val	Ala	Leu	Thr	Pro	His	Ser	Gly	
		200					205					210					
	atg	gga	ttg	gaa	aca	aga	gct	gag	aca	tgg	atg	tca	tcg	gaa	ggg	gct	786
	Met	Gly	Leu	Glu	Thr	Arg	Ala	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	
	215					220					225				230		
	tgg	aag	cat	gct	cag	aga	gta	gag	agc	tgg	ata	ctc	aga	aac	cca	gga	834
	Trp	Lys	His	Ala	Gln	Arg	Val	Glu	Ser	Trp	Ile	Leu	Arg	Asn	Pro	Gly	
				235						240					245		
	ttc	gcg	ctc	ttg	gca	gga	ttt	atg	gct	tat	atg	att	ggg	caa	aca	gga	882
	Phe	Ala	Leu	Leu	Ala	Gly	Phe	Met	Ala	Tyr	Met	Ile	Gly	Gln	Thr	Gly	
				250					255					260			
	atc	cag	cga	act	gtc	ttc	ttt	gtc	cta	atg	atg	ctg	gtc	gcc	cca	tcc	930
	Ile	Gln	Arg	Thr	Val	Phe	Phe	Val	Leu	Met	Met	Leu	Val	Ala	Pro	Ser	
				265				270					275				
	tac	gga	atg	cga	tgc	gta	gga	gta	gga	aac	aga	gac	ttt	gtg	gaa	gga	978
	Tyr	Gly	Met	Arg	Cys	Val	Gly	Val	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	
		280					285					290					
	gtc	tca	ggg	gga	gca	tgg	gtc	gat	ctg	gtg	cta	gaa	cat	gga	gga	tgc	1026
	Val	Ser	Gly	Gly	Ala	Trp	Val	Asp	Leu	Val	Leu	Glu	His	Gly	Gly	Cys	
	295					300					305				310		
	gtc	aca	acc	atg	gcc	cag	gga	aaa	cca	acc	ttg	gat	ttt	gaa	ctg	act	1074
	Val	Thr	Thr	Met	Ala	Gln	Gly	Lys	Pro	Thr	Leu	Asp	Phe	Glu	Leu	Thr	
				315						320					325		
	aag	aca	aca	gcc	aag	gaa	gtg	gct	ctg	tta	aga	acc	tat	tgc	att	gaa	1122
	Lys	Thr	Thr	Ala	Lys	Glu	Val	Ala	Leu	Leu	Arg	Thr	Tyr	Cys	Ile	Glu	
				330				335						340			
	gcc	tca	ata	tca	aac	ata	acc	acg	gca	aca	aga	tgt	cca	acg	caa	gga	1170
	Ala	Ser	Ile	Ser	Asn	Ile	Thr	Thr	Ala	Thr	Arg	Cys	Pro	Thr	Gln	Gly	
				345				350					355				
	gag	cct	tat	cta	aaa	gag	gaa	caa	gac	caa	cag	tac	att	tgc	cgg	aga	1218
	Glu	Pro	Tyr	Leu	Lys	Glu	Glu	Gln	Asp	Gln	Gln	Tyr	Ile	Cys	Arg	Arg	
		360					365					370					
	gat	gtg	gta	gac	aga	ggg	tgg	ggc	aat	ggc	tgt	ggc	ttg	ttt	gga	aaa	1266
	Asp	Val	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
						380					385				390		
	gga	gga	gtt	gtg	aca	tgt	gcg	aag	ttt	tca	tgt	tcg	ggg	aag	ata	aca	1314
	Gly	Gly	Val	Val	Thr	Cys	Ala	Lys	Phe	Ser	Cys	Ser	Gly	Lys	Ile	Thr	
					395					400					405		
	ggc	aat	ttg	gtc	caa	att	gag	aac	ctt	gaa	tac	aca	gtg	gtt	gta	aca	1362
	Gly	Asn	Leu	Val	Gln	Ile	Glu	Asn	Leu	Glu	Tyr	Thr	Val	Val	Val	Thr	
				410					415					420			
	gtc	cac	aat	gga	gac	acc	cat	gca	gta	gga	aat	gac	acg	aat	cat		1410
	Val	His	Asn	Gly	Asp	Thr	His	Ala	Val	Gly	Asn	Asp	Thr	Ser	Asn	His	

1160	atg tcc ttt aga gac ctg gga aga gtg atg gtt atg gta ggc gcc act	1165	Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly Ala Thr	1170	3666
1175	atg acg gat gac ata ggt atg ggc gtg act tat ctt gcc cta cta gca	1180	Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala Leu Leu Ala	1185	3714
1195	gcc ttc aaa gtc aga cca act ttt gca gct gga cta ctc ttg aga aag	1200	Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu Leu Leu Arg Lys	1205	3762
1210	ctg acc tcc aag gaa ttg atg atg act act ata gga att gta ctc ctc	1215	Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile Gly Ile Val Leu Leu	1220	3810
1225	tcc cag agc acc ata cca gag acc att ctt gag ttg act gat gcg tta	1230	Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu Thr Asp Ala Leu	1235	3858
1240	gcc tta ggc atg atg gtc ctc aaa atg gtg aga aat atg gaa aag tat	1245	Ala Leu Gly Met Met Val Leu Lys Met Val Arg Asn Met Glu Lys Tyr	1250	3906
1255	caa ttg gca gtg act atc atg gct atc ttg tgc gtc cca aac gca gtg	1260	Gln Leu Ala Val Thr Ile Met Ala Ile Leu Cys Val Pro Asn Ala Val	1265	3954
1275	ata tta caa aac gca tgg aaa gtg agt tgc aca ata ttg gca gtg gtg	1280	Ile Leu Gln Asn Ala Trp Lys Val Ser Cys Thr Ile Leu Ala Val Val	1285	4002
1290	tcc gtt tcc cca ctg ttc tta aca tcc tca cag caa aaa aca gat tgg	1295	Ser Val Ser Pro Leu Phe Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp	1300	4050
1305	ata cca tta gca ttg acg atc aaa ggt ctc aat cca aca gct att ttt	1310	Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe	1315	4098
1320	cta aca acc ctc tca aga acc agc aag aaa agg agc tgg cca tta aat	1325	Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn	1330	4146
1335	gag gct atc atg gca gtc ggg atg gtg agc att tta gcc agt tct ctc	1340	Glu Ala Ile Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu	1345	4194
1355	cta aaa aat gat att ccc atg aca gga cca tta gtg gct gga ggg ctc	1360	Leu Lys Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu	1365	4242
1370	ctc act gtg tgc tac gtg ctc act gga cga tgc gcc gat ttg gaa ctg	1375	Leu Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu	1380	4290
1385	gag aga gca gcc gat gtc aaa tgg gaa gac cag gca gag ata tca gga	1390	Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly	1395	4338
1400	agc agt cca atc ctg tca ata aca ata tca gaa gat ggt agc atg tcg	1405	Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser	1410	4386
1415	ata aaa aat gaa gag gaa gaa caa aca ctg acc ata ctc att aga aca	1420	Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg Thr	1425	4434
1435	gga ttg ctg gtg atc tca gga ctt ttt cct gta tca ata cca atc acg	1440	Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro Ile Thr	1445	4482
1450	gca gca gca tgg tac ctg tgg gaa gtg aag aaa caa cgg gcc gga gta	1455	Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg Ala Gly Val	1460	4530
1465	ttg tgg gat gtt cct tca ccc cca ccc atg gga aag gct gaa ctg gaa	1470	Leu Trp Asp Val Pro Ser Pro Pro Met Gly Lys Ala Glu Leu Glu	1475	4578
1480	gat gga gcc tat aga att aag caa aaa ggg att ctt gga tat tcc cag	1485	Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile Leu Gly Tyr Ser Gln	1490	4626
1495	atc gga gcc gga gtt tac aaa gaa gga aca ttc cat aca atg tgg cat	1500	Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr Phe His Thr Met Trp His	1505	4674
1515	gtc aca cgt ggc gct gtt cta atg cat aaa gga aag agg att gaa cca	1520	Val Thr Arg Gly Ala Val Leu Met His Lys Gly Lys Arg Ile Glu Pro	1525	4722

tca	tg	gcg	gac	gtc	aag	aaa	gac	cta	ata	tca	tat	gga	gga	ggc	tg	4770
Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	
		1545					1550					1555				
aag	tta	gaa	gga	gaa	tg	aag	gaa	gga	gaa	gaa	gtc	cag	gta	ttg	gca	4818
Lys	Leu	Glu	Gly	Glu	Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	
		1560					1565				1570					
ctg	gag	cct	gga	aaa	aat	cca	aga	gcc	gtc	caa	acg	aaa	cct	ggt	ctt	4866
Leu	Glu	Pro	Gly	Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	
		1575			1580				1585						1590	
ttc	aaa	acc	aac	gcc	gga	aca	ata	ggt	gct	gta	tct	ctg	gac	ttt	tct	4914
Phe	Lys	Thr	Asn	Gly	Thr	Ile	Gly	Ile	Ala	Val	Ser	Leu	Asp	Phe	Ser	
			1595				1600							1605		
cct	gga	acg	tca	gga	tct	cca	att	atc	gac	aaa	aaa	gga	aaa	ggt	gtg	4962
Pro	Gly	Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	
			1610				1615					1620				
ggt	ctt	tat	ggt	aat	ggt	gtt	gtt	aca	agg	agt	gga	gca	tat	gtg	agt	5010
Gly	Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser	
		1625				1630						1635				
gct	ata	gcc	cag	act	gaa	aaa	agc	att	gaa	gac	aac	cca	gag	atc	gaa	5058
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile	Glu	
		1640			1645				1650							
gat	gac	att	ttc	cga	aag	aga	aga	ctg	acc	atc	atg	gac	ctc	cac	cca	5106
Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu	His	Pro	
		1655			1660				1665					1670		
gga	gcg	gga	aag	acg	aag	aga	tac	ctt	ccg	gcc	ata	gtc	aga	gaa	gct	5154
Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val	Arg	Glu	Ala	
			1675					1680						1685		
ata	aaa	cgg	ggt	ttg	aga	aca	tta	atc	ttg	gcc	ccc	act	aga	gtt	gtg	5202
Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro	Thr	Arg	Val	Val	
		1690					1695					1700				
gca	gct	gaa	atg	gag	gaa	gcc	ctt	aga	gga	ctt	cca	ata	aga	tac	cag	5250
Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu	Pro	Ile	Arg	Tyr	Gln	
		1705				1710						1715				
acc	cca	gcc	atc	aga	gct	gtg	cac	acc	ggg	cgg	gag	att	gtg	gac	cta	5298
Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly	Arg	Glu	Ile	Val	Asp	Leu	
		1720			1725				1730							
atg	tgt	cat	gcc	aca	ttt	acc	atg	agg	ctg	cta	tca	cca	gtt	aga	gtg	5346
Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg	Leu	Leu	Ser	Pro	Val	Arg	Val	
		1735			1740				1745					1750		
cca	aac	tac	aac	ctg	att	atc	atg	gac	gaa	gcc	cat	ttc	aca	gac	cca	5394
Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met	Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	
			1755					1760						1765		
gca	agt	ata	gca	gct	aga	gga	tac	atc	tca	act	cga	gtg	gag	atg	ggt	5442
Ala	Ser	Ile	Ala	Ala	Arg	Gly	Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	
		1770				1775						1780				
gag	gca	gct	ggg	att	ttt	atg	aca	gcc	act	ccc	ccg	gga	agc	aga	gac	5490
Glu	Ala	Ala	Gly	Ile	Phe	Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	
		1785				1790						1795				
cca	ttt	cct	cag	agc	aat	gca	cca	atc	ata	gat	gaa	gaa	aga	gaa	atc	5538
Pro	Phe	Pro	Gln	Ser	Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	
		1800			1805				1810							
cct	gaa	cgc	tcg	tg	aat	tcc	gga	cat	gaa	tg	gtc	acg	gat	ttt	aaa	5586
Pro	Glu	Arg	Ser	Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	
		1815			1820				1825					1830		
ggg	aag	act	gtt	tg	ttc	gtt	cca	agt	ata	aaa	gca	gga	aat	gat	ata	5634
Gly	Lys	Thr	Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	
			1835					1840						1845		
gca	gct	tcg	ctg	agg	aaa	aat	gga	aag	aaa	gtg	ata	caa	ctc	agt	agg	5682
Ala	Ala	Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	
			1850				1855					1860				
aag	acc	ttt	gat	tct	gag	tat	gtc	aag	act	aga	acc	aat	gat	tg	gac	5730
Lys	Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	
		1865					1870					1875				
ttc	gtg	gtt	aca	act	gac	att	tca	gaa	atg	ggt	gcc	aat	ttc	aag	gct	5778
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	
		1880			1885				1890							
gag	agg	gtt	ata	gac	ccc	aga	cgc	tcg	atg	aaa	cca	gtc	ata	cta	aca	5826
Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr	

1895	gat	ggt	gaa	gag	cgg	gtg	att	ctg	gca	gga	cct	atg	cca	gtg	acc	cac	5874
Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr	His		
				1915						1920							
tct	agt	gca	gca	caa	aga	aga	ggg	aga	ata	gga	aga	aat	cca	aaa	aat	5922	
Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro	Lys	Asn		
				1930						1935							
gag	aat	gac	cag	tac	ata	tac	atg	ggg	gaa	cct	ctg	gaa	aat	gat	gaa	5970	
Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu	Asn	Asp	Glu		
				1945						1950							
gac	tgt	gca	cac	tgg	aaa	gaa	gct	aaa	atg	ctc	cta	gat	aac	atc	aac	6018	
Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu	Asp	Asn	Ile	Asn		
				1960						1970							
acg	cca	gaa	gga	atc	att	cct	agc	atg	ttc	gaa	cca	gag	cgt	gaa	aag	6066	
Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu	Pro	Glu	Arg	Glu	Lys		
				1975						1985						1990	
gtg	gat	gcc	att	gat	ggc	gaa	tac	cgc	ttg	aga	gga	gaa	gca	agg	aaa	6114	
Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr	Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys		
				1995						2000							
acc	ttt	gta	gac	tta	atg	aga	aga	gga	gac	cta	cca	gtc	tgg	ttg	gcc	6162	
Thr	Phe	Val	Asp	Leu	Met	Arg	Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala		
				2010						2015							
tac	aga	gtg	gca	gct	gaa	ggc	atc	aac	tac	gca	gac	aga	agg	tgg	tgt	6210	
Tyr	Arg	Val	Ala	Ala	Glu	Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys		
				2025						2030							
ttt	gat	gga	gtc	aag	aac	aac	caa	atc	cta	gaa	gaa	aac	gtg	gaa	grr	6258	
Phe	Asp	Gly	Val	Lys	Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val		
				2040						2050							
gaa	atc	tgg	aca	aaa	gaa	ggg	gaa	agg	aag	aaa	ttg	aaa	ccc	aga	tgg	6306	
Glu	Ile	Trp	Thr	Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp		
				2055						2065						2070	
ttg	gat	gct	agg	atc	tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	6354	
Leu	Asp	Ala	Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys		
				2075						2080							
gaa	ttt	gca	gcc	gga	aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	6402	
Glu	Phe	Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu		
				2090						2095							
atg	ggt	agg	ctc	cca	acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	6450	
Met	Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu		
				2105						2110							
gac	aac	tta	gca	gtg	ctg	cac	acg	gct	gag	gca	ggt	gga	agg	gcg	tac	6498	
Asp	Asn	Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr		
				2120						2130							
aac	cat	gct	ctc	agt	gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	6546	
Asn	His	Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu		
				2135						2145						2150	
ctg	aca	ctt	ctg	gct	aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	6594	
Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met		
				2155						2160						2165	
agc	gca	agg	ggc	ata	ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	6642	
Ser	Ala	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile		
				2170						2175						2180	
acg	gct	agc	atc	ctc	cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	6690	
Thr	Ala	Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile		
				2185						2190							
gca	gct	tca	ata	ata	ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	6738	
Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro		
				2200						2205							
gaa	cct	gaa	aaa	cag	aga	aca	ccc	caa	gac	aac	caa	ctg	acc	tac	gtt	6786	
Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln	Asp	Asn	Gln	Leu	Thr	Tyr	Val		
				2215						2225						2230	
gtc	ata	gcc	atc	ctc	aca	gtg	gtg	gcc	gca	acc	atg	gca	aac	gag	atg	6834	
Val	Ile	Ala	Ile	Leu	Thr	Val	Val	Ala	Ala	Thr	Met	Ala	Asn	Glu	Met		
				2235						2240						2245	
ggt	ttc	cta	gaa	aaa	acg	aag	aaa	gat	ctc	gga	ttg	gga	agc	att	gca	6882	
Gly	Phe	Leu	Glu	Lys	Thr	Lys	Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala		
				2250						2255						2260	
acc	cag	caa	ccc	gag	agc	aac	atc	ctg	gac	ata	gat	cta	cgt	cct	gca	6930	
Thr	Gln	Gln	Pro	Glu	Ser	Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala		

tca	gca	tgg	acg	ctg	tat	gcc	gtg	gcc	aca	aca	ttt	gtt	aca	cca	atg	6978
Ser	Ala	Trp	Thr	Leu	Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	
2280						2285					2290					
ttg	aga	cat	agc	att	gaa	aat	tcc	tca	gtg	aat	gtg	tcc	cta	aca	gct	7026
Leu	Arg	His	Ser	Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	
2295					2300					2305					2310	
ata	gcc	aac	caa	gcc	aca	gtg	tta	atg	ggg	ctc	ggg	aaa	gga	tgg	cca	7074
Ile	Ala	Asn	Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	
				2315					2320					2325		
ttg	tca	aag	atg	gac	atc	gga	gtt	ccc	ctt	ctc	gcc	att	gga	tgc	tac	7122
Leu	Ser	Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	
			2330					2335					2340			
tca	caa	gtc	aac	ccc	ata	act	ctc	aca	gca	gct	ctt	ttc	tta	ttg	gta	7170
Ser	Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	
		2345					2350					2355				
gca	cat	tat	gcc	atc	ata	ggg	cca	gga	ctc	caa	gca	aaa	gca	acc	aga	7218
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	
		2360				2365					2370					
gaa	gct	cag	aaa	aga	gca	gcg	gcg	ggc	atc	atg	aaa	aac	cca	act	gtc	7266
Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val	
2375					2380					2385					2390	
gat	gga	ata	aca	gtg	att	gac	cta	gat	cca	ata	cct	tat	gat	cca	aag	7314
Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr	Asp	Pro	Lys	
				2395				2400						2405		
ttt	gaa	aag	cag	ttg	gga	caa	gta	atg	ctc	cta	gtc	ctc	tgc	gtg	act	7362
Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val	Leu	Cys	Val	Thr	
			2410					2415					2420			
caa	gta	ttg	atg	atg	agg	act	aca	tgg	gct	ctg	tgt	gag	gct	tta	acc	7410
Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu	Cys	Glu	Ala	Leu	Thr	
		2425					2430					2435				
tta	gct	acc	ggg	ccc	atc	tcc	aca	ttg	tgg	gaa	gga	aat	cca	ggg	agg	7458
Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp	Glu	Gly	Asn	Pro	Gly	Arg	
		2440				2445				2450						
ttt	tgg	aac	act	acc	att	gcg	gtg	tca	atg	gct	aac	att	ttt	aga	ggg	7506
Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser	Met	Ala	Asn	Ile	Phe	Arg	Gly	
2455					2460					2465				2470		
agt	tac	ttg	gcc	gga	gct	gga	ctt	ctc	ttt	tct	att	atg	aag	aac	aca	7554
Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu	Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	
			2475					2480					2485			
acc	aac	aca	aga	agg	gga	act	ggc	aac	ata	gga	gag	acg	ctt	gga	gag	7602
Thr	Asn	Thr	Arg	Arg	Gly	Thr	Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	
		2490					2495						2500			
aaa	tgg	aaa	agc	cga	ttg	aac	gca	ttg	gga	aaa	agt	gaa	ttc	cag	atc	7650
Lys	Trp	Lys	Ser	Arg	Leu	Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	
		2505					2510					2515				
tac	aag	aaa	agt	gga	atc	cag	gaa	gtg	gat	aga	acc	tta	gca	aaa	gaa	7698
Tyr	Lys	Lys	Ser	Gly	Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	
		2520				2525					2530					
ggc	att	aaa	aga	gga	gaa	acg	gac	cat	cac	gct	gtg	tcg	cga	ggc	tca	7746
Gly	Ile	Lys	Arg	Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	
		2535			2540					2545				2550		
gca	aaa	ctg	aga	tgg	ttc	gtt	gag	aga	aac	atg	gtc	aca	cca	gaa	ggg	7794
Ala	Lys	Leu	Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	
		2555					2560						2565			
aaa	gta	gtg	gac	ctc	ggt	tgt	ggc	aga	gga	ggc	tgg	tca	tac	tat	tgt	7842
Lys	Val	Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	
		2570					2575					2580				
gga	gga	cta	aag	aat	gta	aga	gaa	gtc	aaa	ggc	cta	aca	aaa	gga	gga	7890
Gly	Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly	
		2585					2590				2595					
cca	gga	cac	gaa	gaa	ccc	atc	ccc	atg	tca	aca	tat	ggg	tgg	aat	cta	7938
Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn	Leu	
		2600				2605					2610					
gtg	cgt	ctt	caa	agt	gga	gtt	gac	gtt	ttc	ttc	atc	ccg	cca	gaa	aag	7986
Val	Arg	Leu	Gln	Ser	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro	Glu	Lys	
		2615			2620					2625				2630		
tgt	gac	aca	tta	ttg	tgt	gac	ata	ggg	gag	tca	tca	cca	aat	ccc	aca	8034
Cys	Asp	Thr	Leu	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro	Thr	

gtg	gaa	gca	gga	cga	aca	ctc	aga	gtc	ctt	aac	tta	gta	gaa	aat	tgg	8082
Val	Glu	Ala	Gly	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu	Asn	Trp	
2650								2655					2660			
ttg	aac	aac	aac	act	caa	ttt	tgc	ata	aag	gtt	ctc	aac	cca	tat	atg	8130
Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn	Pro	Tyr	Met	
2665							2670					2675				
ccc	tca	gtc	ata	gaa	aaa	atg	gaa	gca	cta	caa	agg	aaa	tat	gga	gga	8178
Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg	Lys	Tyr	Gly	Gly	
2680						2685					2690					
gcc	tta	gtg	agg	aat	cca	ctc	tca	cga	aac	tcc	aca	cat	gag	atg	tac	8226
Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser	Thr	His	Glu	Met	Tyr	
2695					2700					2705					2710	
tgg	gta	tcc	aat	gct	tcc	ggg	aac	ata	gtg	tca	tca	gtg	aac	atg	att	8274
Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val	Ser	Ser	Val	Asn	Met	Ile	
2715									2720						2725	
tca	agg	atg	ttg	atc	aac	aga	ttt	aca	atg	aga	tac	aag	aaa	gcc	act	8322
Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	
2730							2735						2740			
tac	gag	ccg	gat	gtt	gac	ctc	gga	agc	gga	acc	cgt	aac	atc	ggg	att	8370
Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	
2745						2750						2755				
gaa	agt	gag	ata	cca	aac	cta	gat	ata	att	ggg	aaa	aga	ata	gaa	aaa	8418
Glu	Ser	Glu	Ile	Pro	Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	
2760						2765				2770						
ata	aag	caa	gag	cat	gaa	aca	tca	tgg	cac	tat	gac	caa	gac	cac	cca	8466
Ile	Lys	Gln	Glu	His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	
2775					2780					2785					2790	
tac	aaa	acg	tgg	gca	tac	cat	ggt	agc	tat	gaa	aca	aaa	cag	act	gga	8514
Tyr	Lys	Thr	Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	
2795							2800						2805			
tca	gca	tca	tcc	atg	gtc	aac	gga	gtg	gtc	agg	ctg	ctg	aca	aaa	cct	8562
Ser	Ala	Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	
2810						2815						2820				
tgg	gac	gtc	gtc	ccc	atg	gtg	aca	cag	atg	gca	atg	aca	gac	acg	act	8610
Trp	Asp	Val	Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr	
2825						2830						2835				
cca	ttt	gga	caa	cag	cgc	gtt	ttt	aaa	gag	aaa	gtg	gac	acg	aga	acc	8658
Pro	Phe	Gly	Gln	Gln	Arg	Gly	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	
2840						2845					2850					
caa	gaa	ccg	aaa	gaa	ggc	acg	aag	aaa	cta	atg	aaa	ata	aca	gca	gag	8706
Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu	
2855					2860					2865		2870				
tgg	ctt	tgg	aaa	gaa	tta	ggg	aag	aaa	aag	aca	ccc	agg	atg	tgc	acc	8754
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys	Thr	
2875						2880						2885				
aga	gaa	gaa	ttc	aca	aga	aag	gtg	aga	agc	aat	gca	gcc	ttg	ggg	gcc	8802
Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu	Gly	Ala	
2890						2895						2900				
ata	ttc	act	gat	gag	aac	aag	tgg	aag	tcg	gca	cgt	gag	gct	gtt	gaa	8850
Ile	Phe	Thr	Asp	Glu	Asn	Lys	Trp	Lys	Ser	Ala	Arg	Glu	Ala	Val	Glu	
2905						2910						2915				
gat	agt	agg	ttt	tgg	gag	ctg	gtt	gac	aag	gaa	agg	aat	ctc	cat	ctt	8898
Asp	Ser	Arg	Phe	Trp	Glu	Leu	Val	Asp	Lys	Glu	Arg	Asn	Leu	His	Leu	
2920						2925						2930				
gaa	gga	aag	tgt	gaa	aca	tgt	gtg	tac	aac	atg	atg	gga	aaa	aga	gag	8946
Glu	Gly	Lys	Cys	Glu	Thr	Cys	Val	Tyr	Asn	Met	Met	Gly	Lys	Arg	Glu	
2935					2940					2945					2950	
aag	aag	cta	ggg	gaa	ttc	ggc	aag	gca	aaa	ggc	agc	aga	gcc	ata	tgg	8994
Lys	Lys	Leu	Gly	Glu	Phe	Gly	Lys	Ala	Lys	Gly	Ser	Arg	Ala	Ile	Trp	
2955					2960							2965				
tac	atg	tgg	ctt	gga	gca	cgc	ttc	tta	gag	ttt	gaa	gcc	cta	gga	ttc	9042
Tyr	Met	Trp	Leu	Gly	Ala	Arg	Phe	Leu	Glu	Phe	Glu	Ala	Leu	Gly	Phe	
2970						2975						2980				
tta	aat	gaa	gat	cac	tgg	ttc	tcc	aga	gag	aac	tcc	ctg	agt	gga	gtg	9090
Leu	Asn	Glu	Asp	His	Trp	Phe	Ser	Arg	Glu	Asn	Ser	Leu	Ser	Gly	Val	
2985						2990						2995				
gaa	gga	gaa	ggg	ctg	cac	aag	cta	ggt	tac	att	cta	aga	gac	gtg	agc	9138
Glu	Gly	Glu	Gly	Leu	His	Lys	Leu	Gly	Tyr	Ile	Leu	Arg	Asp	Val	Ser	

3000	aag aaa gag gga gga gca atg tat gcc gat gac acc gca gga tgg gat	3005	3010	9186
Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp	3015	3020	3025	3030
aca aga atc aca cta gaa gac cta aaa aat gaa gaa atg gta aca aac	3035	3040	3045	9234
Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn	3050	3055	3060	9282
cac atg gaa gga gaa cac aag aaa cta gcc gag gcc att ttc aaa cta	3065	3070	3075	9330
His Met Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu	3080	3085	3090	9378
acg tac caa aac aag gtg gtg cgt gtg caa aga cca aca cca aga ggc	3095	3100	3105	9426
Thr Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly	3115	3120	3125	9474
aca gta atg gac atc ata tct aga aga gac caa aga ggt agt gga caa	3130	3135	3140	9522
Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln	3145	3150	3155	9570
gtt ggc acc tat gga ctc aat act ttc acc aat atg gaa gcc caa cta	3160	3165	3170	9618
Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu	3175	3180	3185	9666
atc aga cag atg gag gga gaa gga gtc ttt aaa agc att cag cac cta	3190	3195	3200	9714
Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His Leu	3205	3210	3215	9762
aca atc aca gaa gaa atc gct gtg caa aac tgg tta gca aga gtg ggg	3220	3225	3230	9810
Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg Val Gly	3235	3240	3245	9858
cgc gaa agg tta tca aga atg gcc atc agt gga gat gat tgt gtt gtg	3250	3255	3260	9906
Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp Cys Val Val	3265	3270	3275	9954
aaa cct tta gat gac agg ttc gca agc gct tta aca gct cta aat gac	3280	3285	3290	10002
Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr Ala Leu Asn Asp	3295	3300	3305	10050
atg gga aag att agg aaa gac ata caa caa tgg gaa cct tca aga gga	3310	3315	3320	10098
Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg Gly	3325	3330	3335	10146
tgg aat gat tgg aca caa gtg ccc ttc tgt tca cac cat ttc cat gag	3340	3345	3350	10194
Trp Asn Asp Trp Thr Gln Val Pro Phe Cys Ser His His Phe His Glu	3355	3360	3365	10242
tta atc atg aaa gac ggt cgc gta ctc gtt gtt cca tgt aga aac caa	3370	3375	3380	
Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln	3385	3390	3395	
gat gaa ctg att ggc aga gcc cga atc tcc caa gga gca ggg tgg tct	3400	3405	3410	
Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser	3415	3420	3425	
ttg cgg gag acg gcc tgt ttg ggg aag tct tac gcc caa atg tgg agc	3430	3435	3440	
Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser	3445	3450	3455	
ttg atg tac ttc cac aga cgc gac ctc agg ctg gcg gca aat gct att	3460	3465	3470	
Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile	3475	3480	3485	
tgc tct gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg	3490	3495	3500	
Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp	3505	3510	3515	
tcc ata cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca	3520	3525	3530	
Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr	3535	3540	3545	
gtc tgg aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa	3550	3555	3560	
Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys	3565	3570	3575	
act cca gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa	3580	3585	3590	
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu	3595	3600	3605	
gac caa tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg	3610	3615	3620	
Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp	3625	3630	3635	
gca aag aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc	3640	3645	3650	
Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly	3655	3660	3665	
aat gaa gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga	3670	3675	3680	
Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg	3685	3690	3695	

```

          3370          3375          3380
gaa gag gaa gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa 10292
Glu Glu Glu Glu Ala Gly Val Leu Trp *
          3385          3390
caaggctaga agtcaggctg gattaagcca tagtacggaa aaaactatgc tacctgtgag 10352
ccccgtccaa ggacgttaaa agaagtcagg ccatcataaa tgccatagct tgagtaaact 10412
atgcagcctg tagctccacc tgagaagggtg taaaaaatcc gggaggccac aaaccatgga 10472
agctgtacgc atggcgtagt ggactagcgg ttagaggaga cccctccctt acaaatcgca 10532
gcaacaatgg gggcccaagg cgagatgaag ctgtagtctc gctggaagga ctagagggtta 10592
gaggagaccc ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg 10652
ctgtctcctc agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat 10712
caacaggttc t 10723

```

<210> SEQ ID NO 12

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 12

Met 1	Asn	Asn	Gln	Arg 5	Lys	Lys	Ala	Lys	Asn 10	Thr	Pro	Phe	Asn 15	Met	Leu
Lys	Arg	Glu	Arg 20	Asn	Arg	Val	Ser	Thr 25	Val	Gln	Gln	Leu	Thr 30	Lys	Arg
Phe	Ser	Leu 35	Gly	Met	Leu	Gln	Gly 40	Arg	Gly	Pro	Leu	Lys 45	Leu	Phe	Met
Ala	Leu	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly
Ile 65	Leu	Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	Lys	Ser 75	Lys	Ala	Ile	Asn	Val 80
Leu	Arg	Gly	Phe	Arg 85	Lys	Glu	Ile	Gly	Arg 90	Met	Leu	Asn	Ile	Leu	Asn 95
Arg	Arg	Arg	Ser 100	Ser	Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120	Asp	Gly	Glu	Pro	Leu 125	Met	Ile	Val
Ala	Lys 130	His	Glu	Arg	Gly	Arg 135	Pro	Leu	Leu	Phe	Lys 140	Thr	Thr	Glu	Gly
Ile 145	Asn	Lys	Cys	Thr	Leu 150	Ile	Ala	Met	Asp	Leu 155	Gly	Glu	Met	Cys	Glu 160
Asp	Thr	Val	Thr	Tyr 165	Lys	Cys	Pro	Leu	Leu 170	Val	Asn	Thr	Glu	Pro 175	Glu
Asp	Ile	Asp	Cys 180	Trp	Cys	Asn	Leu	Thr 185	Ser	Thr	Trp	Val	Met 190	Tyr	Gly
Thr	Cys	Thr 195	Gln	Ser	Gly	Glu	Arg 200	Arg	Arg	Glu	Lys	Arg 205	Ser	Val	Ala
Leu	Thr 210	Pro	His	Ser	Gly	Met 215	Gly	Leu	Glu	Thr	Arg 220	Ala	Glu	Thr	Trp
Met 225	Ser	Ser	Glu	Gly	Ala 230	Trp	Lys	His	Ala	Gln 235	Arg	Val	Glu	Ser	Trp 240
Ile	Leu	Arg	Asn 245	Pro	Gly	Phe	Ala	Leu	Leu 250	Ala	Gly	Phe	Met	Ala 255	Tyr
Met	Ile	Gly	Gln 260	Thr	Gly	Ile	Gln	Arg 265	Thr	Val	Phe	Phe	Val 270	Leu	Met
Met	Leu	Val 275	Ala	Pro	Ser	Tyr	Gly 280	Met	Arg	Cys	Val	Gly 285	Val	Gly	Asn
Arg	Asp 290	Phe	Val	Glu	Gly	Val 295	Ser	Gly	Gly	Ala	Trp 300	Val	Asp	Leu	Val
Leu 305	Glu	His	Gly	Gly	Cys 310	Val	Thr	Thr	Met	Ala 315	Gln	Gly	Lys	Pro	Thr 320
Leu	Asp	Phe	Glu 325	Leu	Thr	Lys	Thr	Thr	Ala 330	Lys	Glu	Val	Ala	Leu 335	Leu
Arg	Thr	Tyr	Cys 340	Ile	Glu	Ala	Ser	Ile 345	Ser	Asn	Ile	Thr	Thr 350	Ala	Thr
Arg	Cys	Pro 355	Thr	Gln	Gly	Glu	Pro 360	Tyr	Leu	Lys	Glu	Glu 365	Gln	Asp	Gln
Gln	Tyr 370	Ile	Cys	Arg	Arg	Asp 375	Val	Val	Asp	Arg	Gly 380	Trp	Gly	Asn	Gly

Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
 385 390 395 400
 Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
 405 410 415
 Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
 420 425 430
 Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
 435 440 445
 Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
 450 455 460
 Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
 465 470 475 480
 Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
 485 490 495
 Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
 500 505 510
 Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
 515 520 525
 Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
 530 535 540
 Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
 545 550 555 560
 Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
 565 570 575
 Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
 580 585 590
 Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
 595 600 605
 Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
 610 615 620
 Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
 625 630 635 640
 Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
 645 650 655
 Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
 660 665 670
 Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
 675 680 685
 Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
 690 695 700
 Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
 705 710 715 720
 Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
 725 730 735
 Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
 740 745 750
 Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
 755 760 765
 Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
 770 775 780
 Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 785 790 795 800
 His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 805 810 815
 Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
 820 825 830
 Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
 835 840 845
 Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860
 Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875 880
 Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895
 Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910
 Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925
 Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu

930 935 940
 Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960
 Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975
 Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990
 Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005
 Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly
 1010 1015 1020
 Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035 1040
 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 1050 1055
 Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 1065 1070
 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 1080 1085
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys
 1090 1095 1100
 Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu
 1105 1110 1115 1120
 Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
 1125 1130 1135
 Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
 1140 1145 1150
 Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
 1155 1160 1165
 Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
 1170 1175 1180
 Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
 1185 1190 1195 1200
 Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
 1205 1210 1215
 Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr
 1220 1225 1230
 Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
 1235 1240 1245
 Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
 1250 1255 1260
 Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
 1265 1270 1275 1280
 Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
 1285 1290 1295
 Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser
 1300 1305 1310
 Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
 1315 1320 1325
 Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
 1330 1335 1340
 Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
 1345 1350 1355 1360
 Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
 1365 1370 1375
 Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
 1380 1385 1390
 Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
 1395 1400 1405
 Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
 1410 1415 1420
 Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Gln Thr Leu
 1425 1430 1435 1440
 Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
 1445 1450 1455
 Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
 1460 1465 1470
 Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Met
 1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
 1490 1495 1500
 Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
 1505 1510 1515
 Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
 1525 1530 1535
 Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
 1540 1545 1550
 Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu
 1555 1560 1565
 Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val
 1570 1575 1580
 Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
 1585 1590 1595 1600
 Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
 1605 1610 1615
 Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
 1620 1625 1630
 Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
 1635 1640 1645
 Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr
 1650 1655 1660
 Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1665 1670 1675 1680
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1685 1690 1695
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1700 1705 1710
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
 1715 1720 1725
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1730 1735 1740
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775
 Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr
 1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu

2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125
 Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205
 Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp
 2210 2215 2220
 Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 2230 2235 2240
 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu
 2245 2250 2255
 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp
 2260 2265 2270
 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr
 2275 2280 2285
 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val
 2290 2295 2300
 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly
 2305 2310 2315 2320
 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu
 2325 2330 2335
 Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala
 2340 2345 2350
 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu
 2355 2360 2365
 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile
 2370 2375 2380
 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro
 2385 2390 2395 2400
 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu
 2405 2410 2415
 Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala
 2420 2425 2430
 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
 2435 2440 2445
 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met
 2450 2455 2460
 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
 2465 2470 2475 2480
 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile
 2485 2490 2495
 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly
 2500 2505 2510
 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp
 2515 2520 2525
 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His
 2530 2535 2540
 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn
 2545 2550 2555 2560
 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly
 2565 2570 2575
 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys
 2580 2585 2590

Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser
 2595 2600 2605
 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe
 2610 2615 2620
 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu
 2625 2630 2635 2640
 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu
 2645 2650 2655
 Asn Leu Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys
 2660 2665 2670
 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
 2675 2680 2685
 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn
 2690 2695 2700
 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
 2705 2710 2715 2720
 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met
 2725 2730 2735
 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly
 2740 2745 2750
 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile
 2755 2760 2765
 Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 2770 2775 2780
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 2785 2790 2795 2800
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 2805 2810 2815
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 2820 2825 2830
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 2835 2840 2845
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 2850 2855 2860
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2865 2870 2875 2880
 Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu
 2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe
 3105 3110 3115 3120
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser

```

      3140      3145      3150
Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
      3155      3160      3165
Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
      3170      3175      3180
Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
3185      3190      3195      3200
Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
      3205      3210      3215
Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
      3220      3225      3230
Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
      3235      3240      3245
Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
      3250      3255      3260
Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
3265      3270      3275      3280
Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
      3285      3290      3295
Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
      3300      3305      3310
Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Ile Pro
      3315      3320      3325
Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
      3330      3335      3340
Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
3345      3350      3355      3360
Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
      3365      3370      3375
Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
      3380      3385      3390

```

<210> SEQ ID NO 13

<211> COMPRIMENTO: 10723

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10272)

<400> SEQUÊNCIA: 13

agttgttagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaacgta	60
gttctaacag	ttttttaatt	agagagcaga	tctctg	atg	aat aac caa cgg aaa	114
				Met	Asn Asn Gln Arg Lys	
				1	5	
aag gcg aaa aac acg cct ttc aat atg ctg aaa cgc gag aga aac cgc	162					
Lys Ala Lys Asn Thr Pro Phe Asn Met Leu Lys Arg Glu Arg Asn Arg						
	10					
gtg tcg act gtg caa cag ctg aca aag aga ttc tca ctt gga atg ctg	210					
Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu						
	25					
cag gga cga gga cca tta aaa ctg ttc atg gcc ctg gtg gcg ttc ctt	258					
Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu						
	40					
cgt ttc cta aca atc cca cca aca gca ggg ata ttg aag aga tgg gga	306					
Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly Ile Leu Lys Arg Trp Gly						
	55					
aca att aaa aaa tca aaa gct att aat gtt ttg aga ggg ttc agg aaa	354					
Thr Ile Lys Lys Ser Lys Ala Ile Asn Val Leu Arg Gly Phe Arg Lys						
	75					
gag att gga agg atg ctg aac atc ttg aat agg aga cgc aga tct gca	402					
Glu Ile Gly Arg Met Leu Asn Ile Leu Asn Arg Arg Arg Arg Ser Ala						
	90					
ggc atg atc att atg ctg att cca aca gtg atg gcg ttc cat tta acc	450					
Gly Met Ile Ile Met Leu Ile Pro Thr Val Met Ala Phe His Leu Thr						
	105					
aca cgt aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg	498					
	110					
	115					

Thr	Arg	Asn	Gly	Glu	Pro	His	Met	Ile	Val	Ser	Arg	Gln	Glu	Lys	Gly	
aaa	agt	ctt	ctg	ttt	aaa	aca	gag	gat	ggc	gtg	aac	atg	tgt	acc	ctc	546
Lys	Ser	Leu	Leu	Phe	Lys	Thr	Glu	Asp	Gly	Val	Asn	Met	Cys	Thr	Leu	
135					140					145					150	
atg	gcc	atg	gac	ctt	ggt	gaa	tgt	tgt	gaa	gac	aca	atc	acg	tac	aag	594
Met	Ala	Met	Asp	Leu	Gly	Glu	Leu	Cys	Glu	Asp	Thr	Ile	Thr	Tyr	Lys	
				155					160					165		
tgt	ccc	ctt	ctc	agg	cag	aat	gag	cca	gaa	gac	ata	gac	tgt	tgg	tgc	642
Cys	Pro	Leu	Leu	Arg	Gln	Asn	Glu	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	
				170				175					180			
aac	tct	acg	tcc	acg	tgg	gta	act	tat	ggg	acg	tgt	acc	acc	atg	gga	690
Asn	Ser	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Thr	Thr	Met	Gly	
		185					190					195				
gaa	cat	aga	aga	gaa	aaa	aga	tca	gtg	gca	ctc	gtt	cca	cat	gtg	gga	738
Glu	His	Arg	Arg	Glu	Lys	Arg	Ser	Val	Ala	Leu	Val	Pro	His	Val	Gly	
	200					205					210					
atg	gga	ctg	gag	aca	cga	act	gaa	aca	tgg	atg	tca	tca	gaa	ggg	gcc	786
Met	Gly	Leu	Glu	Thr	Arg	Thr	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	
	215				220					225					230	
tgg	aaa	cat	gtc	cag	aga	att	gaa	act	tgg	atc	ttg	aga	cat	cca	ggc	834
Trp	Lys	His	Val	Gln	Arg	Ile	Glu	Thr	Trp	Ile	Leu	Arg	His	Pro	Gly	
				235					240					245		
ttc	acc	atg	atg	gca	atc	ctg	gca	tac	acc	ata	gga	acg	aca	cat		882
Phe	Thr	Met	Met	Ala	Ala	Ile	Leu	Ala	Tyr	Thr	Ile	Gly	Thr	Thr	His	
				250				255				260				
ttc	caa	aga	gcc	ctg	att	ttc	atc	tta	ctg	aca	gct	gtc	act	cct	tca	930
Phe	Gln	Arg	Ala	Leu	Ile	Phe	Ile	Leu	Leu	Thr	Ala	Val	Thr	Pro	Ser	
	265					270					275					
atg	aca	atg	cgt	tgc	ata	gga	atg	tca	aat	aga	gac	ttt	gtg	gaa	ggg	978
Met	Thr	Met	Arg	Cys	Ile	Gly	Met	Ser	Asn	Arg	Asp	Phe	Val	Glu	Gly	
	280				285					290						
gtt	tca	gga	gga	agc	tgg	gtt	gac	ata	gtc	tta	gaa	cat	gga	agc	tgt	1026
Val	Ser	Gly	Gly	Ser	Trp	Val	Asp	Ile	Val	Leu	Glu	His	Gly	Ser	Cys	
	295				300				305						310	
gtg	acg	acg	atg	gca	aaa	aac	aaa	cca	aca	ttg	gat	ttt	gaa	ctg	ata	1074
Val	Thr	Thr	Met	Lys	Asn	Lys	Pro	Pro	Leu	Leu	Asp	Phe	Glu	Leu	Ile	
				315				320						325		
aaa	aca	gaa	gcc	aaa	cag	cct	gcc	acc	cta	agg	aag	tac	tgt	ata	gag	1122
Lys	Thr	Glu	Ala	Lys	Gln	Pro	Ala	Thr	Leu	Arg	Lys	Tyr	Cys	Ile	Glu	
			330					335					340			
gca	aag	cta	acc	aac	aca	aca	aca	gaa	tct	cgc	tgc	cca	aca	caa	ggg	1170
Ala	Lys	Leu	Thr	Asn	Thr	Thr	Thr	Glu	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
		345					350					355				
gaa	ccc	agc	cta	aat	gaa	gag	cag	gac	aaa	agg	ttc	gtc	tgc	aaa	cac	1218
Glu	Pro	Ser	Leu	Asn	Glu	Glu	Gln	Asp	Lys	Arg	Phe	Val	Cys	Lys	His	
	360				365						370					
tcc	atg	gta	gac	aga	gga	tgg	gga	aat	gga	tgt	gga	cta	ttt	gga	aag	1266
Ser	Met	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
	375				380				385					390		
gga	ggc	att	gtg	acc	tgt	gct	atg	ttc	aga	tgc	aaa	aag	aac	atg	gaa	1314
Gly	Gly	Ile	Val	Thr	Cys	Ala	Met	Phe	Arg	Cys	Lys	Lys	Asn	Met	Glu	
				395				400						405		
gga	aaa	gtt	gtg	caa	gaa	aac	ttg	gaa	tac	acc	att	gtg	ata	aca		1362
Gly	Lys	Val	Val	Gln	Pro	Glu	Asn	Leu	Glu	Tyr	Thr	Ile	Val	Ile	Thr	
		410					415					420				
cct	cac	tca	ggg	gaa	gag	cat	gca	gtc	gga	aat	gac	aca	gga	aaa	cat	1410
Pro	His	Ser	Gly	Glu	Glu	His	Ala	Val	Gly	Asn	Asp	Thr	Gly	Lys	His	
		425					430					435				
ggc	aag	gaa	atc	aaa	ata	aca	cca	cag	agt	tcc	atc	aca	gaa	gca	gaa	1458
Gly	Lys	Glu	Ile	Lys	Ile	Thr	Pro	Gln	Ser	Ser	Ile	Thr	Glu	Ala	Glu	
	440				445					450						
ttg	aca	ggt	tat	ggc	act	gtc	aca	atg	gag	tgc	tct	cca	aga	acg	ggc	1506
Leu	Thr	Gly	Tyr	Gly	Thr	Val	Thr	Met	Glu	Cys	Ser	Pro	Arg	Thr	Gly	
	455				460				465					470		
ctc	gac	ttc	aat	gag	atg	gtg	ttg	ctg	cag	atg	gaa	aat	aaa	gct	tgg	1554
Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Gln	Met	Glu	Asn	Lys	Ala	Trp	
				475				480						485		
ctg	gtg	cac	agg	caa	tgg	ttc	cta	gac	ctg	ccg	tta	cca	tgg	ttg	ccc	1602

Leu	Val	His	Arg	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Trp	Leu	Pro		
gga	gcg	gac	aca	caa	ggg	tca	aat	tgg	ata	cag	aaa	gag	aca	ttg	gtc	1650	
Gly	Ala	Asp	Thr	Gln	Gly	Ser	Asn	Trp	Ile	Gln	Lys	Glu	Thr	Leu	Val		
		505					510					515					
act	ttc	aaa	aat	ccc	cat	gcg	aag	aaa	cag	gat	gtt	gtt	gtt	tta	gga	1698	
Thr	Phe	Lys	Asn	Pro	His	525	Lys	Lys	Gln	Asp	Val	Val	Val	Leu	Gly		
	520										530						
tcc	caa	gaa	ggg	gcc	atg	cac	aca	gca	ctt	aca	ggg	gcc	aca	gaa	atc	1746	
Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile		
					540										550		
caa	atg	tca	tca	gga	aac	tta	ctc	ttc	aca	gga	cat	ctc	aag	tgc	agg	1794	
Gln	Met	Ser	Ser	Gly	Asn	Leu	Leu	Phe	Thr	Gly	His	Leu	Lys	Cys	Arg		
				555											565		
ctg	aga	atg	gac	cta	cag	ctc	aaa	gga	atg	tca	tac	tct	atg	tgc		1842	
Leu	Arg	Met	Asp	Lys	Leu	Gln	Leu	Lys	Gly	Met	Ser	Tyr	Ser	Met	Cys		
			570					575					580				
aca	gga	aag	ttt	aaa	gtt	gtg	aag	gaa	ata	gca	gaa	aca	caa	cat	gga	1890	
Thr	Gly	Lys	Phe	Lys	Val	Val	Lys	Glu	Ile	Ala	Glu	Thr	Gln	His	Gly		
		585					590					595					
aca	ata	gtt	atc	aga	gtg	caa	tat	gaa	ggg	gac	ggc	tct	cca	tgc	aag	1938	
Thr	Ile	Val	Ile	Arg	Val	Gln	Tyr	Glu	Gly	Asp	Gly	Ser	Pro	Cys	Lys		
	600					605					610						
atc	cct	ttt	gag	ata	atg	gat	ttg	gaa	aaa	aga	cat	gtc	tta	ggt	cgc	1986	
Ile	Pro	Phe	Glu	Ile	Met	Asp	Leu	Glu	Lys	Arg	His	Val	Leu	Gly	Arg		
					620					625					630		
ctg	att	aca	gtc	aac	cca	att	gtg	aca	gaa	aaa	gat	agc	cca	gtc	aac	2034	
Leu	Ile	Thr	Val	Asn	Pro	Ile	Val	Thr	Glu	Lys	Asp	Ser	Pro	Val	Asn		
				635					640					645			
ata	gaa	gca	gaa	cct	cca	ttc	gga	gac	agc	tac	atc	atc	ata	gga	gta	2082	
Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Asp	Ser	Tyr	Ile	Ile	Ile	Gly	Val		
			650					655					660				
gag	ccg	gga	caa	ctg	aag	ctc	aac	tgg	ttt	aag	aaa	gga	agt	tct	atc	2130	
Glu	Pro	Gly	Gln	Leu	Lys	Leu	Asn	Trp	Phe	Lys	Lys	Gly	Ser	Ser	Ile		
		665					670					675					
ggc	caa	atg	ttt	gag	aca	aca	atg	agg	ggg	gcg	aag	aga	atg	gcc	att	2178	
Gly	Gln	Met	Phe	Glu	Thr	Thr	Met	Arg	Gly	Ala	Lys	Arg	Met	Ala	Ile		
		680				685					690						
tta	ggg	gac	aca	gcc	tgg	gat	ttt	gga	tcc	ttg	gga	gga	gtg	ttt	aca	2226	
Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Leu	Gly	Gly	Val	Phe	Thr		
					700					705				710			
tct	ata	gga	aag	gct	ctc	caa	gtc	ttt	gga	gca	atc	tat	gga	gct	Ala	2274	
Ser	Ile	Gly	Lys	Ala	Leu	His	Gln	Val	Phe	Gly	Ala	Ile	Tyr	Gly			
				715				720					725				
gcc	ttc	agt	ggg	gtt	tca	tgg	act	atg	aaa	atc	ctc	ata	gga	gtc	att	2322	
Ala	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	Ile	Leu	Ile	Gly	Val	Ile		
			730					735					740				
atc	aca	tgg	ata	gga	atg	aat	tca	cgc	agc	acc	tca	ctg	tct	gtg	aca	2370	
Ile	Thr	Trp	Ile	Gly	Met	Asn	Ser	Arg	Ser	Thr	Ser	Leu	Ser	Val	Thr		
		745				750						755					
cta	gta	ttg	gtg	gga	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	2418	
Leu	Val	Leu	Val	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln		
					765					770							
gcc	gat	agt	ggg	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	2466	
Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys		
				780						785				790			
ggc	agt	ggg	att	ttc	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	2514	
Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln		
				795					800					805			
tac	aag	ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	2562	
Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ala	Ile	Gln		
				810				815					820				
aaa	gcc	cat	gaa	gag	ggc	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	2610	
Lys	Ala	His	Glu	Glu	Gly	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu		
		825				830					835						
gag	aat	ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	2658	
Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu		
		840				845					850						
tca	gaa	aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	2706	

Ser 855	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	
atc	atg	cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	2754
Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu	
				875					880					885		
aag	tat	tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	2802
Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu	
			890					895					900			
tct	cat	aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	2850
Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys	
		905					910					915				
ccc	aac	aca	aat	aga	gct	tgg	aat	tgc	ttg	gaa	gtt	gaa	gac	tat	ggc	2898
Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly	
	920					925				930						
ttt	gga	gta	ttc	acc	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	cag	2946
Phe	Gly	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln	
	935				940				945						950	
gat	gta	ttc	tgc	gac	tca	aaa	ctc	atg	tca	gcg	gcc	ata	aaa	gac	aac	2994
Asp	Val	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn	
			955					960						965		
aga	gcc	gtc	cat	gcc	gat	atg	ggt	tat	tgg	ata	gaa	agt	gca	ctc	aat	3042
Arg	Ala	Val	His	Ala	Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn	
			970					975					980			
gac	aca	tgg	aag	ata	gag	aaa	gcc	tct	ttc	att	gaa	gtt	aaa	aac	tgc	3090
Asp	Thr	Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys	
		985					990					995				
cac	tgg	cca	aaa	tca	cac	acc	ctc	tgg	agc	aat	gga	gtg	cta	gaa	agt	3138
His	Trp	Pro	Lys	Ser	His	Thr	Leu	Trp	Ser	Asn	Gly	Val	Leu	Glu	Ser	
	1000					1005					1010					
gag	atg	ata	att	cca	aag	aat	ctc	gct	gga	cca	gtg	tct	caa	cac	aac	3186
Glu	Met	Ile	Ile	Pro	Lys	Asn	Leu	Ala	Gly	Pro	Val	Ser	Gln	His	Asn	
	1015				1020				1025						1030	
tat	aga	cca	ggc	tac	cat	aca	caa	ata	aca	gga	cca	tgg	cat	cta	ggt	3234
Tyr	Arg	Pro	Gly	Tyr	His	Thr	Gln	Ile	Thr	Gly	Pro	Trp	His	Leu	Gly	
			1035					1040						1045		
aag	ctt	gag	atg	gac	ttt	gat	ttc	tgt	gat	gga	aca	aca	gtg	gta	gtg	3282
Lys	Leu	Glu	Met	Asp	Phe	Asp	Phe	Cys	Asp	Gly	Thr	Thr	Val	Val	Val	
			1050					1055					1060			
act	gag	gac	tgc	gga	aat	aga	gga	ccc	tct	ttg	aga	aca	acc	act	gcc	3330
Thr	Glu	Asp	Cys	Gly	Asn	Arg	Gly	Pro	Ser	Leu	Arg	Thr	Thr	Thr	Ala	
		1065					1070					1075				
tct	gga	aaa	ctc	ata	aca	gaa	tgg	tgc	tgc	cga	tct	tgc	aca	tta	cca	3378
Ser	Gly	Lys	Leu	Ile	Thr	Glu	Trp	Cys	Cys	Arg	Ser	Cys	Thr	Leu	Pro	
	1080					1085						1090				
ccg	cta	aga	tac	aga	ggt	gag	gat	ggg	tgc	tgg	tac	ggg	atg	gaa	atc	3426
Pro	Leu	Arg	Tyr	Arg	Gly	Glu	Asp	Gly	Cys	Trp	Tyr	Gly	Met	Glu	Ile	
	1095				1100				1105					1110		
aga	cca	ttg	aag	gag	aaa	gaa	gag	aat	ttg	gtc	aac	tcc	ttg	gtc	aca	3474
Arg	Pro	Leu	Lys	Glu	Lys	Glu	Glu	Asn	Leu	Val	Asn	Ser	Leu	Val	Thr	
			1115					1120						1125		
gct	gga	cat	ggg	cag	gtc	gac	aac	ttt	tca	cta	gga	gtc	ttg	gga	atg	3522
Ala	Gly	His	Gly	Gln	Val	Asp	Asn	Phe	Ser	Leu	Gly	Val	Leu	Gly	Met	
			1130					1135						1140		
gca	ttg	ttc	ctg	gag	gaa	atg	ctt	agg	acc	cga	gta	gga	acg	aaa	cat	3570
Ala	Leu	Phe	Leu	Glu	Glu	Met	Leu	Arg	Thr	Arg	Val	Gly	Thr	Lys	His	
		1145					1150					1155				
gca	ata	cta	cta	gtt	gca	gtt	tct	ttt	gtg	aca	ttg	atc	aca	ggg	aac	3618
Ala	Ile	Leu	Leu	Val	Ala	Val	Ser	Phe	Val	Thr	Leu	Ile	Thr	Gly	Asn	
	1160					1165						1170				
atg	tcc	ttt	aga	gac	ctg	gga	aga	gtg	atg	gtt	atg	gta	ggc	gcc	act	3666
Met	Ser	Phe	Arg	Asp	Leu	Gly	Arg	Val	Met	Val	Met	Val	Gly	Ala	Thr	
	1175				1180						1185				1190	
atg	acg	gat	gac	ata	ggt	atg	ggc	gtg	act	tat	ctt	gcc	cta	cta	gca	3714
Met	Thr	Asp	Asp	Ile	Gly	Met	Gly	Val	Thr	Tyr	Leu	Ala	Leu	Leu	Ala	
			1195					1200						1205		
gcc	ttc	aaa	gtc	aga	cca	act	ttt	gca	gct	gga	cta	ctc	ttg	aga	aag	3762
Ala	Phe	Lys	Val	Arg	Pro	Thr	Phe	Ala	Ala	Gly	Leu	Leu	Leu	Arg	Lys	
			1210				1215						1220			
ctg	acc	tcc	aag	gaa	ttg	atg	atg	act	act	ata	gga	att	gta	ctc	ctc	3810

Leu	Thr	Ser	Lys	Glu	Leu	Met	Met	Thr	Thr	Ile	Gly	Ile	Val	Leu	Leu	
tcc	cag	agc	acc	ata	cca	gag	acc	att	ctt	gag	ttg	act	gat	gcg	tta	3858
Ser	Gln	Ser	Thr	Ile	Pro	Glu	Thr	Ile	Leu	Glu	Leu	Thr	Asp	Ala	Leu	
1225	1240					1245				1250						
gcc	tta	ggc	atg	atg	gtc	ctc	aaa	atg	gtg	aga	aat	atg	gaa	aag	tat	3906
Ala	Leu	Gly	Met	Met	Val	Leu	Lys	Met	Val	Arg	Asn	Met	Glu	Lys	Tyr	
1255					1260					1265					1270	
caa	ttg	gca	gtg	act	atc	atg	gct	atc	ttg	tgc	gtc	cca	aac	gca	gtg	3954
Gln	Leu	Ala	Val	Thr	Ile	Met	Ala	Ile	Leu	Cys	Val	Pro	Asn	Ala	Val	
1275					1280					1285						
ata	tta	caa	aac	gca	tgg	aaa	gtg	agt	tgc	aca	ata	ttg	gca	gtg	gtg	4002
Ile	Leu	Gln	Asn	Ala	Trp	Lys	Val	Ser	Cys	Thr	Ile	Leu	Ala	Val	Val	
1290					1295							1300				
tcc	gtt	tcc	cca	ctg	ctc	tta	aca	tcc	tca	cag	caa	aaa	aca	gat	tgg	4050
Ser	Val	Ser	Pro	Leu	Leu	Leu	Thr	Ser	Ser	Gln	Gln	Lys	Thr	Asp	Trp	
1305						1310						1315				
ata	cca	tta	gca	ttg	acg	atc	aaa	ggg	ctc	aat	cca	aca	gct	att	ttt	4098
Ile	Pro	Leu	Ala	Leu	Thr	Ile	Lys	Gly	Leu	Asn	Pro	Thr	Ala	Ile	Phe	
1320					1325					1330						
cta	aca	acc	ctc	tca	aga	acc	agc	aag	aaa	agg	agc	tgg	cca	tta	aat	4146
Leu	Thr	Thr	Leu	Ser	Arg	Thr	Ser	Lys	Lys	Arg	Ser	Trp	Pro	Leu	Asn	
1335					1340					1345					1350	
gag	gct	atc	atg	gca	gtc	ggg	atg	gtg	agc	att	tta	gcc	agt	tct	ctc	4194
Glu	Ala	Ile	Met	Ala	Val	Gly	Met	Val	Ser	Ile	Leu	Ala	Ser	Ser	Leu	
1355					1360							1365				
cta	aaa	aat	gat	att	ccc	atg	aca	gga	cca	tta	gtg	gct	gga	ggg	ctc	4242
Leu	Lys	Asn	Asp	Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	
1370					1375							1380				
ctc	act	gtg	tgc	tac	gtg	ctc	act	gga	cga	tgc	gcc	gat	ttg	gaa	ctg	4290
Leu	Thr	Val	Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu	
1385					1390							1395				
gag	aga	gca	gcc	gat	gtc	aaa	tgg	gaa	gac	cag	gca	gag	ata	tca	gga	4338
Glu	Arg	Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	
1400					1405					1410						
agc	agt	cca	atc	ctg	tca	ata	aca	ata	tca	gaa	gat	ggg	agc	atg	tgc	4386
Ser	Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser	
1415					1420					1425					1430	
ata	aaa	aat	gaa	gag	gaa	gaa	caa	aca	ctg	acc	ata	ctc	att	aga	aca	4434
Ile	Lys	Asn	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg	Thr		
1435					1440							1445				
gga	ttg	ctg	gtg	atc	tca	gga	ctt	ttt	cct	gta	tca	ata	cca	atc	acg	4482
Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro	Ile	Thr	
1450					1455							1460				
gca	gca	gca	tgg	tac	ctg	tgg	gaa	gtg	aag	aaa	caa	cgg	gcc	gga	gta	4530
Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg	Ala	Gly	Val	
1465					1470							1475				
ttg	tgg	gat	gtt	cct	tca	ccc	cca	ccc	atg	gga	aag	gct	gaa	ctg	gaa	4578
Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys	Ala	Glu	Leu	Glu	
1480					1485					1490						
gat	gga	gcc	tat	aga	att	aag	caa	aaa	ggg	att	ctt	gga	tat	tcc	cag	4626
Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile	Leu	Gly	Tyr	Ser	Gln	
1495					1500					1505					1510	
atc	gga	gcc	gga	gtt	tac	aaa	gaa	gga	aca	ttc	cat	aca	atg	tgg	cat	4674
Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr	Phe	His	Thr	Met	Trp	His	
1515					1520							1525				
gtc	aca	cgt	ggc	gct	gtt	cta	atg	cat	aaa	gga	aag	agg	att	gaa	cca	4722
Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	
1530					1535							1540				
tca	tgg	gcg	gac	gtc	aag	aaa	gac	cta	ata	tca	tat	gga	gga	ggc	tgg	4770
Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	
1545					1550							1555				
aag	tta	gaa	gga	gaa	tgg	aag	gaa	gga	gaa	gaa	gtc	cag	gta	ttg	gca	4818
Lys	Leu	Glu	Gly	Glu	Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	
1560					1565					1570						
ctg	gag	cct	gga	aaa	aat	cca	aga	gcc	gtc	caa	acg	aaa	cct	ggg	ctt	4866
Leu	Glu	Pro	Gly	Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	
1575					1580					1585					1590	
ttc	aaa	acc	aac	gcc	gga	aca	ata	ggg	gct	gta	tct	ctg	gac	ttt	tct	4914

Phe	Lys	Thr	Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	
cct	gga	acg	tca	gga	tct	cca	att	atc	gac	aaa	aaa	gga	aaa	gtt	gtg	4962
Pro	Gly	Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	
ggt	ctt	tat	ggt	aat	ggt	gtt	gtt	aca	agg	agt	gga	gca	tat	gtg	agt	5010
Gly	Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser	
gct	ata	gcc	cag	act	gaa	aaa	agc	att	gaa	gac	aac	cca	gag	atc	gaa	5058
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile	Glu	
gat	gac	att	ttc	cga	aag	aga	aga	ctg	acc	atc	atg	gac	ctc	cac	cca	5106
Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu	His	Pro	
gga	gcg	gga	aag	acg	aag	aga	tac	ctt	ccg	gcc	ata	gtc	aga	gaa	gct	5154
Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val	Arg	Glu	Ala	
ata	aaa	cgg	ggt	ttg	aga	aca	tta	atc	ttg	gcc	ccc	act	aga	gtt	gtg	5202
Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro	Thr	Arg	Val	Val	
gca	gct	gaa	atg	gag	gaa	gcc	ctt	aga	gga	ctt	cca	ata	aga	tac	cag	5250
Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu	Pro	Ile	Arg	Tyr	Gln	
acc	cca	gcc	atc	aga	gct	gag	cac	acc	ggg	cgg	gag	att	gtg	gac	cta	5298
Thr	Pro	Ala	Ile	Arg	Ala	Glu	His	Thr	Gly	Arg	Glu	Ile	Val	Asp	Leu	
atg	tgt	cat	gcc	aca	ttt	acc	atg	agg	ctg	cta	tca	cca	gtt	aga	gtg	5346
Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg	Leu	Leu	Ser	Pro	Val	Arg	Val	
cca	aac	tac	aac	ctg	att	atc	atg	gac	gaa	gcc	cat	ttc	aca	gac	cca	5394
Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met	Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	
gca	agt	ata	gca	gct	aga	gga	tac	atc	tca	act	cga	gtg	gag	atg	ggt	5442
Ala	Ser	Ile	Ala	Ala	Arg	Gly	Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	
gag	gca	gct	ggg	att	ttt	atg	aca	gcc	act	ccc	ccg	gga	agc	aga	gac	5490
Glu	Ala	Ala	Gly	Ile	Phe	Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	
cca	ttt	cct	cag	agc	aat	gca	cca	atc	ata	gat	gaa	gaa	aga	gaa	atc	5538
Pro	Phe	Pro	Gln	Ser	Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	
cct	gaa	cgt	tcg	tgg	aat	tcc	gga	cat	gaa	tgg	gtc	acg	gat	ttt	aaa	5586
Pro	Glu	Arg	Ser	Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	
ggg	aag	act	gtt	tgg	ttc	gtt	cca	agt	ata	aaa	gca	gga	aat	gat	ata	5634
Gly	Lys	Thr	Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	
gca	gct	tgc	ctg	agg	aaa	aat	gga	aag	aaa	gtg	ata	caa	ctc	agt	agg	5682
Ala	Ala	Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	
aag	acc	ttt	gat	tct	gag	tat	gtc	aag	act	aga	acc	aat	gat	tgg	gac	5730
Lys	Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	
ttc	gtg	gtt	aca	act	gac	att	tca	gaa	atg	ggt	gcc	aat	ttc	aag	gct	5778
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	
gag	agg	gtt	ata	gac	ccc	aga	cgc	tgc	atg	aaa	cca	gtc	ata	cta	aca	5826
Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr	
gat	ggt	gaa	gag	cgg	gtg	att	ctg	gca	gga	cct	atg	cca	gtg	acc	cac	5874
Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr	His	
tct	agt	gca	gca	caa	aga	aga	ggg	aga	ata	gga	aga	aat	cca	aaa	aat	5922
Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro	Lys	Asn	
gag	aat	gac	cag	tac	ata	tac	atg	ggg	gaa	cct	ctg	gaa	aat	gat	gaa	5970
Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu	Asn	Asp	Glu	
gac	tgt	gca	cac	tgg	aaa	gaa	gct	aaa	atg	ctc	cta	gat	aac	atc	aac	6018

Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu	Asp	Asn	Ile	Asn	
1960	1960					1965				1970						
acg	cca	gaa	gga	atc	att	cct	agc	atg	ttc	gaa	cca	gag	cgt	gaa	aag	6066
Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu	Pro	Glu	Arg	Glu	Lys	
1975					1980					1985					1990	
gtg	gat	gcc	att	gat	ggc	gaa	tac	cgc	ttg	aga	gga	gaa	gca	agg	aaa	6114
Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr	Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	
				1995					2000						2005	
acc	ttt	gta	gac	tta	atg	aga	aga	gga	gac	cta	cca	gtc	tgg	ttg	gcc	6162
Thr	Phe	Val	Asp	Leu	Met	Arg	Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	
			2010						2015					2020		
tac	aga	gtg	gca	gct	gaa	ggc	atc	aac	tac	gca	gac	aga	agg	tgg	tgt	6210
Tyr	Arg	Val	Ala	Ala	Glu	Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	
			2025				2030						2035			
ttt	gat	gga	gtc	aag	aac	aac	caa	atc	cta	gaa	gaa	aac	gtg	gaa	gtt	6258
Phe	Asp	Gly	Val	Lys	Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	
	2040				2045					2050						
gaa	atc	ttg	aca	aaa	gaa	ggg	gaa	agg	aag	aaa	ttg	aaa	ccc	aga	ttg	6306
Glu	Ile	Trp	Thr	Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	
	2055				2060					2065					2070	
ttg	gat	gct	agg	atc	tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	6354
Leu	Asp	Ala	Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	
			2075					2080						2085		
gaa	ttt	gca	gcc	gga	aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	6402
Glu	Phe	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu		
			2090					2095					2100			
atg	ggt	agg	ctc	cca	acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	6450
Met	Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu	
			2105				2110						2115			
gac	aac	tta	gca	gtg	ctg	cac	acg	gct	gag	gca	ggt	gga	agg	gcg	tac	6498
Asp	Asn	Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr	
	2120					2125					2130					
aac	cat	gct	ctc	agt	gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	6546
Asn	His	Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu	
	2135				2140					2145					2150	
ctg	aca	ctt	ctg	gct	aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	6594
Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met	
			2155					2160						2165		
agc	gga	agg	ggc	ata	ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	6642
Ser	Gly	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile	
			2170					2175					2180			
acg	gct	agc	atc	ctc	cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	6690
Thr	Ala	Ser	Leu	Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile		
		2185			2190						2195					
gca	gct	tca	ata	ata	ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	6738
Ala	Ala	Ser	Ile	Ile	Leu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro		
		2200			2205					2210						
gaa	cct	gaa	aaa	cag	aga	aca	ccc	caa	gac	aac	caa	ctg	acc	tac	gtt	6786
Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln	Asp	Asn	Gln	Leu	Thr	Tyr	Val	
	2215				2220					2225					2230	
gtc	ata	gcc	atc	ctc	aca	gtg	gtg	gcc	gca	acc	atg	gca	aac	gag	atg	6834
Val	Ile	Ala	Ile	Leu	Thr	Val	Val	Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	
			2235					2240						2245		
ggt	ttc	cta	gaa	aaa	acg	aag	aaa	gat	ctc	gga	ttg	gga	agc	att	gca	6882
Gly	Phe	Leu	Glu	Lys	Thr	Lys	Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	
			2250					2255					2260			
acc	cag	caa	ccc	gag	agc	aac	atc	ctg	gac	ata	gat	cta	cgt	cct	gca	6930
Thr	Gln	Gln	Pro	Glu	Ser	Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	
		2265				2270						2275				
tca	gca	ttg	acg	ctg	tat	gcc	gtg	gcc	aca	aca	ttt	gtt	aca	cca	atg	6978
Ser	Ala	Trp	Thr	Leu	Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	
		2280			2285						2290					
ttg	aga	cat	agc	att	gaa	aat	tcc	tca	gtg	aat	gtg	tcc	cta	aca	gct	7026
Leu	Arg	His	Ser	Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	
	2295				2300				2305						2310	
ata	gcc	aac	caa	gcc	aca	gtg	tta	atg	ggt	ctc	ggg	aaa	gga	ttg	cca	7074
Ile	Ala	Asn	Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	
			2315					2320						2325		
ttg	tca	aag	atg	gac	atc	gga	gtt	ccc	ctt	ctc	gcc	att	gga	tgc	tac	7122

Leu	Ser	Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	
tca	caa	gtc	aac	ccc	ata	act	ctc	aca	gca	gct	ctt	ttc	tta	ttg	gta	7170
Ser	Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	
		2345					2350					2355				
gca	cat	tat	gcc	atc	ata	ggg	cca	gga	ctc	caa	gca	aaa	gca	acc	aga	7218
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	
		2360				2365					2370					
gaa	gct	cag	aaa	aga	gca	gcg	gcg	ggc	atc	atg	aaa	aac	cca	act	gtc	7266
Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val	
		2375			2380					2385					2390	
gat	gga	ata	aca	gtg	att	gac	cta	gat	cca	ata	cct	tat	gat	cca	aag	7314
Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr	Asp	Pro	Lys	
				2395					2400					2405		
ttt	gaa	aag	cag	ttg	gga	caa	gta	atg	ctc	cta	gtc	ctc	tgc	gtg	act	7362
Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val	Leu	Cys	Val	Thr	
				2410				2415					2420			
caa	gta	ttg	atg	agg	act	aca	tgg	gct	ctg	tgt	gag	gct	tta	acc		7410
Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu	Cys	Glu	Ala	Leu	Thr	
		2425					2430					2435				
tta	gct	acc	ggg	ccc	atc	tcc	aca	ttg	tgg	gaa	gga	aat	cca	ggg	agg	7458
Leu	Ala	Thr	Gly	Pro	Ile	tcc	Thr	Leu	Trp	Glu	Gly	Asn	Pro	Gly	Arg	
		2440				2445					2450					
ttt	tgg	aac	act	acc	att	gcg	gtg	tca	atg	gct	aac	att	ttt	aga	ggg	7506
Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser	Met	Ala	Asn	Ile	Phe	Arg	Gly	
		2455				2460				2465					2470	
agt	tac	ttg	gcc	gga	gct	gga	ctt	ctc	ttt	tct	att	atg	aag	aac	aca	7554
Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu	Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	
				2475				2480					2485			
acc	aac	aca	aga	agg	gga	act	ggc	aac	ata	gga	gag	acg	ctt	gga	gag	7602
Thr	Asn	Thr	Arg	Arg	Gly	Thr	Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	
			2490				2495						2500			
aaa	tgg	aaa	agc	cga	ttg	aac	gca	ttg	gga	aaa	agt	gaa	ttc	cag	atc	7650
Lys	Trp	Lys	Ser	Arg	Leu	Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	
		2505					2510					2515				
tac	aag	aaa	agt	gga	atc	cag	gaa	gtg	gat	aga	acc	tta	gca	aaa	gaa	7698
Tyr	Lys	Lys	Ser	Gly	Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	
		2520				2525					2530					
ggc	att	aaa	aga	gga	gaa	acg	gac	cat	cac	gct	gtg	tcg	cga	ggc	tca	7746
Gly	Ile	Lys	Arg	Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	
		2535			2540					2545					2550	
gca	aaa	ctg	aga	tgg	ttc	gtt	gag	aga	aac	atg	gtc	aca	cca	gaa	ggg	7794
Ala	Lys	Leu	Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	
				2555				2560						2565		
aaa	gta	gtg	gac	ctc	ggt	tgt	ggc	aga	gga	ggc	tgg	tca	tac	tat	tgt	7842
Lys	Val	Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	
			2570				2575						2580			
gga	gga	cta	aag	aat	gta	aga	gaa	gtc	aaa	ggc	cta	aca	aaa	gga	gga	7890
Gly	Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly	
		2585					2590					2595				
cca	gga	cac	gaa	gaa	ccc	atc	ccc	atg	tca	aca	tat	ggg	tgg	aat	cta	7938
Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn	Leu	
		2600				2605					2610					
gtg	cgt	ctt	caa	agt	gga	gtt	gac	gtt	ttc	ttc	atc	ccg	cca	gaa	aag	7986
Val	Arg	Leu	Gln	Ser	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro	Glu	Lys	
		2615			2620					2625					2630	
tgt	gac	aca	tta	ttg	gtg	gac	ata	ggg	gag	tca	tca	cca	aat	ccc	aca	8034
Cys	Asp	Thr	Leu	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro	Thr	
				2635				2640						2645		
gtg	gaa	gca	gga	cga	aca	ctc	aga	gtc	ctt	aac	tta	gta	gaa	aat	tgg	8082
Val	Glu	Ala	Gly	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu	Asn	Trp	
		2650					2655					2660				
ttg	aac	aac	aac	act	caa	ttt	tgc	ata	aag	gtt	ctc	aac	cca	tat	atg	8130
Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn	Pro	Tyr	Met	
		2665					2670					2675				
ccc	tca	gtc	ata	gaa	aaa	atg	gaa	gca	cta	caa	agg	aaa	tat	gga	gga	8178
Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg	Lys	Tyr	Gly	Gly	
		2680				2685					2690					
gcc	tta	gtg	agg	aat	cca	ctc	tca	cga	aac	tcc	aca	cat	gag	atg	tac	8226

Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser	Thr	His	Glu	Met	Tyr	
2695					2700					2705					2710	
tgg	gta	tcc	aat	gct	tcc	ggg	aac	ata	gtg	tca	tca	gtg	aac	atg	att	8274
Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val	Ser	Ser	Val	Asn	Met	Ile	
				2715					2720						2725	
tca	agg	atg	ttg	atc	aac	aga	ttt	aca	atg	aga	tac	aag	aaa	gcc	act	8322
Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	
			2730					2735					2740			
tac	gag	ccg	gat	gtt	gac	ctc	gga	agc	gga	acc	cgt	aac	atc	ggg	att	8370
Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	
		2745					2750					2755				
gaa	agt	gag	ata	cca	aac	cta	gat	ata	att	ggg	aaa	aga	ata	gaa	aaa	8418
Glu	Ser	Glu	Ile	Pro	Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	
	2760				2765					2770						
ata	aag	caa	gag	cat	gaa	aca	tca	tgg	cac	tat	gac	caa	gac	cac	cca	8466
Ile	Lys	Gln	Glu	His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	
	2775				2780					2785					2790	
tac	aaa	acg	tgg	gca	tac	cat	ggt	agc	tat	gaa	aca	aaa	cag	act	gga	8514
Tyr	Lys	Thr	Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	
			2795						2800						2805	
tca	gca	tca	tcc	atg	gtc	aac	gga	gtg	gtc	agg	ctg	ctg	aca	aaa	cct	8562
Ser	Ala	Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	
			2810					2815					2820			
tgg	gac	gtc	gtc	ccc	atg	gtg	aca	cag	atg	gca	atg	aca	gac	acg	act	8610
Trp	Asp	Val	Val	Pro	Met	Val	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr		
		2825					2830					2835				
cca	ttt	gga	caa	cag	cgc	gtt	ttt	aaa	gag	aaa	gtg	gac	acg	aga	acc	8658
Pro	Phe	Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	
	2840					2845					2850					
caa	gaa	ccg	aaa	gaa	ggc	acg	aag	aaa	cta	atg	aaa	ata	aca	gca	gag	8706
Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu	
	2855				2860				2865						2870	
tgg	ctt	tgg	aaa	gaa	tta	ggg	aag	aaa	aag	aca	ccc	agg	atg	tgc	acc	8754
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys	Thr	
			2875						2880						2885	
aga	gaa	gaa	ttc	aca	aga	aag	gtg	aga	agc	aat	gca	gcc	ttg	ggg	gcc	8802
Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu	Gly	Ala	
			2890					2895					2900			
ata	ttc	act	gat	gag	aac	aag	tgg	aag	tcg	gca	cgt	gag	gct	gtt	gaa	8850
Ile	Phe	Thr	Asp	Glu	Asn	Lys	Trp	Lys	Ser	Ala	Arg	Glu	Ala	Val	Glu	
		2905					2910					2915				
gat	agt	agg	ttt	tgg	gag	ctg	gtt	gac	aag	gaa	agg	aat	ctc	cat	ctt	8898
Asp	Ser	Arg	Phe	Trp	Glu	Val	Val	Asp	Lys	Glu	Arg	Asn	Leu	His	Leu	
		2920				2925					2930					
gaa	gga	aag	tgt	gaa	aca	tgt	gtg	tac	aac	atg	atg	gga	aaa	aga	gag	8946
Glu	Gly	Lys	Cys	Glu	Thr	Cys	Val	Tyr	Asn	Met	Met	Gly	Lys	Arg	Glu	
	2935				2940					2945					2950	
aag	aag	cta	ggg	gaa	ttc	ggc	aag	gca	aaa	ggc	agc	aga	gcc	ata	tgg	8994
Lys	Lys	Leu	Gly	Glu	Phe	Gly	Lys	Ala	Lys	Gly	Ser	Arg	Ala	Ile	Trp	
			2955					2960							2965	
tac	atg	tgg	ctt	gga	gca	cgc	ttc	tta	gag	ttt	gaa	gcc	cta	gga	ttc	9042
Tyr	Met	Trp	Leu	Gly	Ala	Arg	Phe	Leu	Glu	Phe	Glu	Ala	Leu	Gly	Phe	
			2970					2975					2980			
tta	aat	gaa	gat	cac	tgg	ttc	tcc	aga	gag	aac	tcc	ctg	agt	gga	gtg	9090
Leu	Asn	Glu	Asp	His	Trp	Phe	Ser	Arg	Glu	Asn	Ser	Leu	Ser	Gly	Val	
		2985					2990					2995				
gaa	gga	gaa	ggg	ctg	cac	aag	cta	ggt	tac	att	cta	aga	gac	gtg	agc	9138
Glu	Gly	Glu	Gly	Leu	His	Lys	Leu	Gly	Tyr	Ile	Leu	Arg	Asp	Val	Ser	
	3000					3005				3010						
aag	aaa	gag	gga	gga	gca	atg	tat	gcc	gat	gac	acc	gca	gga	tgg	gat	9186
Lys	Lys	Glu	Gly	Gly	Ala	Met	Tyr	Ala	Asp	Asp	Thr	Ala	Gly	Trp	Asp	
	3015				3020					3025					3030	
aca	aga	atc	aca	cta	gaa	gac	cta	aaa	aat	gaa	gaa	atg	gta	aca	aac	9234
Thr	Arg	Ile	Thr	Leu	Glu	Asp	Leu	Lys	Asn	Glu	Glu	Met	Val	Thr	Asn	
			3035					3040							3045	
cac	atg	gaa	gga	gaa	cac	aag	aaa	cta	gcc	gag	gcc	att	ttc	aaa	cta	9282
His	Met	Glu	Gly	Glu	His	Lys	Lys	Leu	Ala	Glu	Ala	Ile	Phe	Lys	Leu	
			3050					3055					3060			
acg	tac	caa	aac	aag	gtg	gtg	cgt	gtg	caa	aga	cca	aca	cca	aga	ggc	9330

Thr	Tyr	Gln	Asn	Lys	Val	Val	Arg	Val	Gln	Arg	Pro	Thr	Pro	Arg	Gly	
		3065					3070					3075				
aca	gta	atg	gac	atc	ata	tcg	aga	aga	gac	caa	aga	ggg	agt	gga	caa	9378
Thr	Val	Met	Asp	Ile	Ile	Ser	Arg	Arg	Asp	Gln	Arg	Gly	Ser	Gly	Gln	
		3080					3085					3090				
ggt	ggc	acc	tat	gga	ctc	aat	act	ttc	acc	aat	atg	gaa	gcc	caa	cta	9426
Val	Gly	Thr	Tyr	Gly	Leu	Asn	Thr	Phe	Thr	Asn	Met	Glu	Ala	Gln	Leu	
		3095					3100				3105				3110	
atc	aga	cag	atg	gag	gga	gaa	gga	gtc	ttt	aaa	agc	att	cag	cac	cta	9474
Ile	Arg	Gln	Met	Glu	Gly	Glu	Gly	Val	Phe	Lys	Ser	Ile	Gln	His	Leu	
				3115					3120					3125		
aca	atc	aca	gaa	atc	gct	gtg	caa	aac	tgg	tta	gca	aga	gtg	ggg		9522
Thr	Ile	Thr	Glu	Glu	Ile	Ala	Val	Gln	Asn	Trp	Leu	Ala	Arg	Val	Gly	
			3130				3135					3140				
cgc	gaa	agg	tta	tca	aga	atg	gcc	atc	agt	gga	gat	gat	tgt	gtt	gtg	9570
Arg	Glu	Arg	Leu	Ser	Arg	Met	Ala	Ile	Ser	Gly	Asp	Asp	Cys	Val	Val	
		3145					3150				3155					
aaa	cct	tta	gat	gac	agg	ttc	gca	agc	gct	tta	aca	gct	cta	aat	gac	9618
Lys	Pro	Leu	Asp	Asp	Arg	Phe	Ala	Ser	Ala	Leu	Thr	Ala	Leu	Asn	Asp	
		3160				3165					3170					
atg	gga	aag	att	agg	aaa	gac	ata	caa	caa	tgg	gaa	cct	tca	aga	gga	9666
Met	Gly	Lys	Ile	Arg	Lys	Asp	Ile	Gln	Gln	Trp	Glu	Pro	Ser	Arg	Gly	
		3175				3180				3185					3190	
tgg	aat	gat	tgg	aca	caa	gtg	ccc	ttc	tgt	tca	cac	cat	ttc	cat	gag	9714
Trp	Asn	Asp	Trp	Thr	Gln	Val	Pro	Phe	Cys	Ser	His	His	Phe	His	Glu	
				3195					3200					3205		
tta	atc	atg	aaa	gac	ggg	cgc	gta	ctc	gtt	gtt	cca	tgt	aga	aac	caa	9762
Leu	Ile	Met	Lys	Asp	Gly	Arg	Val	Leu	Val	Val	Pro	Cys	Arg	Asn	Gln	
			3210					3215					3220			
gat	gaa	ctg	att	ggc	aga	gcc	cga	atc	tcc	caa	gga	gca	ggg	tgg	tct	9810
Asp	Glu	Leu	Ile	Gly	Arg	Ala	Arg	Ile	Ser	Gln	Gly	Ala	Gly	Trp	Ser	
		3225					3230					3235				
ttg	cgg	gag	acg	gcc	tgt	ttg	ggg	aag	tct	tac	gcc	caa	atg	tgg	agc	9858
Leu	Arg	Glu	Thr	Ala	Cys	Leu	Gly	Lys	Ser	Tyr	Ala	Gln	Met	Trp	Ser	
		3240				3245					3250					
ttg	atg	tac	ttc	cac	aga	cgc	gac	ctc	agg	ctg	ggg	gca	aat	gct	att	9906
Leu	Met	Tyr	Phe	His	Arg	Arg	Asp	Leu	Arg	Leu	Ala	Ala	Asn	Ala	Ile	
		3255			3260					3265					3270	
tgc	tcg	gca	gta	cca	tca	cat	tgg	gtt	cca	aca	agt	cga	aca	acc	tgg	9954
Cys	Ser	Ala	Val	Pro	Ser	His	Trp	Val	Pro	Thr	Ser	Arg	Thr	Thr	Trp	
				3275					3280					3285		
tcc	ata	cat	gct	aaa	cat	gaa	tgg	atg	aca	acg	gaa	gac	atg	ctg	aca	10002
Ser	Ile	His	Ala	Lys	His	Glu	Trp	Met	Thr	Thr	Glu	Asp	Met	Leu	Thr	
			3290					3295					3300			
gtc	tgg	aac	agg	gtg	tgg	att	caa	gaa	aac	cca	tgg	atg	gaa	gac	aaa	10050
Val	Trp	Asn	Arg	Val	Trp	Ile	Gln	Glu	Asn	Pro	Trp	Met	Glu	Asp	Lys	
		3305					3310					3315				
act	cca	gtg	gaa	tca	tgg	gag	gaa	atc	cca	tac	ttg	ggg	aaa	aga	gaa	10098
Thr	Pro	Val	Glu	Ser	Trp	Glu	Glu	Ile	Pro	Tyr	Leu	Gly	Lys	Arg	Glu	
		3320				3325					3330					
gac	caa	tgg	tgc	ggc	tca	ttg	att	ggg	tta	aca	agc	agg	gcc	acc	tgg	10146
Asp	Gln	Trp	Cys	Gly	Ser	Leu	Ile	Gly	Leu	Thr	Ser	Arg	Ala	Thr	Trp	
		3335			3340				3345						3350	
gca	aag	aac	atc	caa	gca	gca	ata	aat	caa	gtt	aga	tcc	ctt	ata	ggc	10194
Ala	Lys	Asn	Ile	Gln	Ala	Ala	Ile	Asn	Gln	Val	Arg	Ser	Leu	Ile	Gly	
			3355					3360					3365			
aat	gaa	gaa	tac	aca	gat	tac	atg	cca	tcc	atg	aaa	aga	ttc	aga	aga	10242
Asn	Glu	Glu	Tyr	Thr	Asp	Tyr	Met	Pro	Ser	Met	Lys	Arg	Phe	Arg	Arg	
			3370					3375					3380			
gaa	gag	gaa	gaa	gca	gga	gtt	ctg	tgg	tag	aaagcaaaac	taacatgaaa					10292
Glu	Glu	Glu	Glu	Ala	Gly	Val	Leu	Trp	*							
		3385				3390										
caaggctaga	agtcaggctcg	gattaagcca	tagtacggaa	aaaactatgc	tacctgtgag											10352
ccccgtccaa	ggacgttaaa	agaagtcagg	ccatcataaa	tgccatagct	tgagtaaaact											10412
atgcagctg	tagctccacc	tgagaagggt	taaaaaatcc	gggaggccac	aaaccatgga											10472
agctgtacgc	atggcgtagt	ggactagcgg	ttagaggaga	ccccctccctt	acaaatcgca											10532
gcaacaatgg	gggcccagg	cgagatgaag	ctgtagtctc	gctggaagga	ctagagggtta											10592
gaggagaccc	ccccgaaaca	aaaaacagca	tattgacgct	gggaaagacc	agagatcctg											10652
ctgtctcttc	agcatcattc	caggcacaga	acgccagaaa	atggaatggt	gctgttgaat											10712

caacagggttc t

10723

EP2278012B1

<210> SEQ ID NO 14

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 14

Met	Asn	Asn	Gln	Arg	Lys	Lys	Ala	Lys	Asn	Thr	Pro	Phe	Asn	Met	Leu
1				5					10					15	
Lys	Arg	Glu	Arg	Asn	Arg	Val	Ser	Thr	Val	Gln	Gln	Leu	Thr	Lys	Arg
			20					25					30		
Phe	Ser	Leu	Gly	Met	Leu	Gln	Gly	Arg	Gly	Pro	Leu	Lys	Leu	Phe	Met
		35					40					45			
Ala	Leu	Val	Ala	Phe	Leu	Arg	Phe	Leu	Thr	Ile	Pro	Pro	Thr	Ala	Gly
	50					55					60				
Ile	Leu	Lys	Arg	Trp	Gly	Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val
65					70					75					80
Leu	Arg	Gly	Phe	Arg	Lys	Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn
				85					90					95	
Arg	Arg	Arg	Arg	Ser	Ala	Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val
				100				105					110		
Met	Ala	Phe	His	Leu	Thr	Thr	Arg	Asn	Gly	Glu	Pro	His	Met	Ile	Val
		115					120					125			
Ser	Arg	Gln	Glu	Lys	Gly	Lys	Ser	Leu	Leu	Phe	Lys	Thr	Glu	Asp	Gly
	130					135					140				
Val	Asn	Met	Cys	Thr	Leu	Met	Ala	Met	Asp	Leu	Gly	Glu	Leu	Cys	Glu
145					150					155					160
Asp	Thr	Ile	Thr	Tyr	Lys	Cys	Pro	Leu	Leu	Arg	Gln	Asn	Glu	Pro	Glu
				165					170					175	
Asp	Ile	Asp	Cys	Trp	Cys	Asn	Ser	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly
			180					185					190		
Thr	Cys	Thr	Thr	Met	Gly	Glu	His	Arg	Arg	Glu	Lys	Arg	Ser	Val	Ala
		195					200					205			
Leu	Val	Pro	His	Val	Gly	Met	Gly	Leu	Glu	Thr	Arg	Thr	Glu	Thr	Trp
	210					215					220				
Met	Ser	Ser	Glu	Gly	Ala	Trp	Lys	His	Val	Gln	Arg	Ile	Glu	Thr	Trp
225					230					235					240
Ile	Leu	Arg	His	Pro	Gly	Phe	Thr	Met	Met	Ala	Ala	Ile	Leu	Ala	Tyr
				245					250					255	
Thr	Ile	Gly	Thr	Thr	His	Phe	Gln	Arg	Ala	Leu	Ile	Phe	Ile	Leu	Leu
			260					265					270		
Thr	Ala	Val	Thr	Pro	Ser	Met	Thr	Met	Arg	Cys	Ile	Gly	Met	Ser	Asn
		275					280					285			
Arg	Asp	Phe	Val	Glu	Gly	Val	Ser	Gly	Gly	Ser	Trp	Val	Asp	Ile	Val
	290					295					300				
Leu	Glu	His	Gly	Ser	Cys	Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr
305					310					315					320
Leu	Asp	Phe	Glu	Leu	Ile	Lys	Thr	Glu	Ala	Lys	Gln	Pro	Ala	Thr	Leu
				325					330					335	
Arg	Lys	Tyr	Cys	Ile	Glu	Ala	Lys	Leu	Thr	Asn	Thr	Thr	Thr	Glu	Ser
			340					345					350		
Arg	Cys	Pro	Thr	Gln	Gly	Glu	Pro	Ser	Leu	Asn	Glu	Glu	Gln	Asp	Lys
		355					360					365			
Arg	Phe	Val	Cys	Lys	His	Ser	Met	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly
	370					375					380				
Cys	Gly	Leu	Phe	Gly	Lys	Gly	Gly	Ile	Val	Thr	Cys	Ala	Met	Phe	Arg
385					390					395					400
Cys	Lys	Lys	Asn	Met	Glu	Gly	Lys	Val	Val	Gln	Pro	Glu	Asn	Leu	Glu
			405					410						415	
Tyr	Thr	Ile	Val	Ile	Thr	Pro	His	Ser	Gly	Glu	Glu	His	Ala	Val	Gly
			420					425					430		
Asn	Asp	Thr	Gly	Lys	His	Gly	Lys	Glu	Ile	Lys	Ile	Thr	Pro	Gln	Ser
		435					440					445			
Ser	Ile	Thr	Glu	Ala	Glu	Leu	Thr	Gly	Tyr	Gly	Thr	Val	Thr	Met	Glu
	450					455					460				
Cys	Ser	Pro	Arg	Thr	Gly	Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Gln

465	Met	Glu	Asn	Lys	Ala	470	Trp	Leu	Val	His	Arg	475	Gln	Trp	Phe	Leu	Asp	480	Leu
	Pro	Leu	Pro	Trp	485	Pro	Gly	Ala	Asp	490	Thr	Gln	Gly	Ser	Asn	495	Trp	Ile	
	Gln	Lys	Glu	Thr	500	Leu	Val	Thr	Phe	505	Lys	Asn	Pro	His	Ala	510	Lys	Lys	Gln
	Asp	Val	Val	Val	515	Leu	Gly	Ser	Gln	520	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	
	Thr	Gly	Ala	Thr	530	Glu	Ile	Gln	Met	535	Ser	Ser	Gly	Asn	Leu	Leu	Phe	Thr	
	Gly	His	Leu	Lys	545	Cys	Arg	Leu	Arg	550	Met	Asp	Lys	Leu	Gln	Leu	Lys	Gly	
	Met	Ser	Tyr	Ser	565	Met	Cys	Thr	Gly	570	Lys	Phe	Lys	Val	Val	Lys	Glu	Ile	
	Ala	Glu	Thr	Gln	580	His	Gly	Thr	Ile	585	Val	Ile	Arg	Val	Gln	Tyr	Glu	Gly	
	Asp	Gly	Ser	Pro	595	Cys	Lys	Ile	Pro	600	Phe	Glu	Ile	Met	Asp	Leu	Glu	Lys	
	Arg	His	Val	Leu	610	Gly	Arg	Leu	Ile	615	Thr	Val	Asn	Pro	Ile	Val	Thr	Glu	
	Lys	Asp	Ser	Pro	625	Val	Asn	Ile	Glu	630	Ala	Glu	Pro	Pro	Phe	Gly	Asp	Ser	
	Tyr	Ile	Ile	Ile	645	Gly	Val	Glu	Pro	650	Gly	Gln	Leu	Lys	Leu	Asn	Trp	Phe	
	Lys	Lys	Gly	Ser	660	Ser	Ile	Gly	Gln	665	Met	Phe	Glu	Thr	Thr	Met	Arg	Gly	
	Ala	Lys	Arg	Met	675	Ala	Ile	Leu	Gly	680	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	
	Leu	Gly	Gly	Val	690	Phe	Thr	Ser	Ile	695	Gly	Lys	Ala	Leu	His	Gln	Val	Phe	
	Gly	Ala	Ile	Tyr	705	Gly	Ala	Ala	Phe	710	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	
	Ile	Leu	Ile	Gly	725	Val	Ile	Ile	Thr	730	Trp	Ile	Gly	Met	Asn	Ser	Arg	Ser	
	Thr	Ser	Leu	Ser	740	Val	Thr	Leu	Val	745	Leu	Val	Gly	Ile	Val	Thr	Leu	Tyr	
	Leu	Gly	Val	Met	755	Val	Gln	Ala	Asp	760	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	
	Asn	Lys	Glu	Leu	770	Lys	Cys	Gly	Ser	775	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	
	His	Thr	Trp	Thr	785	Glu	Tyr	Lys	Phe	790	Gln	Pro	Glu	Ser	Pro	Ser	Lys		
	Leu	Ala	Ser	Ala	805	Ile	Gln	Lys	Ala	810	Glu	Glu	Gly	Ile	Cys	Gly	Ile		
	Arg	Ser	Val	Thr	820	Arg	Leu	Glu	Asn	825	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	
	Glu	Leu	Asn	His	835	Ile	Leu	Ser	Glu	840	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	
	Thr	Gly	Asp	Ile	850	Lys	Gly	Ile	Met	855	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	
	Pro	Gln	Pro	Thr	865	Glu	Leu	Lys	Tyr	870	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	
	Lys	Met	Leu	Ser	885	Thr	Glu	Ser	His	890	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	
	Pro	Glu	Thr	Ala	900	Glu	Cys	Pro	Asn	905	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	
	Glu	Val	Glu	Asp	915	Tyr	Gly	Phe	Gly	920	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	
	Lys	Leu	Lys	Glu	930	Lys	Gln	Asp	Val	935	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	
	Ala	Ala	Ile	Lys	945	Asp	Asn	Arg	Ala	950	Val	His	Ala	Asp	Met	Gly	Tyr	Trp	
	Ile	Glu	Ser	Ala	965	Leu	Asn	Asp	Thr	970	Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	
	Ile	Glu	Val	Lys	980	Asn	Cys	His	Trp	985	Pro	Lys	Ser	His	Thr	Leu	Trp	Ser	
	Asn	Gly	Val	Leu	995	Glu	Ser	Glu	Met	1000	Ile	Ile	Pro	Lys	Asn	Leu	Ala	Gly	
					1010					1015									

Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035 1040
 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 1050 1055
 Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 1065 1070
 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 1080 1085
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys
 1090 1095 1100
 Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu
 1105 1110 1115 1120
 Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
 1125 1130 1135
 Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
 1140 1145 1150
 Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
 1155 1160 1165
 Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
 1170 1175 1180
 Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
 1185 1190 1195 1200
 Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
 1205 1210 1215
 Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr
 1220 1225 1230
 Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
 1235 1240 1245
 Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
 1250 1255 1260
 Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
 1265 1270 1275 1280
 Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
 1285 1290 1295
 Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Leu Leu Thr Ser Ser
 1300 1305 1310
 Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
 1315 1320 1325
 Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
 1330 1335 1340
 Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
 1345 1350 1355 1360
 Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
 1365 1370 1375
 Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
 1380 1385 1390
 Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
 1395 1400 1405
 Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
 1410 1415 1420
 Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu
 1425 1430 1435 1440
 Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
 1445 1450 1455
 Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
 1460 1465 1470
 Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
 1475 1480 1485
 Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
 1490 1495 1500
 Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
 1505 1510 1515 1520
 Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
 1525 1530 1535
 Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
 1540 1545 1550
 Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu
 1555 1560 1565
 Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val

1570 1575 1580
 Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
 1585 1590 1595 1600
 Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
 1605 1610 1615
 Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
 1620 1625 1630
 Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
 1635 1640 1645
 Asp Asn Pro Glu Ile Glu Asp Ile Phe Arg Lys Arg Arg Leu Thr
 1650 1655 1660
 Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1665 1670 1675 1680
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1685 1690 1695
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1700 1705 1710
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Glu His Thr Gly
 1715 1720 1725
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1730 1735 1740
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775
 Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr
 1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu
 2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125

Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Gly Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205
 Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp
 2210 2215 2220
 Asn Gln Leu Thr Tyr Val Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 2230 2235 2240
 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu
 2245 2250 2255
 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp
 2260 2265 2270
 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr
 2275 2280 2285
 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val
 2290 2295 2300
 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly
 2305 2310 2315 2320
 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu
 2325 2330 2335
 Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala
 2340 2345 2350
 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu
 2355 2360 2365
 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Gly Ile
 2370 2375 2380
 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro
 2385 2390 2395 2400
 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu
 2405 2410 2415
 Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala
 2420 2425 2430
 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
 2435 2440 2445
 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met
 2450 2455 2460
 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
 2465 2470 2475 2480
 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile
 2485 2490 2495
 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly
 2500 2505 2510
 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp
 2515 2520 2525
 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His
 2530 2535 2540
 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn
 2545 2550 2555 2560
 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly
 2565 2570 2575
 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys
 2580 2585 2590
 Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser
 2595 2600 2605
 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe
 2610 2615 2620
 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu
 2625 2630 2635 2640
 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu
 2645 2650 2655
 Asn Leu Val Glu Asn Trp Leu Asn Asn Thr Gln Phe Cys Ile Lys
 2660 2665 2670
 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu

2675 2680 2685
 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn
 2690 2695 2700
 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
 2705 2710 2715 2720
 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met
 2725 2730 2735
 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly
 2740 2745 2750
 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile
 2755 2760 2765
 Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 2770 2775 2780
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 2785 2790 2795 2800
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 2805 2810 2815
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 2820 2825 2830
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 2835 2840 2845
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 2850 2855 2860
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2865 2870 2875 2880
 Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu
 2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe
 3105 3110 3115
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3120 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser
 3140 3145 3150
 Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
 3170 3175 3180
 Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195 3200
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
 3205 3210 3215
 Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
 3220 3225 3230

Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
 3235 3240 3245
 Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
 3250 3255 3260
 Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
 3265 3270 3275 3280
 Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
 3285 3290 3295
 Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
 3300 3305 3310
 Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro
 3315 3320 3325
 Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
 3330 3335 3340
 Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
 3345 3350 3355 3360
 Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
 3365 3370 3375
 Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

<210> SEQ ID NO 15

<211> COMPRIMENTO: 10723

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10272)

<400> SEQUÊNCIA: 15

ägttggtagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta	60
gttctaacag tttttaatt agagagcaga tctctg atg aat aac caa cgg aaa	114
Met Asn Asn Gln Arg Lys	
1	
aag gcg aaa aac acg cct ttc aat atg ctg aaa cgc gag aga aac cgc	162
Lys Ala Lys Asn Thr Pro Phe Asn Met Leu Lys Arg Glu Arg Asn Arg	
10	
gtg tcg act gtg caa cag ctg aca aag aga ttc tca ctt gga atg ctg	210
Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu	
25	
cag gga cga gga cca tta aaa ctg ttc atg gcc ctg gtg gcg ttc ctt	258
Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu	
40	
cgt ttc cta aca atc cca cca aca gca ggg ata ttg aag aga tgg gga	306
Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly Ile Leu Lys Arg Trp Gly	
55	
aca att aaa aaa tca aaa gct att aat gtt ttg aga ggg ttc agg aaa	354
Thr Ile Lys Lys Ser Lys Ala Ile Asn Val Leu Arg Gly Phe Arg Lys	
75	
gag att gga agg atg ctg aac atc ttg aat agg aga cgc aga tct gca	402
Glu Ile Gly Arg Met Leu Asn Ile Leu Asn Arg Arg Arg Arg Ser Ala	
90	
ggc atg atc att atg ctg att cca aca gtg atg gcg ttc cat tta acc	450
Gly Met Ile Ile Met Leu Ile Pro Thr Val Met Ala Phe His Leu Thr	
105	
aca cgt aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg	498
Thr Arg Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly	
120	
aaa agt ctt ctg ttt aaa aca gag gtt ggc gtg aac atg tgt acc ctc	546
Lys Ser Leu Leu Phe Lys Thr Glu Val Gly Val Asn Met Cys Thr Leu	
135	
atg gcc atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag	594
Met Ala Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys	
155	
tgt ccc ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc	642
Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys	
170	
175	
180	

aac	tct	acg	tcc	acg	tgg	gta	act	tat	ggg	acg	tgt	acc	acc	atg	gga	690
Asn	Ser	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Thr	Thr	Met	Gly	
	185						190					195				
gaa	cat	aga	aga	gaa	aaa	aga	tca	gtg	gca	ctc	gtt	cca	cat	gtg	gga	738
Glu	His	Arg	Arg	Glu	Lys	Arg	Ser	Val	Ala	Leu	Val	Pro	His	Val	Gly	
	200					205					210					
atg	gga	ctg	gag	aca	cga	act	gaa	aca	tgg	atg	tca	tca	gaa	ggg	gcc	786
Met	Gly	Leu	Glu	Thr	Arg	Thr	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	
	215				220					225					230	
tgg	aaa	cat	gtc	cag	aga	att	gaa	act	tgg	atc	ttg	aga	cat	cca	ggc	834
Trp	Lys	His	Val	Gln	Arg	Ile	Glu	Thr	Trp	Ile	Leu	Arg	His	Pro	Gly	
				235					240					245		
ttc	acc	atg	atg	gca	gca	atc	ctg	gca	tac	acc	ata	gga	acg	aca	cat	882
Phe	Thr	Met	Met	Ala	Ala	Ile	Leu	Ala	Tyr	Thr	Ile	Gly	Thr	Thr	His	
				250				255					260			
ttc	caa	aga	gcc	ctg	att	ttc	atc	tta	ctg	aca	gct	gtc	act	cct	tca	930
Phe	Gln	Arg	Ala	Leu	Ile	Phe	Ile	Leu	Leu	Thr	Ala	Val	Thr	Pro	Ser	
		265					270					275				
atg	aca	atg	cgt	tgc	ata	gga	atg	tca	aat	aga	gac	ttt	gtg	gaa	ggg	978
Met	Thr	Met	Arg	Cys	Ile	Gly	Met	Ser	Asn	Arg	Asp	Phe	Val	Glu	Gly	
	280					285					290					
gtt	tca	gga	gga	agc	tgg	gac	ata	gtc	tta	gaa	cat	gga	agc	tgt		1026
Val	Ser	Gly	Gly	Ser	Trp	Val	Asp	Ile	Val	Leu	Glu	His	Gly	Ser	Cys	
	295				300					305					310	
gtg	acg	acg	atg	gca	aaa	aac	aaa	cca	aca	ttg	gat	ttt	gaa	ctg	ata	1074
Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Phe	Glu	Leu	Ile	
				315				320						325		
aaa	aca	gaa	gcc	aaa	cag	cct	gcc	acc	cta	agg	aag	tac	tgt	ata	gag	1122
Lys	Thr	Glu	Ala	Lys	Gln	Pro	Ala	Thr	Leu	Arg	Lys	Tyr	Cys	Ile	Glu	
			330					335					340			
gca	aag	cta	acc	aac	aca	aca	aca	gaa	tct	cgc	tgc	cca	aca	caa	ggg	1170
Ala	Lys	Leu	Thr	Asn	Thr	Thr	Thr	Glu	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
		345					350					355				
gaa	ccc	agc	cta	aat	gaa	gag	cag	gac	aaa	agg	ttc	gtc	tgc	aaa	cac	1218
Glu	Pro	Ser	Leu	Asn	Glu	Glu	Gln	Asp	Lys	Arg	Phe	Val	Cys	Lys	His	
	360					365					370					
tcc	atg	gta	gac	aga	gga	tgg	gga	aat	gga	tgt	gga	cta	ttt	gga	aag	1266
Ser	Met	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
	375				380					385					390	
gga	ggc	att	gtg	acc	tgt	gct	atg	ttc	aga	tgc	aaa	aag	aac	atg	gaa	1314
Gly	Gly	Ile	Val	Thr	Cys	Ala	Met	Phe	Arg	Cys	Lys	Lys	Asn	Met	Glu	
				395				400						405		
gga	aaa	gtt	gtg	caa	gaa	aac	ttg	gaa	tac	acc	att	gtg	ata	aca		1362
Gly	Lys	Val	Val	Gln	Pro	Glu	Asn	Leu	Glu	Tyr	Thr	Ile	Val	Ile	Thr	
			410				415						420			
cct	cac	tca	ggg	gaa	gag	cat	gca	gtc	gga	aat	gac	aca	gga	aaa	cat	1410
Pro	His	Ser	Gly	Glu	Glu	His	Ala	Val	Gly	Asn	Asp	Thr	Gly	Lys	His	
		425					430					435				
ggc	aag	gaa	atc	aaa	ata	aca	cca	cag	agt	tcc	atc	aca	gaa	gca	gaa	1458
Gly	Lys	Glu	Ile	Lys	Ile	Thr	Pro	Gln	Ser	Ser	Ile	Thr	Glu	Ala	Glu	
	440				445						450					
ttg	aca	ggt	tat	ggc	act	gtc	aca	atg	gag	tgc	tct	cca	aga	acg	ggc	1506
Leu	Thr	Gly	Tyr	Gly	Thr	Val	Thr	Met	Glu	Cys	Ser	Pro	Arg	Thr	Gly	
	455				460					465					470	
ctc	gac	ttc	aat	gag	atg	gtg	ttg	ctg	cag	atg	gaa	aat	aaa	gct	tgg	1554
Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Gln	Met	Glu	Asn	Lys	Ala	Trp	
				475					480					485		
ctg	gtg	cac	agg	caa	tgg	ttc	cta	gac	ctg	ccg	tta	cca	tgg	ttg	ccc	1602
Leu	Val	His	Arg	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Trp	Leu	Pro	
			490					495					500			
gga	gcg	gac	aca	caa	ggg	tca	aat	tgg	ata	cag	aaa	gag	aca	ttg	gtc	1650
Gly	Ala	Asp	Thr	Gln	Gly	Ser	Asn	Trp	Ile	Gln	Lys	Glu	Thr	Leu	Val	
		505					510					515				
act	ttc	aaa	aat	ccc	cat	gcg	aaa	cag	gat	gtt	gtt	gtt	tta	gga		1698
Thr	Phe	Lys	Asn	Pro	His	Ala	Lys	Lys	Gln	Asp	Val	Val	Val	Leu	Gly	
		520			525						530					
tcc	caa	gaa	ggg	gcc	atg	cac	aca	gca	ctt	aca	ggg	gcc	aca	gaa	atc	1746
Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile	
					540					545					550	

caa	atg	tca	tca	gga	aac	tta	ctc	ttc	aca	gga	cat	ctc	aag	tgc	agg	1794
Gln	Met	Ser	Ser	Gly	Asn	Leu	Leu	Phe	Thr	Gly	His	Leu	Lys	Cys	Arg	
				555					560					565		
ctg	aga	atg	gac	aag	cta	cag	ctc	aaa	gga	atg	tca	tac	tct	atg	tgc	1842
Leu	Arg	Met	Asp	Lys	Leu	Gln	Leu	Lys	Gly	Met	Ser	Tyr	Ser	Met	Cys	
			570					575					580			
aca	gga	aag	ttt	aaa	gtt	gtg	aag	ata	gca	gaa	aca	cat	gga			1890
Thr	Gly	Lys	Phe	Lys	Val	Val	Lys	Glu	Ile	Ala	Glu	Thr	Gln	His	Gly	
		585					590					595				
aca	ata	gtt	atc	aga	gtg	caa	tat	gaa	ggg	gac	ggc	tct	cca	tgc	aag	1938
Thr	Ile	Val	Ile	Arg	Val	Gln	Tyr	Glu	Gly	Asp	Gly	Ser	Pro	Cys	Lys	
	600				605					610						
atc	cct	ttt	gag	ata	atg	gat	ttg	gaa	aaa	aga	cat	gtc	tta	ggt	cgc	1986
Ile	Pro	Phe	Glu	Ile	Met	Asp	Leu	Glu	Lys	Arg	His	Val	Leu	Gly	Arg	
	615				620					625					630	
ctg	att	aca	gtc	aac	cca	att	gtg	aca	gaa	gat	agc	cca	gtc	aac		2034
Leu	Ile	Thr	Val	Asn	Pro	Ile	Val	Thr	Glu	Lys	Asp	Ser	Pro	Val	Asn	
				635					640					645		
ata	gaa	gca	gaa	cct	cca	ttt	gga	gac	agc	tac	atc	atc	ata	gga	gta	2082
Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Asp	Ser	Tyr	Ile	Ile	Ile	Gly	Val	
			650					655					660			
gag	ccg	gga	caa	ctg	aag	ctc	aac	tgg	ttt	aag	aaa	gga	agt	tct	atc	2130
Glu	Pro	Gly	Gln	Leu	Lys	Leu	Asn	Trp	Phe	Lys	Lys	Gly	Ser	Ser	Ile	
		665					670					675				
ggc	caa	atg	ttt	gag	aca	aca	atg	agg	ggg	ggc	aag	aga	atg	gcc	att	2178
Gly	Gln	Met	Phe	Glu	Thr	Met	Arg	Gly	Gly	Ala	Lys	Arg	Met	Ala	Ile	
		680				685					690					
tta	ggt	gac	aca	gcc	tgg	gat	ttt	gga	tcc	ttg	gga	gga	gtg	ttt	aca	2226
Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Leu	Gly	Gly	Val	Phe	Thr	
	695				700					705					710	
tct	ata	gga	aag	gct	ctc	cac	caa	gtc	ttt	gga	gca	atc	tat	gga	gct	2274
Ser	Ile	Gly	Lys	Ala	Leu	His	Gln	Val	Phe	Gly	Ala	Ile	Tyr	Gly	Ala	
				715					720					725		
gcc	ttc	agt	ggg	gtt	tca	tgg	act	atg	aaa	atc	ctc	ata	gga	gtc	att	2322
Ala	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	Ile	Leu	Ile	Gly	Val	Ile	
			730					735					740			
atc	aca	tgg	ata	gga	atg	aat	tca	cgc	agc	acc	tca	ctg	tct	gtg	aca	2370
Ile	Thr	Trp	Ile	Gly	Met	Asn	Ser	Arg	Ser	Thr	Ser	Leu	Ser	Val	Thr	
		745					750					755				
cta	gta	ttg	gtg	gga	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	2418
Leu	Val	Leu	Val	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln	
		760				765					770					
gcc	gat	agt	ggt	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	2466
Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys	
				780						785					790	
ggc	agt	ggg	att	ttc	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	2514
Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln	
				795					800					805		
tac	aag	ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	2562
Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ala	Ile	Gln	
			810					815					820			
aaa	gcc	cat	gaa	gag	gac	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	2610
Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu	
		825					830					835				
gag	aat	ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	2658
Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu	
		840				845					850					
tca	gaa	aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	2706
Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	
		855			860					865					870	
atc	atg	cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	2754
Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu	
				875					880					885		
aag	tat	tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	2802
Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu	
			890					895					900			
tct	cat	aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	2850
Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys	
		905					910					915				

ccc Pro	aac Asn	aca Thr	aat Asn	aga Arg	gct Ala	tgg Trp	aat Asn	tcg Ser	ttg Leu	gaa Glu	gtt Val	gaa Glu	gac Asp	tat Tyr	ggc Gly	2898
920						925					930					
ttt Phe	gga Gly	gta Val	ttc Phe	acc Thr	acc Thr	aat Asn	ata Ile	tgg Trp	cta Leu	aaa Lys	ttg Leu	aaa Lys	gaa Glu	aaa Lys	cag Gln	2946
935					940					945					950	
gat Asp	gta Val	ttc Phe	tgc Cys	gac Asp	tca Ser	aaa Lys	ctc Leu	atg Met	tca Ser	gcg Ala	gcc Ala	ata Ile	aaa Lys	gac Asp	aac Asn	2994
				955					960					965		
aga Arg	gcc Ala	gtc Val	cat His	gcc Ala	gat Asp	atg Met	ggg Gly	tat Tyr	tgg Trp	ata Ile	gaa Glu	agt Ser	gca Ala	ctc Leu	aat Asn	3042
			970					975					980			
gac Asp	aca Thr	tgg Trp	aag Lys	ata Ile	gag Glu	aaa Lys	gcc Ala	tct Ser	ttc Phe	att Ile	gaa Glu	gtt Val	aaa Lys	aac Asn	tgc Cys	3090
		985					990					995				
cac His	tgg Pro	cca Lys	aaa Ser	tca His	cac His	acc Thr	ctc Leu	tgg Trp	agc Ser	aat Asn	gga Gly	gtg Val	cta Leu	gaa Glu	agt Ser	3138
	1000					1005					1010					
gag Glu	atg Met	ata Ile	att Ile	cca Pro	aag Lys	aat Asn	ctc Leu	gct Ala	gga Gly	cca Pro	gtg Val	tct Ser	caa Gln	cac His	aac Asn	3186
	1015				1020					1025					1030	
tat Tyr	aga Arg	cca Pro	ggc Gly	tac Tyr	cat His	aca Thr	caa Gln	ata Ile	aca Thr	gga Gly	cca Pro	tgg Trp	cat His	cta Leu	ggg Gly	3234
				1035				1040						1045		
aag Lys	ctt Leu	gag Glu	atg Met	gac Asp	ttt Phe	gat Asp	ttc Phe	tgt Cys	gat Asp	gga Gly	aca Thr	aca Thr	gtg Val	gta Val	gtg Val	3282
			1050					1055					1060			
act Thr	gag Glu	gac Asp	tgc Cys	gga Gly	aat Asn	aga Arg	gga Gly	ccc Pro	tct Ser	ttg Leu	aga Arg	aca Thr	acc Thr	act Thr	gcc Ala	3330
		1065					1070					1075				
tct Ser	gga Gly	aaa Lys	ctc Leu	ata Ile	aca Thr	gaa Trp	tgg Cys	tgc Cys	tgc Cys	cga Arg	tct Ser	tgc Cys	aca Thr	tta Leu	cca Pro	3378
	1080					1085					1090					
ccg Pro	cta Leu	aga Arg	tac Tyr	aga Arg	ggt Glu	gat Asp	ggg Gly	tgc Cys	tgg Trp	tac Tyr	ggg Gly	atg Met	gaa Glu	atc Ile		3426
	1095				1100					1105				1110		
aga Arg	cca Pro	ttg Leu	aag Lys	gag Glu	aaa Lys	gaa Glu	gag Glu	aat Asn	ttg Leu	gtc Val	aac Asn	tcc Ser	ttg Leu	gtc Val	aca Thr	3474
				1115					1120				1125			
gct Ala	gga Gly	cat His	ggg Gly	gtc Gln	gac Val	aac Asp	ttt Phe	tca Ser	cta Leu	gga Gly	gtc Val	ttg Leu	gga Gly	atg Met		3522
		1130					1135					1140				
gca Ala	ttg Leu	ttc Phe	ctg Leu	gag Glu	gaa Glu	atg Met	ctt Leu	agg Arg	acc Thr	cga Arg	gta Val	gga Gly	acg Thr	aaa Lys	cat His	3570
		1145				1150					1155					
gca Ala	ata Ile	cta Leu	cta Leu	ggt Val	gca Ala	ggt Val	tct Ser	ttt Phe	gtg Val	aca Thr	ttg Leu	atc Ile	aca Thr	ggg Gly	aac Asn	3618
	1160				1165					1170						
atg Met	tcc Ser	ttt Phe	aga Arg	gac Asp	ctg Leu	gga Gly	aga Arg	gtg Val	atg Met	gtt Val	atg Met	gta Val	ggc Gly	gcc Ala	act Thr	3666
	1175				1180				1185					1190		
atg Met	acg Thr	gat Asp	gac Asp	ata Ile	ggt Gly	atg Met	ggc Gly	gtg Val	act Thr	tat Tyr	ctt Leu	gcc Ala	cta Leu	cta Leu	gca Ala	3714
				1195					1200				1205			
gcc Ala	ttc Phe	aaa Lys	gtc Val	aga Arg	cca Pro	act Thr	ttt Phe	gca Ala	gct Ala	gga Gly	cta Leu	ctc Leu	ttg Leu	aga Arg	aag Lys	3762
			1210					1215				1220				
ctg Leu	acc Thr	tcc Ser	aag Lys	gaa Glu	ttg Leu	atg Met	atg Met	act Thr	act Thr	ata Ile	gga Gly	att Ile	gta Val	ctc Leu	ctc Leu	3810
		1225				1230						1235				
tcc Ser	cag Gln	agc Ser	acc Thr	ata Ile	cca Pro	gag Glu	acc Thr	att Ile	ctt Leu	gag Glu	ttg Leu	act Thr	gat Asp	gcg Ala	tta Leu	3858
	1240					1245					1250					
gcc Ala	tta Leu	ggc Gly	atg Met	atg Met	gtc Val	ctc Leu	aaa Lys	atg Met	gtg Val	aga Arg	aat Asn	atg Met	gaa Glu	aag Lys	tat Tyr	3906
	1255				1260					1265					1270	
caa Gln	ttg Leu	gca Ala	gtg Val	act Thr	atc Ile	atg Met	gct Ala	atc Ile	ttg Leu	tgc Cys	gtc Val	cca Pro	aac Asn	gca Ala	gtg Val	3954
				1275					1280					1285		

ata tta caa aac gca tgg aaa gtg agt tgc aca ata ttg gca gtg gtg	4002
Ile Leu Gln Asn Ala Trp Lys Val Ser Cys Thr Ile Leu Ala Val Val	
1290 1295 1300	
tcc gtt tcc cca ctg ttc tta aca tcc tca cag caa aaa aca gat tgg	4050
Ser Val Ser Pro Leu Phe Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp	
1305 1310 1315	
ata cca tta gca ttg acg atc aaa ggt ctc aat cca aca gct att ttt	4098
Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe	
1320 1325 1330	
cta aca acc ctc tca aga acc agc aag aaa agg agc tgg cca tta aat	4146
Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn	
1335 1340 1345 1350	
gag gct atc atg gca gtc ggg atg gtg agc att tta gcc agt tct ctc	4194
Glu Ala Ile Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu	
1355 1360 1365	
cta aaa aat gat att ccc atg aca gga cca tta gtg gct gga ggg ctc	4242
Leu Lys Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu	
1370 1375 1380	
ctc act gtg tgc tac gtg ctc act gga cga tgc gcc gat ttg gaa ctg	4290
Leu Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu	
1385 1390 1395	
gag aga gca gcc gat gtc aaa tgg gaa gac cag gca gag ata tca gga	4338
Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly	
1400 1405 1410	
agc agt cca atc ctg tca ata aca ata tca gaa gat ggt agc atg tgc	4386
Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser	
1415 1420 1425 1430	
ata aaa aat gaa gag gaa gaa caa aca ctg acc ata ctc att aga aca	4434
Ile Lys Asn Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg Thr	
1435 1440 1445	
gga ttg ctg gtg atc tca gga ctt ttt cct gta tca ata cca atc acg	4482
Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro Ile Thr	
1450 1455 1460	
gca gca gca tgg tac ctg tgg gaa gtg aag aaa caa cgg gcc gga gta	4530
Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg Ala Gly Val	
1465 1470 1475	
ttg tgg gac gtt cct tca ccc cca atg gga aag gct gaa ctg gaa	4578
Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys Ala Glu Leu Glu	
1480 1485 1490	
gat gga gcc tat aga att aag caa aaa ggg att ctt gga tat tcc cag	4626
Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile Leu Gly Tyr Ser Gln	
1495 1500 1505 1510	
atc gga gcc gga gtt tac aaa gaa gga aca ttc cat aca atg tgg cat	4674
Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr Phe His Thr Met Trp His	
1515 1520 1525	
gtc aca cgt ggc gct gtt cta atg cat aaa gga aag agg att gaa cca	4722
Val Thr Arg Gly Ala Val Leu Met His Lys Gly Lys Arg Ile Glu Pro	
1530 1535 1540	
tca tgg gcg gac gtc aag aaa gac cta ata tca tat gga gga ggc tgg	4770
Ser Trp Ala Asp Val Lys Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp	
1545 1550 1555	
aag tta gaa gga gaa tgg aag gaa gga gaa gtc cag gta ttg gca	4818
Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala	
1560 1565 1570	
ctg gag cct gga aaa aat cca aga gcc gtc caa acg aaa cct ggt ctt	4866
Leu Glu Pro Gly Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu	
1575 1580 1585 1590	
ttc aaa acc aac gcc gga aca ata ggt gct gta tct ctg gac ttt tct	4914
Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser	
1595 1600 1605	
cct gga acg tca gga tct cca att atc gac aaa aaa gga aaa gtt gtg	4962
Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val	
1610 1615 1620	
ggt ctt tat ggt aat ggt gtt gtt aca agg agt gga gca tat gtg agt	5010
Gly Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser	
1625 1630 1635	
gct ata gcc cag act gaa aaa agc att gaa gac aac cca gag atc gaa	5058
Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile Glu	
1640 1645 1650	

gat gac att ttc cga aag aga aga ctg acc atc atg gac ctc cac cca	5106
Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro	
1655 1660 1665 1670	
gga gcg gga aag acg aag aga tac ctt ccg gcc ata gtc aga gaa gct	5154
Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu Ala	
1675 1680 1685	
ata aaa cgg ggt ttg aga aca tta atc ttg gcc ccc act aga gtt gtg	5202
Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg Val Val	
1690 1695 1700	
gca gct gaa atg gag gaa gcc ctt aga gga ctt cca ata aga tac cag	5250
Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile Arg Tyr Gln	
1705 1710 1715	
acc cca gcc atc aga gct gtg cac acc ggg cgg gag att gtg gac cta	5298
Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu Ile Val Asp Leu	
1720 1725 1730	
atg tgt cat gcc aca ttt acc atg agg ctg cta tca cca gtt aga gtg	5346
Met Cys His Ala Thr Phe Thr Met Arg Leu Leu Ser Pro Val Arg Val	
1735 1740 1745 1750	
cca aac tac aac ctg att atc atg gac gaa gcc cat ttc aca gac cca	5394
Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu Ala His Phe Thr Asp Pro	
1755 1760 1765	
gca agt ata gca gct aga gga tac atc tca act cga gtg gag atg ggt	5442
Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg Val Glu Met Gly	
1770 1775 1780	
gag gca gct ggg att ttt atg aca gcc act ccc ccg gga agc aga gac	5490
Glu Ala Ala Gly Ile Phe Met Thr Ala Thr Pro Pro Gly Ser Arg Asp	
1785 1790 1795	
cca ttt cct cag agc aat gca cca atc ata gat gaa gaa aga gaa atc	5538
Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile	
1800 1805 1810	
cct gaa cgc tcg tgg aat tcc gga cat gaa tgg gtc acg gat ttt aaa	5586
Pro Glu Arg Ser Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys	
1815 1820 1825 1830	
ggg aag act gtt tgg ttc gtt cca agt ata aaa gca gga aat gat ata	5634
Gly Lys Thr Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile	
1835 1840 1845	
gca gct tgc ctg agg aaa aat gga aag aaa gtg ata caa ctc agt agg	5682
Ala Ala Cys Leu Arg Lys Asn Gly Lys Val Ile Gln Leu Ser Arg	
1850 1855 1860	
aag acc ttt gat tct gag tat gtc aag act aga acc aat gat tgg gac	5730
Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp	
1865 1870 1875	
ttc gtg gtt aca act gac att tca gaa atg ggt gcc aat ttc aag gct	5778
Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala	
1880 1885 1890	
gag agg gtt ata gac ccc aga cgc tgc atg aaa cca gtc ata cta aca	5826
Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu Thr	
1895 1900 1905 1910	
gat ggt gaa gag cgg gtg att ctg gca gga cct atg cca gtg acc cac	5874
Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr His	
1915 1920 1925	
tct agt gca gca caa aga aga ggg aga ata gga aga aat cca aaa aat	5922
Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro Lys Asn	
1930 1935 1940	
gag aat gac cag tac ata tac atg ggg gaa cct ctg gaa aat gat gaa	5970
Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu Asn Asp Glu	
1945 1950 1955	
gac tgt gca cac tgg aaa gaa gct aaa atg ctc cta gat aac atc aac	6018
Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu Asp Asn Ile Asn	
1960 1965 1970	
acg cca gaa gga atc att cct agc atg ttc gaa cca gag cgt gaa aag	6066
Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu Pro Glu Arg Glu Lys	
1975 1980 1985 1990	
gtg gat gcc att gat ggc gaa tac cgc ttg aga gga gaa gca agg aaa	6114
Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu Arg Gly Glu Ala Arg Lys	
1995 2000 2005	
acc ttt gta gac tta atg aga aga gga gac cta cca gtc tgg ttg gcc	6162
Thr Phe Val Asp Leu Met Arg Arg Gly Asp Leu Pro Val Trp Leu Ala	
2010 2015 2020	

tac	aga	gtg	gca	gct	gaa	ggc	atc	aac	tac	gca	gac	aga	agg	tgg	tgt	6210
Tyr	Arg	Val	Ala	Ala	Glu	Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	
		2025					2030					2035				
ttt	gat	gga	gtc	aag	aac	aac	caa	atc	cta	gaa	gaa	aac	gtg	gaa	gtt	6258
Phe	Asp	Gly	Val	Lys	Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	
		2040				2045					2050					
gaa	atc	tgg	aca	aaa	gaa	ggg	gaa	agg	aag	aaa	ttg	aaa	ccc	aga	tgg	6306
Glu	Ile	Trp	Thr	Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	
		2055			2060					2065					2070	
ttg	gat	gct	agg	atc	tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	6354
Leu	Asp	Ala	Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	
			2075						2080					2085		
gaa	ttt	gca	gcc	gga	aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	6402
Glu	Phe	Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	
			2090					2095					2100			
atg	ggg	agg	ctc	cca	acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	6450
Met	Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu	
		2105					2110					2115				
gac	aac	tta	gca	gtg	ctg	cac	acg	gct	gag	gca	ggg	gga	agg	gcg	tac	6498
Asp	Asn	Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr	
		2120				2125					2130					
aac	cat	gct	ctc	agt	gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	6546
Asn	His	Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu	
		2135			2140					2145					2150	
ctg	aca	ctt	ctg	gct	aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	6594
Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met	
			2155					2160						2165		
agc	gca	agg	ggc	ata	ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	6642
Ser	Ala	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile	
			2170				2175						2180			
acg	gct	agc	atc	ctc	cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	6690
Thr	Ala	Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile	
		2185					2190					2195				
gca	gct	tca	ata	ata	ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	6738
Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro	
		2200				2205					2210					
gaa	cct	gaa	aaa	cag	aga	aca	ccc	caa	gac	aac	caa	ctg	acc	tac	gtt	6786
Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln	Asp	Asn	Gln	Leu	Thr	Tyr	Val	
		2215			2220					2225					2230	
gtc	ata	gcc	atc	ctc	aca	gtg	gtg	gcc	gca	acc	atg	gca	aac	gag	atg	6834
Val	Ile	Ala	Ile	Leu	Thr	Val	Val	Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	
			2235					2240						2245		
ggg	ttc	cta	gaa	aaa	acg	aag	aaa	gat	ctc	gga	ttg	gga	agc	att	gca	6882
Gly	Phe	Leu	Glu	Lys	Thr	Lys	Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	
		2250					2255						2260			
acc	cag	caa	ccc	gag	agc	aac	atc	ctg	gac	ata	gat	cta	cgt	cct	gca	6930
Thr	Gln	Gln	Pro	Glu	Ser	Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	
		2265					2270					2275				
tca	gca	tgg	acg	ctg	tat	gcc	gtg	gcc	aca	aca	ttt	gtt	aca	cca	atg	6978
Ser	Ala	Trp	Thr	Leu	Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	
		2280			2285						2290					
ttg	aga	cat	agc	att	gaa	aat	tcc	tca	gtg	aat	gtg	tcc	cta	aca	gct	7026
Leu	Arg	His	Ser	Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	
		2295			2300					2305					2310	
ata	gcc	aac	caa	gcc	aca	gtg	tta	atg	ggg	ctc	ggg	aaa	gga	tgg	cca	7074
Ile	Ala	Asn	Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	
			2315					2320						2325		
ttg	tca	aag	atg	gac	atc	gga	gtt	ccc	ctt	ctc	gcc	att	gga	tgc	tac	7122
Leu	Ser	Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	
			2330					2335					2340			
tca	caa	gtc	aac	ccc	ata	act	ctc	aca	gca	gct	ctt	ttc	tta	ttg	gta	7170
Ser	Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	
		2345					2350					2355				
gca	cat	tat	gcc	atc	ata	ggg	cca	gga	ctc	caa	gca	aaa	gca	acc	aga	7218
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	
		2360			2365					2370						
gaa	gct	cag	aaa	aga	gca	gcg	ggc	atc	atg	aaa	aac	cca	act	gtc		7266
Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val	
		2375			2380				2385					2390		

gat gga ata aca gtg att gac cta gat cca ata cct tat gat cca aag	7314
Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro Lys	
2395	
ttt gaa aag cag ttg gga caa gta atg ctc cta gtc ctc tgc gtg act	7362
Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys Val Thr	
2410	
caa gta ttg atg atg agg act aca tgg gct ctg tgt gag gct tta acc	7410
Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu Ala Leu Thr	
2425	
tta gct acc ggg ccc atc tcc aca ttg tgg gaa gga aat cca ggg agg	7458
Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly Asn Pro Gly Arg	
2440	
ttt tgg aac act acc att gcg gtg tca atg gct aac att ttt aga ggg	7506
Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala Asn Ile Phe Arg Gly	
2455	
agt tac ttg gcc gga gct gga ctt ctc ttt tct att atg aag aac aca	7554
Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe Ser Ile Met Lys Asn Thr	
2475	
acc aac aca aga agg gga act ggc aac ata gga gag acg ctt gga gag	7602
Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile Gly Glu Thr Leu Gly Glu	
2490	
aaa tgg aaa agc cga ttg aac gca ttg gga aaa agt gaa ttc cag atc	7650
Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile	
2505	
tac aag aaa agt gga atc cag gaa gtg gat aga acc tta gca aaa gaa	7698
Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu	
2520	
ggc att aaa aga gga gaa acg gac cat cac gct gtg tgc cga ggc tca	7746
Gly Ile Lys Arg Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser	
2535	
gca aaa ctg aga tgg ttc gtt gag aga aac atg gtc aca cca gaa ggg	7794
Ala Lys Leu Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly	
2555	
aaa gta gtg gac ctc ggt tgt ggc aga gga ggc tgg tca tac tat tgt	7842
Lys Val Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys	
2570	
gga gga cta aag aat gta aga gaa gtc aaa ggc cta aca aaa gga gga	7890
Gly Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly	
2585	
cca gga cac gaa gaa ccc atc ccc atg tca aca tat ggg tgg aat cta	7938
Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu	
2600	
gtg cgt ctt caa agt gga gtt gac gtt ttc ttc atc ccg cca gaa aag	7986
Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys	
2615	
tgt gac aca tta ttg tgt gac ata ggg gag tca cca aat ccc aca	8034
Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro Thr	
2635	
gtg gaa gca gga cga aca ctc aga gtc ctt aac tta gta gaa aat tgg	8082
Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu Asn Trp	
2650	
ttg aac aac aac act caa ttt tgc ata aag gtt ctc aac cca tat atg	8130
Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn Pro Tyr Met	
2665	
ccc tca gtc ata gaa aaa atg gaa gca cta caa agg aaa tat gga gga	8178
Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg Lys Tyr Gly Gly	
2680	
gcc tta gtg agg aat cca ctc tca cga aac tcc aca cat gag atg tac	8226
Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser Thr His Glu Met Tyr	
2695	
tgg gta tcc aat gct tcc ggg aac ata gtg tca tca gtg aac atg att	8274
Trp Val Ser Asn Ala Ser Gly Asn Ile Val Ser Ser Val Asn Met Ile	
2715	
tca agg atg ttg atc aac aga ttt aca atg aga tac aag aaa gcc act	8322
Ser Arg Met Leu Ile Asn Arg Phe Thr Met Arg Tyr Lys Lys Ala Thr	
2730	
tac gag ccg gat gtt gac ctc gga agc gga acc cgt aac atc ggg att	8370
Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile	
2745	
2750	
2755	

gaa agt gag ata cca aac cta gat ata att ggg aaa aga ata gaa aaa	8418
Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys	
2760 2765 2770	
ata aag caa gag cat gaa aca tca tgg cac tat gac caa gac cac cca	8466
Ile Lys Gln Glu His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro	
2775 2780 2785 2790	
tac aaa acg tgg gca tac cat ggt agc tat gaa aca aaa cag act gga	8514
Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly	
2795 2800 2805	
tca gca tca tcc atg gtc aac gga gtg gtc agg ctg ctg aca aaa cct	8562
Ser Ala Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro	
2810 2815 2820	
tgg gac gtt gtc ccc atg gtg aca cag atg gca atg aca gac acg act	8610
Trp Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr	
2825 2830 2835	
cca ttt gga caa cag cgc gtt ttt aaa gag aaa gtg gac acg aga acc	8658
Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg Thr	
2840 2845 2850	
caa gaa ccg aaa gaa ggc acg aag aaa cta atg aaa ata aca gca gag	8706
Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu	
2855 2860 2865 2870	
tgg ctt tgg aaa gaa tta ggg aag aaa aag aca ccc agg atg tgc acc	8754
Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys Thr	
2875 2880 2885	
aga gaa gaa ttc aca aga aag gtg aga agc aat gca gcc ttg ggg gcc	8802
Arg Glu Glu Thr Arg Lys Val Arg Ser Asn Ala Ala Leu Gly Ala	
2890 2895 2900	
ata ttc act gat gag aac aag tgg aag tgg gca cgt gag gct gtt gaa	8850
Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu Ala Val Glu	
2905 2910 2915	
gat agt agg ttt tgg gag ctg gtt gac aag gaa agg aat ctc cat ctt	8898
Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg Asn Leu His Leu	
2920 2925 2930	
gaa gga aag tgt gaa aca tgt gtg tac aac atg atg gga aaa aga gag	8946
Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met Met Gly Lys Arg Glu	
2935 2940 2945 2950	
aag aag cta ggg gaa ttc ggc aag gca aaa ggc agc aga gcc ata tgg	8994
Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys Gly Ser Arg Ala Ile Trp	
2955 2960 2965	
tac atg tgg ctt gga gca cgc ttc tta gag ttt gaa gcc cta gga ttc	9042
Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu Phe Glu Ala Leu Gly Phe	
2970 2975 2980	
tta aat gaa gat cac tgg ttc tcc aga gag aac tcc ctg agt gga gtg	9090
Leu Asn Glu Asp His Trp Phe Ser Arg Glu Asn Ser Leu Ser Gly Val	
2985 2990 2995	
gaa gga gaa ggg ctg cac aag cta ggt tac att cta aga gac gtg agc	9138
Glu Gly Glu Gly Leu His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser	
3000 3005 3010	
aag aaa gag gga gga gca atg tat gcc gat gac acc gca gga tgg gat	9186
Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp	
3015 3020 3025 3030	
aca aga atc aca cta gaa gac cta aaa aat gaa gaa atg gta aca aac	9234
Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn	
3035 3040 3045	
cac atg gaa gga gaa cac aag aaa cta gcc gag gcc att ttc aaa cta	9282
His Met Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu	
3050 3055 3060	
acg tac caa aac aag gtg gtg cgt gtg caa aga cca aca cca aga ggc	9330
Thr Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly	
3065 3070 3075	
aca gta atg gac atc ata tcg aga aga gac caa aga ggt agt gga caa	9378
Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln	
3080 3085 3090	
gtt ggc acc tat gga ctc aat act ttc acc aat atg gaa gcc caa cta	9426
Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu	
3095 3100 3105 3110	
atc aga cag atg gag gga gaa gga gtc ttt aaa agc att cag cac cta	9474
Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His Leu	
3115 3120 3125	

```

aca atc aca gaa gaa atc gct gtg caa aac tgg tta gca aga gtg ggg 9522
Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg Val Gly
3130 3135 3140
cgc gaa agg tta tca aga atg gcc atc agt gga gat gat tgt gtt gtg 9570
Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp Cys Val Val
3145 3150 3155
aaa cct tta gat gac agg ttc gca agc gct tta aca gct cta aat gac 9618
Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr Ala Leu Asn Asp
3160 3165 3170
atg gga aag att agg aaa gac ata caa caa tgg gaa cct tca aga gga 9666
Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg Gly
3175 3180 3185 3190
tgg aat gat tgg aca caa gtg ccc ttc tgt tca cac cat ttc cat gag 9714
Trp Asn Asp Trp Thr Gln Val Pro Phe Cys Ser His His Phe His Glu
3195 3200 3205
tta atc atg aaa gac ggt cgc gta ctc gtt gtt cca tgt aga aac caa 9762
Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln
3210 3215 3220
gat gaa ctg att ggc aga gcc cga atc tcc caa gga gca ggg tgg tct 9810
Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser
3225 3230 3235
ttg cgg gag acg gcc tgt ttg ggg aag tct tac gcc caa atg tgg agc 9858
Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser
3240 3245 3250
ttg atg tac ttc cac aga cgc gac ctc agg ctg gcg gca aat gct att 9906
Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile
3255 3260 3265 3270
tgc tcg gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg 9954
Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp
3275 3280 3285
tcc ata cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca 10002
Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr
3290 3295 3300
gtc tgg aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa 10050
Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315
act cca gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa 10098
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu
3320 3325 3330
gac caa tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg 10146
Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
3335 3340 3345 3350
gca aag aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc 10194
Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly
3355 3360 3365
aat gaa gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga 10242
Asn Glu Glu Thr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg
3370 3375 3380
gaa gag gaa gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa 10292
Glu Glu Glu Glu Ala Gly Val Leu Trp *
3385 3390
caaggctaga agtcaggctg gattaagcca tagtacggaa aaaactatgc tacctgtgag 10352
ccccgtccaa ggacgttaaa agaagtcagg ccatcataaa tgccatagct tgagtaaaact 10412
atgcagcctg tagctccacc tgagaagggtg taaaaaatcc gggaggccac aaaccatgga 10472
agctgtacgc atggcgtagt ggactagcgg tttagaggaga cccctccctt acaaatcgca 10532
gcaacaatgg gggcccaagg cgagatgaag ctgtagtctc gctggaagga ctagaggtta 10592
gaggagaccc ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg 10652
ctgtctcttc agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat 10712
caacaggttc t 10723

```

<210> SEQ ID NO 16

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

EP2278012B1

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 16

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu

1				5				10				15				
Lys	Arg	Glu	Arg	Asn	Arg	Val	Ser	Thr	Val	Gln	Gln	Leu	Thr	Lys	Arg	
			20					25					30			
Phe	Ser	Leu	Gly	Met	Leu	Gln	Gly	Arg	Gly	Pro	Leu	Lys	Leu	Phe	Met	
		35					40					45				
Ala	Leu	Val	Ala	Phe	Leu	Arg	Phe	Leu	Thr	Ile	Pro	Pro	Thr	Ala	Gly	
		50				55					60					
Ile	Leu	Lys	Arg	Trp	Gly	Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val	
65					70					75				80		
Leu	Arg	Gly	Phe	Arg	Lys	Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn	
				85					90					95		
Arg	Arg	Arg	Arg	Ser	Ala	Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val	
			100					105					110			
Met	Ala	Phe	His	Leu	Thr	Thr	Arg	Asn	Gly	Glu	Pro	His	Met	Ile	Val	
		115					120					125				
Ser	Arg	Gln	Glu	Lys	Gly	Lys	Ser	Leu	Leu	Phe	Lys	Thr	Glu	Val	Gly	
		130				135					140					
Val	Asn	Met	Cys	Thr	Leu	Met	Ala	Met	Asp	Leu	Gly	Glu	Leu	Cys	Glu	
145					150					155				160		
Asp	Thr	Ile	Thr	Tyr	Lys	Cys	Pro	Leu	Leu	Arg	Gln	Asn	Glu	Pro	Glu	
				165						170				175		
Asp	Ile	Asp	Cys	Trp	Cys	Asn	Ser	Thr	Ser	Thr	Trp	Val	Thr	Tyr	Gly	
		180						185					190			
Thr	Cys	Thr	Thr	Met	Gly	Glu	His	Arg	Arg	Glu	Lys	Arg	Ser	Val	Ala	
		195					200					205				
Leu	Val	Pro	His	Val	Gly	Met	Gly	Leu	Glu	Thr	Arg	Thr	Glu	Thr	Trp	
		210			215						220					
Met	Ser	Ser	Glu	Gly	Ala	Trp	Lys	His	Val	Gln	Arg	Ile	Glu	Thr	Trp	
225					230					235				240		
Ile	Leu	Arg	His	Pro	Gly	Phe	Thr	Met	Met	Ala	Ala	Ile	Leu	Ala	Tyr	
				245					250					255		
Thr	Ile	Gly	Thr	Thr	His	Phe	Gln	Arg	Ala	Leu	Ile	Phe	Ile	Leu	Leu	
			260					265					270			
Thr	Ala	Val	Thr	Pro	Ser	Met	Thr	Met	Arg	Cys	Ile	Gly	Met	Ser	Asn	
		275					280					285				
Arg	Asp	Phe	Val	Glu	Gly	Val	Ser	Gly	Gly	Ser	Trp	Val	Asp	Ile	Val	
		290				295					300					
Leu	Glu	His	Gly	Ser	Cys	Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	
305					310					315				320		
Leu	Asp	Phe	Glu	Leu	Ile	Lys	Thr	Glu	Ala	Lys	Gln	Pro	Ala	Thr	Leu	
				325					330					335		
Arg	Lys	Tyr	Cys	Ile	Glu	Ala	Lys	Leu	Thr	Asn	Thr	Thr	Thr	Glu	Ser	
			340					345					350			
Arg	Cys	Pro	Thr	Gln	Gly	Glu	Pro	Ser	Leu	Asn	Glu	Glu	Gln	Asp	Lys	
		355					360					365				
Arg	Phe	Val	Cys	Lys	His	Ser	Met	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	
		370				375					380					
Cys	Gly	Leu	Phe	Gly	Lys	Gly	Gly	Ile	Val	Thr	Cys	Ala	Met	Phe	Arg	
385					390					395				400		
Cys	Lys	Lys	Asn	Met	Glu	Gly	Lys	Val	Val	Gln	Pro	Glu	Asn	Leu	Glu	
				405					410					415		
Tyr	Thr	Ile	Val	Ile	Thr	Pro	His	Ser	Gly	Glu	Glu	His	Ala	Val	Gly	
			420					425					430			
Asn	Asp	Thr	Gly	Lys	His	Gly	Lys	Glu	Ile	Lys	Ile	Thr	Pro	Gln	Ser	
		435					440						445			
Ser	Ile	Thr	Glu	Ala	Glu	Leu	Thr	Gly	Tyr	Gly	Thr	Val	Thr	Met	Glu	
		450				455					460					
Cys	Ser	Pro	Arg	Thr	Gly	Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Gln	
465					470					475				480		
Met	Glu	Asn	Lys	Ala	Trp	Leu	Val	His	Arg	Gln	Trp	Phe	Leu	Asp	Leu	
				485					490					495		
Pro	Leu	Pro	Trp	Leu	Pro	Gly	Ala	Asp	Thr	Gln	Gly	Ser	Asn	Trp	Ile	
			500					505					510			
Gln	Lys	Glu	Thr	Leu	Val	Thr	Phe	Lys	Asn	Pro	His	Ala	Lys	Lys	Gln	
		515					520					525				
Asp	Val	Val	Val	Leu	Gly	Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	
		530				535					540					
Thr	Gly	Ala	Thr	Glu	Ile	Gln	Met	Ser	Ser	Gly	Asn	Leu	Leu	Phe	Thr	
545					550					555					560	

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
 565 570 575
 Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
 580 585 590
 Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
 595 600 605
 Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
 610 615 620
 Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
 625 630 635 640
 Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Phe Gly Asp Ser
 645 650 655
 Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
 660 665 670
 Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
 675 680 685
 Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700
 Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
 705 710 715 720
 Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
 725 730 735
 Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
 740 745 750
 Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
 755 760 765
 Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
 770 775 780
 Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 785 790 795 800
 His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 805 810 815
 Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
 820 825 830
 Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
 835 840 845
 Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860
 Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875 880
 Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895
 Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910
 Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925
 Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940
 Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960
 Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975
 Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990
 Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005
 Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly
 1010 1015 1020
 Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035 1040
 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 1050 1055
 Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 1065 1070
 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 1080 1085
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly Glu Asp Gly Cys
 1090 1095 1100
 Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu

```

1105      1110      1115      1120
Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr
Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser
Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Gln Thr Leu
Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Met
Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
Ser Tyr Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu
Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val
Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr
1650      1655      1660

```

Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1665 1670 1675 1680
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1685 1690 1695
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1700 1705 1710
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
 1715 1720 1725
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1730 1735 1740
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775
 Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr
 1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu
 2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125
 Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205
 Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp

2210 Asn Gln Leu Thr Tyr Val Val 2215 Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu 2240
 2245 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp 2255
 2260 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr 2270
 2275 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val 2285
 2290 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly 2300
 2305 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu 2320
 2325 Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala 2335
 2340 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu 2350
 2355 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile 2365
 2370 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro 2380
 2385 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu 2400
 2405 Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala 2415
 2420 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp 2430
 2435 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met 2445
 2450 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe 2460
 2465 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile 2480
 2485 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly 2495
 2500 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp 2510
 2515 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His 2525
 2530 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn 2540
 2545 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly 2555
 2565 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys 2575
 2580 Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser 2590
 2595 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe 2605
 2610 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu 2620
 2625 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu 2635
 2645 Asn Leu Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys 2655
 2660 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu 2670
 2675 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn 2685
 2690 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val 2700
 2705 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met 2715
 2725 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly 2735
 2740 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile 2750
 2755 2760 2765

Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 2770 2775 2780
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 2785 2790 2795 2800
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 2805 2810 2815
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 2820 2825 2830
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 2835 2840 2845
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 2850 2855 2860
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2865 2870 2875 2880
 Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu
 2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe
 3105 3110 3115 3120
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser
 3140 3145 3150
 Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
 3170 3175 3180
 Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195 3200
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
 3205 3210 3215
 Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
 3220 3225 3230
 Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
 3235 3240 3245
 Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
 3250 3255 3260
 Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
 3265 3270 3275 3280
 Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
 3285 3290 3295
 Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
 3300 3305 3310
 Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro

```

          3315          3320          3325
Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
      3330          3335          3340
Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
3345          3350          3355          3360
Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
          3365          3370          3375
Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
          3380          3385          3390

```

<210> SEQ ID NO 17

<211> COMPRIMENTO: 10756

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência
artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10305)

<400> SEQUÊNCIA: 17

agttgttagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaacgta	60
gttctaacag	ttttttaatt	agagagcaga	tctctg	atg	aat aac caa cgg aaa	114
			Met	Asn Asn Gln Arg Lys		
aag gcg aaa aac acg cct ttc aat atg ctg aaa cgc gag aga aac cgc			1		5	162
Lys Ala Lys Asn Thr Pro Phe Asn Met			15		20	
gtg tgc act gtg caa cag ctg aca aag aga ttc tca ctt gga atg ctg			30		35	210
Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu						
cag gga cga gga cca tta aaa ctg ttc atg gcc ctg gtg gcg ttc ctt			45		50	258
Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu						
cgt ttc cta aca atc cca cca aca gca ggg ata ttg aag aga tgg gga			60		65	306
Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly Ile Leu Lys Arg Trp Gly						
aca att aaa aaa tca aaa gct att aat gtt ttg aga ggg ttc agg aaa			75		80	354
Thr Ile Lys Lys Ser Lys Ala Ile Asn Val Leu Arg Gly Phe Arg Lys						
gag att gga agg atg ctg aac atc ttg aat agg aga cgc aga tcc gcg			90		95	402
Glu Ile Gly Arg Met Leu Asn Ile Leu Asn Arg Arg Arg Arg Ser Ala						
ggc acc gga att gca gtc atg att ggc ctg atc gcc agc gta gga gca			105		110	450
Gly Thr Gly Ile Ala Val Met Ile Gly Leu Ile Ala Ser Val Gly Ala						
gtt acc ctc tct aac ttc caa ggg aag gtg atg atg acg gta aat gct			120		125	498
Val Thr Leu Ser Asn Phe Gln Gly Lys Val Met Met Thr Val Asn Ala						
act gac gtc aca gat gtc atc acg att cca aca gct gct gga aag aac			135		140	546
Thr Asp Val Thr Asp Val Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn						
cta tgc att gtc aga gca atg gat gtg gga tac atg tgc gat gat act			145		150	594
Leu Cys Ile Val Arg Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr						
atc act tat gaa tgc cca gtg ctg tgc gct ggt aat gat cca gaa gac			155		160	642
Ile Thr Tyr Glu Cys Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp						
atc gac tgt tgg tgc aca aag tca gca gtc tac gtc agg tat gga aga			170		175	690
Ile Asp Cys Trp Cys Thr Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg						
tgc acc aag aca cgc cac tca aga cgc agt cgg agg tca ctg aca gtg			185		190	738
Cys Thr Lys Thr Arg His Ser Arg Arg Ser Arg Ser Leu Thr Val						
cag aca cac gga gaa agc act cta gcg aac aag aag ggg gct tgg atg			200		205	786
Gln Thr His Gly Glu Ser Thr Leu Ala Asn Lys Lys Gly Ala Trp Met						
gac agc acc aag gcc aca agg tat ttg gta aaa aca gaa tca tgg atc			215		220	834
Asp Ser Thr Lys Ala Thr Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile						

ttg	agg	aac	cct	gga	tat	gcc	ctg	gtg	gca	gcc	gtc	att	ggt	tgg	atg	882
Leu	Arg	Asn	Pro	Gly	Tyr	Ala	Leu	Val	Ala	Ala	Val	Ile	Gly	Trp	Met	
			250					255					260			
ctt	ggg	agc	aac	acc	atg	cag	aga	gtt	gtg	ttt	gtc	gtg	cta	ttg	ctt	930
Leu	Gly	Ser	Asn	Thr	Met	Gln	Arg	Val	Val	Phe	Val	Val	Leu	Leu	Leu	
			265					270					275			
ttg	gtg	gcc	cca	gct	tac	agc	ttc	aac	tgc	ctt	gga	atg	agc	aac	aga	978
Leu	Val	Ala	Pro	Ala	Tyr	Ser	Phe	Asn	Cys	Leu	Gly	Met	Ser	Asn	Arg	
								285					290			
gac	ttc	ttg	gaa	gga	gtg	tct	gga	gca	aca	tgg	gtg	gat	ttg	gtt	ctc	1026
Asp	Phe	Leu	Glu	Gly	Val	Ser	Gly	Ala	Thr	Trp	Val	Asp	Leu	Val	Leu	
								300							310	
gaa	ggc	gac	agc	tgc	gtg	act	atc	atg	tct	aag	gac	aag	cct	acc	atc	1074
Glu	Gly	Asp	Ser	Cys	Val	Thr	Ile	Met	Ser	Lys	Asp	Lys	Pro	Thr	Ile	
															325	
gat	gtg	aag	atg	aat	atg	gag	gcg	gcc	aac	ctg	gca	gag	gtc	cg		1122
Asp	Val	Lys	Met	Met	Asn	Met	Glu	Ala	Ala	Asn	Leu	Ala	Glu	Val	Arg	
															340	
agt	tat	tgc	tat	ttg	gct	acc	gtc	agc	gat	ctc	tcc	acc	aaa	gct	gcg	1170
Ser	Tyr	Cys	Tyr	Leu	Ala	Thr	Val	Ser	Asp	Leu	Ser	Thr	Lys	Ala	Ala	
															355	
tgc	ccg	acc	atg	gga	gaa	gct	cac	aat	gac	aaa	cgt	gct	gac	cca	gct	1218
Cys	Pro	Thr	Met	Gly	Glu	Ala	His	Asn	Asp	Lys	Arg	Ala	Asp	Pro	Ala	
															370	
ttt	gtg	tgc	aga	caa	gga	gtg	gac	agg	ggc	tgg	ggc	aac	ggc	tgc		1266
Phe	Val	Cys	Arg	Gln	Gly	Val	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	
															390	
gga	cta	ttt	ggc	aaa	gga	agc	att	gac	aca	tgc	gcc	aaa	ttt	gcc	tgc	1314
Gly	Leu	Phe	Gly	Lys	Gly	Ser	Ile	Asp	Thr	Cys	Ala	Lys	Phe	Gly	Cys	
															405	
tct	acc	aag	gca	ata	gga	aga	acc	atc	ttg	aaa	gag	aat	atc	aag	tac	1362
Ser	Thr	Lys	Ala	Ile	Gly	Arg	Thr	Ile	Leu	Lys	Glu	Asn	Ile	Lys	Tyr	
															420	
gaa	gtg	gcc	att	ttt	gtc	cat	gga	cca	act	act	gtg	gag	tgc	cac	gga	1410
Glu	Val	Ala	Ile	Phe	Val	His	Gly	Pro	Thr	Thr	Val	Glu	Ser	His	Gly	
															435	
aat	tac	tcc	aca	cag	gtt	gga	gcc	act	cag	gca	ggg	aga	ttc	agc	atc	1458
Asn	Tyr	Ser	Thr	Gln	Val	Gly	Ala	Thr	Gln	Ala	Gly	Arg	Phe	Ser	Ile	
															450	
act	cct	gcg	gcg	cct	tca	tac	aca	cta	aag	ctt	gga	gaa	tat	gga	gag	1506
Thr	Pro	Ala	Ala	Pro	Ser	Tyr	Thr	Leu	Lys	Leu	Gly	Glu	Tyr	Gly	Glu	
															470	
gtg	aca	gtg	gac	tgt	gaa	cca	cgg	tca	ggg	att	gac	acc	aat	gca	tac	1554
Val	Thr	Val	Asp	Cys	Glu	Pro	Arg	Ser	Gly	Ile	Asp	Thr	Asn	Ala	Tyr	
															485	
tac	gtg	atg	act	gtt	gga	aca	aag	acg	ttc	ttg	gtc	cat	cgt	gag	tgg	1602
Tyr	Val	Met	Thr	Val	Gly	Thr	Lys	Thr	Phe	Leu	Val	His	Arg	Glu	Trp	
															500	
ttc	atg	gac	ctc	aac	ctc	cct	tgg	agc	agt	gct	gga	agt	act	gtg	tgg	1650
Phe	Met	Asp	Leu	Asn	Leu	Pro	Trp	Ser	Ser	Ala	Gly	Ser	Thr	Val	Trp	
															515	
agg	aac	aga	gag	acg	tta	atg	gag	ttt	gag	gaa	cca	cac	gcc	acg	aag	1698
Arg	Asn	Arg	Glu	Thr	Leu	Met	Glu	Phe	Glu	Glu	Pro	His	Ala	Thr	Lys	
															530	
cag	tct	gtg	ata	gca	ttg	ggc	tca	caa	gag	gga	gct	ctg	cat	caa	gct	1746
Gln	Ser	Val	Ile	Ala	Leu	Gly	Ser	Gln	Glu	Gly	Ala	Leu	His	Gln	Ala	
															550	
ttg	gct	gga	gcc	att	cct	gtg	gaa	ttt	tca	agc	aac	act	gtc	aag	ttg	1794
Leu	Ala	Gly	Ala	Ile	Pro	Val	Glu	Phe	Ser	Ser	Asn	Thr	Val	Lys	Leu	
															565	
acg	tgc	ggt	cat	ttg	aag	tgt	aga	gtg	aag	atg	gaa	aaa	ttg	cag	ttg	1842
Thr	Ser	Gly	His	Leu	Lys	Cys	Arg	Val	Lys	Met	Glu	Lys	Leu	Gln	Leu	
															580	
aag	gga	aca	acc	tat	ggc	gtc	tgt	tca	aag	gct	ttc	aag	ttt	ctt	ggg	1890
Lys	Gly	Thr	Thr	Tyr	Gly	Val	Cys	Ser	Lys	Ala	Phe	Lys	Phe	Leu	Gly	
															595	
act	ccc	gca	gac	aca	ggc	cac	ggc	act	gtg	gtg	ttg	gaa	ttg	cag	tac	1938
Thr	Pro	Ala	Asp	Thr	Gly	His	Gly	Thr	Val	Val	Leu	Glu	Leu	Gln	Tyr	

600	act	ggc	acg	gat	gga	cct	tgc	aaa	gtt	cct	atc	tcg	tca	gtg	gct	tca	1986
Thr	Gly	Thr	Asp	Gly	Pro	Cys	Lys	Val	Pro	Ile	Ser	Ser	Val	Ala	Ser		
615	ttg	aac	gac	cta	acg	cca	gtg	ggc	aga	ttg	gtc	act	gtc	aac	cct	ttt	2034
Leu	Asn	Asp	Leu	Thr	Pro	Val	Gly	Arg	Leu	Val	Thr	Val	Asn	Pro	Phe		
635	gtt	tca	gtg	gcc	acg	gcc	aac	gct	aag	gtc	ctg	att	gaa	ttg	gaa	cca	2082
Val	Ser	Val	Ala	Thr	Ala	Asn	Ala	Lys	Val	Leu	Ile	Glu	Leu	Glu	Pro		
650	ccc	ttt	gga	gac	tca	tac	ata	gtg	gtg	ggc	aga	gga	gaa	caa	cag	atc	2130
Pro	Phe	Gly	Asp	Ser	Tyr	Ile	Val	Val	Gly	Arg	Arg	Gly	Glu	Gln	Gln	Ile	
665	aat	cac	cat	tgg	cac	aag	tct	gga	agg	agg	att	ggc	aaa	gcc	ttt	aca	2178
Asn	His	His	Trp	His	Lys	Ser	Gly	Ser	Ser	Ser	Ile	Gly	Lys	Ala	Phe	Thr	
680	acc	acc	ctc	aaa	gga	ggc	cag	aga	cta	gcc	gct	cta	gga	gac	aca	gct	2226
Thr	Thr	Leu	Lys	Gly	Ala	Gln	Arg	Leu	Ala	Ala	Leu	Gly	Asp	Thr	Ala		
695	tgg	gac	ttt	gga	tca	gtt	gga	ggg	gtg	ttc	acc	tca	gtt	ggg	aag	gct	2274
Trp	Asp	Phe	Gly	Ser	Val	Gly	Gly	Val	Phe	Thr	Thr	Ser	Val	Gly	Lys	Ala	
715	gtc	cat	caa	gtg	ttc	gga	gga	gca	ttc	cgc	tca	ctg	ttc	gga	ggc	atg	2322
Val	His	Gln	Val	Phe	Gly	Gly	Ala	Phe	Arg	Ser	Leu	Phe	Gly	Gly	Met		
730	tcc	tgg	ata	acg	caa	gga	ttg	ctg	ggg	gct	ctc	ctg	ttg	tgg	atg	ggc	2370
Ser	Trp	Ile	Thr	Gln	Gly	Leu	Leu	Gly	Ala	Leu	Leu	Leu	Leu	Trp	Met	Gly	
745	atc	aat	gct	cgt	gat	agg	tcc	ata	gct	ctc	acg	ttt	ctc	gca	gcc	ggc	2418
Ile	Asn	Ala	Arg	Asp	Arg	Ser	Ile	Ala	Leu	Thr	Thr	Phe	Leu	Ala	Ala	Gly	
760	att	gtg	aca	ctg	tat	ttg	ggg	gtc	atg	gtg	cag	gcc	gat	agt	ggt	tgc	2466
Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln	Gln	Ala	Asp	Ser	Gly	Cys	
775	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	ggc	agt	ggg	att	ttc	2514
Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys	Gly	Ser	Gly	Ile	Phe		
795	atc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa	tac	aag	ttc	caa	cca	2562
Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Thr	Glu	Gln	Tyr	Lys	Phe	Gln	Pro	
810	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	aaa	gcc	cat	gaa	gag	2610
Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ser	Ala	Ile	Gln	Lys	Ala	His	Glu	Glu	
825	ggc	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	gag	aat	ctg	atg	tgg	2658
Gly	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Thr	Arg	Leu	Glu	Asn	Leu	Met	Trp	
840	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	tca	gaa	aat	gag	gtg	2706
Lys	Gln	Ile	Thr	Pro	Pro	Leu	Asn	His	Ile	Ile	Leu	Ser	Glu	Asn	Glu	Val	
855	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	atc	atg	cag	gca	gga	2754
Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	Ile	Met	Gln	Ala	Gly		
875	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	aag	tat	tca	tgg	aaa	2802
Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Pro	Thr	Glu	Leu	Lys	Tyr	Ser	Trp	Lys	
890	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	tct	cat	aac	cag	acc	2850
Thr	Trp	Gly	Lys	Ala	Lys	Met	Met	Leu	Ser	Thr	Glu	Ser	His	Asn	Gln	Thr	
905	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	ccc	aac	aca	aat	aga	2898
Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Glu	Cys	Pro	Asn	Thr	Asn	Arg	
920	gct	tgg	aat	tcg	ttg	gaa	gtt	gaa	gac	tat	ggc	ttt	gga	gta	ttc	acc	2946
Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly	Phe	Gly	Val	Phe	Thr		
935	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	gat	gta	ttc	tgc	gac		2994
Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln	Asp	Val	Phe	Cys	Asp		
955	tca	aaa	ctc	atg	tca	ggc	gcc	ata	aaa	gac	aac	aga	gcc	gtc	cat	gcc	3042
Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn	Arg	Ala	Val	His	Ala		

gat	atg	ggt	tat	tgg	ata	gaa	agt	gca	ctc	aat	gac	aca	tgg	aag	ata	3090
Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn	Asp	Thr	Trp	Lys	Ile	
970	985	990	995	1000	1005	1010	1015	1020	1025	1030	1035	1040	1045	1050	1055	
gag	aaa	gcc	tct	ttc	att	gaa	ggt	aaa	aac	tgc	cac	tgg	cca	aaa	tca	3138
Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys	His	Trp	Pro	Lys	Ser	
1060	1065	1070	1075	1080	1085	1090	1095	1100	1105	1110	1115	1120	1125	1130	1135	
cac	acc	ctc	tgg	agc	aat	gga	gtg	cta	gaa	agt	gag	atg	ata	att	cca	3186
His	Thr	Leu	Trp	Ser	Asn	Gly	Val	Leu	Glu	Ser	Glu	Met	Ile	Ile	Pro	
1140	1145	1150	1155	1160	1165	1170	1175	1180	1185	1190	1195	1200	1205	1210	1215	
aag	aat	ctc	gct	gga	cca	gtg	tct	caa	cac	aac	tat	aga	cca	ggc	tac	3234
Lys	Asn	Leu	Ala	Gly	Pro	Val	Ser	Gln	His	Asn	Tyr	Arg	Pro	Gly	Tyr	
1220	1225	1230	1235	1240	1245	1250	1255	1260	1265	1270	1275	1280	1285	1290	1295	
cat	aca	caa	ata	aca	gga	cca	tgg	cat	cta	ggt	aag	ctt	gag	atg	gac	3282
His	Thr	Gln	Ile	Thr	Gly	Pro	Trp	His	Leu	Gly	Lys	Leu	Glu	Met	Asp	
1300	1305	1310	1315	1320	1325	1330	1335	1340	1345	1350	1355	1360	1365	1370	1375	
ttt	gat	ttc	tgt	gat	gga	aca	aca	gtg	gta	gtg	act	gag	gac	tgc	gga	3330
Phe	Asp	Phe	Cys	Asp	Gly	Thr	Thr	Val	Val	Val	Thr	Glu	Asp	Cys	Gly	
1380	1385	1390	1395	1400	1405	1410	1415	1420	1425	1430	1435	1440	1445	1450	1455	
aat	aga	gga	ccc	tct	ttg	aga	aca	acc	act	gcc	tct	gga	aaa	ctc	ata	3378
Asn	Arg	Gly	Pro	Ser	Leu	Arg	Thr	Thr	Thr	Ala	Ser	Gly	Lys	Leu	Ile	
1460	1465	1470	1475	1480	1485	1490	1495	1500	1505	1510	1515	1520	1525	1530	1535	
aca	gaa	tgg	tgc	tgc	cga	tct	tgc	aca	tta	cca	ccg	cta	aga	tac	aga	3426
Thr	Glu	Trp	Cys	Cys	Arg	Ser	Cys	Thr	Leu	Pro	Pro	Leu	Arg	Tyr	Arg	
1540	1545	1550	1555	1560	1565	1570	1575	1580	1585	1590	1595	1600	1605	1610	1615	
ggt	gag	gat	ggg	tgc	tgg	tac	ggg	atg	gaa	atc	aga	cca	ttg	aag	gag	3474
Gly	Glu	Asp	Gly	Cys	Trp	Tyr	Gly	Met	Glu	Ile	Arg	Pro	Leu	Lys	Glu	
1620	1625	1630	1635	1640	1645	1650	1655	1660	1665	1670	1675	1680	1685	1690	1695	
aaa	gaa	gag	aat	ttg	gtc	aac	tcc	ttg	gtc	aca	gct	gga	cat	ggg	cag	3522
Lys	Glu	Glu	Asn	Leu	Val	Asn	Ser	Leu	Val	Thr	Ala	Gly	His	Gly	Gln	
1700	1705	1710	1715	1720	1725	1730	1735	1740	1745	1750	1755	1760	1765	1770	1775	
gtc	gac	aac	ttt	tca	cta	gga	gtc	ttg	gga	atg	gca	ttg	ttc	ctg	gag	3570
Val	Asp	Asn	Phe	Ser	Leu	Gly	Val	Leu	Gly	Met	Ala	Leu	Phe	Leu	Glu	
1780	1785	1790	1795	1800	1805	1810	1815	1820	1825	1830	1835	1840	1845	1850	1855	
gaa	atg	ctt	agg	acc	cga	gta	gga	acg	aaa	cat	gca	ata	cta	cta	gtt	3618
Glu	Met	Leu	Arg	Thr	Arg	Val	Gly	Thr	Lys	His	Ala	Ile	Leu	Leu	Val	
1860	1865	1870	1875	1880	1885	1890	1895	1900	1905	1910	1915	1920	1925	1930	1935	
gca	gtt	tct	ttt	gtg	aca	ttg	atc	aca	ggg	aac	atg	tcc	ttt	aga	gac	3666
Ala	Val	Ser	Phe	Val	Thr	Leu	Ile	Thr	Gly	Asn	Met	Ser	Phe	Arg	Asp	
1940	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	
ctg	gga	aga	gtg	atg	gtt	atg	gta	ggc	gcc	act	atg	acg	gat	gac	ata	3714
Leu	Gly	Arg	Val	Met	Val	Met	Val	Gly	Ala	Thr	Met	Thr	Asp	Asp	Ile	
2020	2025	2030	2035	2040	2045	2050	2055	2060	2065	2070	2075	2080	2085	2090	2095	
ggt	atg	ggc	gtg	act	tat	ctt	gcc	cta	cta	gca	gcc	ttc	aaa	gtc	aga	3762
Gly	Met	Gly	Val	Thr	Tyr	Leu	Ala	Leu	Leu	Ala	Ala	Phe	Lys	Val	Arg	
2100	2105	2110	2115	2120	2125	2130	2135	2140	2145	2150	2155	2160	2165	2170	2175	
cca	act	ttt	gca	gct	gga	cta	ctc	ttg	aga	aag	ctg	acc	tcc	aag	gaa	3810
Pro	Thr	Phe	Ala	Ala	Gly	Leu	Leu	Leu	Arg	Lys	Leu	Thr	Ser	Lys	Glu	
2180	2185	2190	2195	2200	2205	2210	2215	2220	2225	2230	2235	2240	2245	2250	2255	
ttg	atg	atg	act	act	ata	gga	att	gta	ctc	ctc	tcc	cag	agc	acc	ata	3858
Leu	Met	Met	Thr	Thr	Ile	Gly	Ile	Val	Leu	Leu	Ser	Gln	Ser	Thr	Ile	
2260	2265	2270	2275	2280	2285	2290	2295	2300	2305	2310	2315	2320	2325	2330	2335	
cca	gag	acc	att	ctt	gag	ttg	act	gat	gcg	tta	gcc	tta	ggc	atg	atg	3906
Pro	Glu	Thr	Ile	Leu	Glu	Leu	Thr	Asp	Ala	Leu	Ala	Leu	Gly	Met	Met	
2340	2345	2350	2355	2360	2365	2370	2375	2380	2385	2390	2395	2400	2405	2410	2415	
gtc	ctc	aaa	atg	gtg	aga	aat	atg	gaa	aag	tat	caa	ttg	gca	gtg	act	3954
Val	Leu	Lys	Met	Val	Arg	Asn	Met	Glu	Lys	Tyr	Gln	Leu	Ala	Val	Thr	
2420	2425	2430	2435	2440	2445	2450	2455	2460	2465	2470	2475	2480	2485	2490	2495	
atc	atg	gct	atc	ttg	tgc	gtc	cca	aac	gca	gtg	ata	tta	caa	aac	gca	4002
Ile	Met	Ala	Ile	Leu	Cys	Val	Pro	Asn	Ala	Val	Ile	Leu	Gln	Asn	Ala	
2500	2505	2510	2515	2520	2525	2530	2535	2540	2545	2550	2555	2560	2565	2570	2575	
tgg	aaa	gtg	agt	tgc	aca	ata	ttg	gca	gtg	gtg	tcc	gtt	tcc	cca	ctg	4050
Trp	Lys	Val	Ser	Cys	Thr	Ile	Leu	Ala	Val	Val	Ser	Val	Ser	Pro	Leu	
2580	2585	2590	2595	2600	2605	2610	2615	2620	2625	2630	2635	2640	2645	2650	2655	
ctc	tta	aca	tcc	tca	cag	caa	aaa	aca	gat	tgg	ata	cca	tta	gca	ttg	4098
Leu	Leu	Thr	Ser	Ser	Gln	Gln	Lys	Thr	Asp	Trp	Ile	Pro	Leu	Ala	Leu	
2660	2665	2670	2675	2680	2685	2690	2695	2700	2705	2710	2715	2720	2725	2730	2735	
acg	atc	aaa	ggt	ctc	aat	cca	aca	gct	att	ttt	cta	aca	acc	ctc	tca	4146
Thr	Ile	Lys	Gly	Leu	Asn	Pro	Thr	Ala	Ile	Phe	Leu	Thr	Thr	Leu	Ser	

1335	aga acc agc aag aaa agg agc tgg cca tta aat gag gct atc atg gca	1340	1345	1350	4194
Arg Thr Ser Lys	Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala	1355	1360	1365	
gtc ggg atg gtg agc att tta gcc agt tct ctc cta aaa aat gat att	4242	Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile	1370	1375	1380
ccc atg aca gga cca tta gtg gct gga ggg ctc ctc act gtg tgc tac	4290	Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr	1385	1390	1395
gtg ctc act gga cga tcg gcc gat ttg gaa ctg gag aga gca gcc gat	4338	Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp	1400	1405	1410
gtc aaa tgg gaa gac cag gca gag ata tca gga agc agt cca atc ctg	4386	Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu	1415	1420	1425
tca ata aca ata tca gaa gat ggt agc atg tcg ata aaa aat gaa gag	4434	Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu	1430	1435	1440
gaa gaa caa aca ctg acc ata ctc att aga aca gga ttg ctg gtg atc	4482	Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile	1445	1450	1455
tca gga ctt ttt cct gta tca ata cca atc acg gca gca gca tgg tac	4530	Ser Gly Leu Phe Pro Val Ser Ile Pro Ile Thr Ala Ala Trp Tyr	1460	1465	1470
ctg tgg gaa gtg aag aaa caa cgg gcc gga gta ttg tgg gat gtt cct	4578	Leu Trp Glu Val Lys Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro	1475	1480	1485
tca ccc cca ccc atg gga aag gct gaa ctg gaa gat gga gcc tat aga	4626	Ser Pro Pro Pro Met Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg	1490	1495	1500
att aag caa aaa ggg att ctt gga tat tcc cag atc gga gcc gga gtt	4674	Ile Lys Gln Lys Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val	1505	1510	1515
tac aaa gaa gga aca ttc cat aca atg tgg cat gtc aca cgt ggc gct	4722	Tyr Lys Glu Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala	1520	1525	1530
gtt cta atg cat aaa gga aag agg att gaa cca tca tgg gcg gac gtc	4770	Val Leu Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val	1535	1540	1545
aag aaa gac cta ata tca tat gga ggc tgg aag tta gaa gga gaa	4818	Lys Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu	1550	1555	1560
tgg aag gaa gga gaa gaa gtc cag gta ttg gca ctg gag cct gga aaa	4866	Trp Lys Glu Gly Glu Gln Val Gln Val Leu Ala Leu Pro Gly Lys	1565	1570	1575
aat cca aga gcc gtc caa acg aaa cct ggt ctt ttc aaa acc aac gcc	4914	Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr Ala Ala	1580	1585	1590
gga aca ata ggt gct gta tct ctg gac ttt tct cct gga acg tca gga	4962	Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly	1595	1600	1605
tct cca att atc gac aaa aaa gga aaa gtt gtg ggt ctt tat ggt aat	5010	Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn	1610	1615	1620
ggt gtt gtt aca agg agt gga gca tat gtg agt gct ata gcc cag act	5058	Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr	1625	1630	1635
gaa aaa agc att gaa gac aac cca gag atc gaa gat gac att ttc cga	5106	Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg	1640	1645	1650
aag aga aga ctg acc atc atg gac ctc cac cca gga gcg gga aag acg	5154	Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr	1655	1660	1665
aag aga tac ctt ccg gcc ata gtc aga gaa gct ata aaa cgg ggt ttg	5202	Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu	1670	1675	1680
aga aca tta atc ttg gcc ccc act aga gtt gtg gca gct gaa atg gag	5250	Arg Thr Leu Ile Leu Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu	1685	1690	1695

2075										2080										2085										
tat	tct	gac	cca	ctg	gcg	cta	aaa	gaa	ttt	aag	gaa	ttt	gca	gcc	gga		6402													
Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	Glu	Phe	Ala	Ala	Gly															
2090																														
aga	aag	tct	ctg	acc	ctg	aac	cta	atc	aca	gaa	atg	ggg	agg	ctc	cca		6450													
Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	Met	Gly	Arg	Leu	Pro															
2105																														
acc	ttc	atg	act	cag	aag	gca	aga	gac	gca	ctg	gac	aac	tta	gca	gtg		6498													
Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu	Asp	Asn	Leu	Ala	Val															
2120																														
ctg	cac	acg	gct	gag	gca	ggg	gga	agg	gcg	tac	aac	cat	gct	ctc	agt		6546													
Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala	Tyr	Asn	His	Ala	Leu	Ser															
2135																														
gaa	ctg	ccg	gag	acc	ctg	gag	aca	ttg	ctt	tta	ctg	aca	ctt	ctg	gct		6594													
Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu	Leu	Leu	Leu	Thr	Leu	Leu	Ala															
2155																														
aca	gtc	acg	gga	ggg	atc	ttt	tta	ttc	ttg	atg	agc	gga	agg	ggc	ata		6642													
Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu	Phe	Leu	Met	Ser	Gly	Arg	Gly	Ile															
2170																														
ggg	aag	atg	acc	ctg	gga	atg	tgc	tgc	ata	atc	acg	gct	agc	atc	ctc		6690													
Gly	Lys	Met	Thr	Leu	Gly	Met	Cys	Cys	Ile	Ile	Thr	Ala	Ser	Ile	Leu															
2185																														
cta	tgg	tac	gca	caa	ata	cag	cca	cac	tgg	ata	gca	gct	tca	ata	ata		6738													
Leu	Trp	Tyr	Ala	Gln	Ile	Gln	Pro	His	Trp	Ile	Ala	Ala	Ser	Ile	Ile															
2200																														
ctg	gag	ttt	ttt	ctc	ata	gtt	ttg	ctt	att	cca	gaa	cct	gaa	aaa	cag		6786													
Leu	Glu	Phe	Phe	Leu	Ile	Val	Leu	Leu	Ile	Pro	Glu	Pro	Glu	Lys	Gln															
2215																														
aga	aca	ccc	caa	gac	aac	caa	ctg	acc	tac	gtt	gtc	ata	gcc	atc	ctc		6834													
Arg	Thr	Pro	Gln	Asp	Asn	Gln	Leu	Thr	Tyr	Val	Val	Ile	Ala	Ile	Leu															
2235																														
aca	gtg	gtg	gcc	gca	acc	atg	gca	aac	gag	atg	ggg	ttc	cta	gaa	aaa		6882													
Thr	Val	Val	Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys															
2250																														
acg	aag	aaa	gat	ctc	gga	ttg	gga	agc	att	gca	acc	cag	caa	ccc	gag		6930													
Thr	Lys	Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu															
2265																														
agc	aac	atc	ctg	gac	ata	gat	cta	cg	cct	gca	tca	gca	tgg	acg	ctg		6978													
Ser	Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu															
2280																														
tat	gcc	gtg	gcc	aca	aca	ttt	gtt	aca	cca	atg	ttg	aga	cat	agc	att		7026													
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser	Ile															
2295																														
gaa	aat	tcc	tca	gtg	aat	gtg	tcc	cta	aca	gct	ata	gcc	aac	caa	gcc		7074													
Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn	Gln	Ala															
2315																														
aca	gtg	tta	atg	ggg	ctc	ggg	aaa	gga	tgg	cca	ttg	tca	aag	atg	gac		7122													
Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser	Lys	Met	Asp															
2330																														
atc	gga	gtt	ccc	ctt	ctc	gcc	att	gga	tgc	tac	tca	caa	gtc	aac	ccc		7170													
Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser	Gln	Val	Asn	Pro															
2345																														
ata	act	ctc	aca	gca	gct	ctt	ttc	tta	ttg	gta	gca	cat	tat	gcc	atc		7218													
Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	Ala	His	Tyr	Ala	Ile															
2360																														
ata	ggg	cca	gga	ctc	caa	gca	aaa	gca	acc	aga	gaa	gct	cag	aaa	aga		7266													
Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	Glu	Ala	Gln	Lys	Arg															
2375																														
gca	gcg	gcg	ggc	atc	atg	aaa	aac	cca	act	gtc	gat	gga	ata	aca	gtg		7314													
Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val	Asp	Gly	Ile	Thr	Val															
2395																														
att	gac	cta	gat	cca	ata	cct	tat	gat	cca	aag	ttt	gaa	aag	cag	ttg		7362													
Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr	Asp	Pro	Lys	Phe	Glu	Lys	Gln	Leu															
2410																														
gga	caa	gta	atg	ctc	cta	gtc	ctc	tgc	gtg	act	caa	gta	ttg	atg	atg		7410													
Gly	Gln	Val	Met	Leu	Leu	Val	Leu	Cys	Val	Thr	Gln	Val	Leu	Met	Met															
2425																														
agg	act	aca	tgg	gct	ctg	tgt	gag	gct	tta	acc	tta	gct	acc	ggg	ccc		7458													
Arg	Thr	Thr	Trp	Ala	Leu	Cys	Glu	Ala	Leu	Thr	Leu	Ala	Thr	Gly	Pro															

2440	atc	tcc	aca	ttg	tgg	gaa	gga	aat	cca	ggg	agg	ttt	tgg	aac	act	acc	7506
2455	Ile	Ser	Thr	Leu	Trp	Glu	Gly	Asn	Pro	Gly	Arg	Phe	Trp	Asn	Thr	Thr	
2460	att	gcg	gtg	tca	atg	gct	aac	att	ttt	aga	ggg	agt	tac	ttg	gcc	gga	7554
2475	Ile	Ala	Val	Ser	Met	Ala	Asn	Ile	Phe	Arg	Gly	Ser	Tyr	Leu	Ala	Gly	
2480	gct	gga	ctt	ctc	ttt	tct	att	atg	aag	aac	aca	acc	aac	aca	aga	agg	7602
2490	Ala	Gly	Leu	Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	Thr	Asn	Thr	Arg	Arg	
2500	gga	act	ggc	aac	ata	gga	gag	acg	ctt	gga	gag	aaa	tgg	aaa	agc	cga	7650
2510	Gly	Thr	Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	Lys	Trp	Lys	Ser	Arg	
2520	ttg	aac	gca	ttg	gga	aaa	agt	gaa	ttc	cag	atc	tac	aag	aaa	agt	gga	7698
2535	Leu	Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	Tyr	Lys	Lys	Ser	Gly	
2540	atc	cag	gaa	gtg	gat	aga	acc	tta	gca	aaa	gaa	ggc	att	aaa	aga	gga	7746
2555	Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	Gly	Ile	Lys	Arg	Gly	
2560	gaa	acg	gac	cat	cac	gct	gtg	tcg	cga	ggc	tca	gca	aaa	ctg	aga	tgg	7794
2570	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	Ala	Lys	Leu	Arg	Trp	
2580	ttc	gtt	gag	aga	aac	atg	gtc	aca	cca	gaa	ggg	aaa	gta	gtg	gac	ctc	7842
2590	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	Lys	Val	Val	Asp	Leu	
2600	ggt	tgt	ggc	aga	gga	ggc	tgg	tca	tac	tat	tgt	gga	gga	cta	aag	aat	7890
2615	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	Gly	Gly	Leu	Lys	Asn	
2620	gta	aga	gaa	gtc	aaa	ggc	cta	aca	aaa	gga	gga	cca	gga	cac	gaa	gaa	7938
2630	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly	Pro	Gly	His	Glu	Glu	
2640	ccc	atc	ccc	atg	tca	aca	tat	ggg	tgg	aat	cta	gtg	cgt	ctt	caa	agt	7986
2655	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn	Leu	Val	Arg	Leu	Gln	Ser	
2660	gga	gtt	gac	gtt	ttc	ttc	atc	ccg	cca	gaa	aag	tgt	gac	aca	tta	ttg	8034
2670	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro	Glu	Lys	Cys	Asp	Thr	Leu	Leu	
2680	tgt	gac	ata	ggg	gag	tca	tca	cca	aat	ccc	aca	gtg	gaa	gca	gga	cga	8082
2695	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro	Thr	Val	Glu	Ala	Gly	Arg	
2700	aca	ctc	aga	gtc	ctt	aac	tta	gta	gaa	aat	tgg	ttg	aac	aac	aac	act	8130
2715	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu	Asn	Trp	Leu	Asn	Asn	Asn	Thr	
2720	caa	ttt	tgc	ata	aag	gtt	ctc	aac	cca	tat	atg	ccc	tca	gtc	ata	gaa	8178
2730	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn	Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	
2740	aaa	atg	gaa	gca	cta	caa	agg	aaa	tat	gga	gga	gcc	tta	gtg	agg	aat	8226
2755	Lys	Met	Glu	Ala	Leu	Gln	Arg	Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	
2760	cca	ctc	tca	cga	aac	tcc	aca	cat	gag	atg	tac	tgg	gta	tcc	aat	gct	8274
2770	Pro	Leu	Ser	Arg	Asn	Ser	Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	
2780	tcc	ggg	aac	ata	gtg	tca	tca	gtg	aac	atg	att	tca	agg	atg	ttg	atc	8322
2795	Ser	Gly	Asn	Ile	Val	Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	
2800	aac	aga	ttt	aca	atg	aga	tac	aag	aaa	gcc	act	tac	gag	ccg	gat	gtt	8370
2815	Asn	Arg	Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	
2820	gac	ctc	gga	agc	gga	acc	cgt	aac	atc	ggg	att	gaa	agt	gag	ata	cca	8418
2835	Asp	Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	Glu	Ser	Glu	Ile	Pro	
2840	aac	cta	gat	ata	att	ggg	aaa	aga	ata	gaa	aaa	ata	aag	caa	gag	cat	8466
2855	Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	Ile	Lys	Gln	Glu	His	
2860	gaa	aca	tca	tgg	cac	tat	gac	caa	gac	cac	cca	tac	aaa	acg	tgg	gca	8514
2875	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	Tyr	Lys	Thr	Trp	Ala	
2880	tac	cat	ggt	agc	tat	gaa	aca	aaa	cag	act	gga	tca	gca	tca	tcc	atg	8562
2895	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	Ser	Ala	Ser	Ser	Met	


```

3175          3180          3185          3190
aaa gac ata caa caa tgg gaa cct tca aga gga tgg aat gat tgg aca 9714
Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr
          3195          3200          3205
caa gtg ccc ttc tgt tca cac cat ttc cat gag tta atc atg aaa gac 9762
Gln Val Pro Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp
          3210          3215          3220
ggg cgc gta ctc gtt gtt cca tgt aga aac caa gat gaa ctg att ggc 9810
Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly
          3225          3230          3235
aga gcc cga atc tcc caa gga gca ggg tgg tct ttg cgg gag acg gcc 9858
Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
          3240          3245          3250
tgt ttg ggg aag tct tac gcc caa atg tgg agc ttg atg tac ttc cac 9906
Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His
          3255          3260          3265          3270
aga cgc gac ctc agg ctg gcg gca aat gct att tgc tcg gca gta cca 9954
Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro
          3275          3280          3285
tca cat tgg gtt cca aca agt cga aca acc tgg tcc ata cat gct aaa 10002
Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys
          3290          3295          3300
cat gaa tgg atg aca acg gaa gac atg ctg aca gtc tgg aac agg gtg 10050
His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val
          3305          3310          3315
tgg att caa gaa aac cca tgg atg gaa gac aaa act cca gtg gaa tca 10098
Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser
          3320          3325          3330
tgg gag gaa atc cca tac ttg ggg aaa aga gaa gac caa tgg tgc ggc 10146
Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly
          3335          3340          3345          3350
tca ttg att ggg tta aca agc agg gcc acc tgg gca aag aac atc caa 10194
Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln
          3355          3360          3365
gca gca ata aat caa gtt aga tcc ctt ata ggc aat gaa gaa tac aca 10242
Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr
          3370          3375          3380
gat tac atg cca tcc atg aaa aga ttc aga aga gaa gag gaa gaa gca 10290
Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala
          3385          3390          3395
gga gtt ctg tgg tag aaagcaaaac taacatgaaa caaggctaga agtcaggctcg 10345
Gly Val Leu Trp *
          3400
gattaagcca tagtacggaa aaaactatgc tacctgtgag ccccgccaa ggacgttaaa 10405
agaagtcagg ccatacaaa tgccatagct tgagtaaaact atgcagcctg tagctccacc 10465
tgagaagggtg taaaaaatcc gggaggccac aaaccatgga agctgtacgc atggcgtagt 10525
ggactagcgg ttagaggaga cccctccctt acaaatcgca gcaacaatgg gggcccaagg 10585
cgagatgaag ctgtagtctc gctggaagga ctagagggtta gaggagaccc ccccgaaaca 10645
aaaaacagca tattgacgct gggaaagacc agagatcctg ctgtctcctc agcatcattc 10705
caggcacaga acgccagaaa atggaatggt gctgttgaat caacagggttc t 10756

```

<210> SEQ ID NO 18

<211> COMPRIMENTO: 3402

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 18

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

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 Arg Arg Arg Arg 85 Ser Ala Gly Thr Gly 90 Ile Ala Val Met Ile Gly Leu
 Ile Ala Ser Val 100 Gly Ala Val Thr 105 Leu Ser Asn Phe Gln Gly Lys Val
 Met Met Thr Val 115 Asn Ala Thr 120 Asp Val Thr Asp Val 125 Thr Ile Pro
 Thr Ala Ala Gly Lys Asn 130 Leu Cys Ile Val Arg Ala Met Asp Val Gly
 Tyr Met Cys Asp Asp Thr 135 Ile Thr Tyr Glu Cys Pro Val Leu Ser Ala
 Gly Asn Asp Pro 145 Glu Asp Ile Asp Cys Trp Cys Thr Lys Ser Ala Val
 Tyr Val Arg Tyr 155 Gly Arg Cys Thr Lys Thr Arg His Ser Arg Arg Ser
 Arg Arg Ser Leu Thr Val Gln 165 Thr His Gly Glu Ser 170 Thr Leu Ala Asn
 Lys Lys Gly Ala Trp Met 185 Asp Ser Thr Lys Ala Thr Arg Tyr Leu Val
 Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly Tyr Ala Leu Val Ala
 Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr Met Gln Arg Val Val
 Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys
 Leu Gly Met Ser Asn Arg Asp 245 Phe Leu Glu Gly Val Ser Gly Ala Thr
 Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr Ile Met Ser
 Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met Glu Ala Ala
 Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val Ser Asp
 Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His Asn Asp
 Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val Asp Arg
 Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr
 Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr Ile Leu
 Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly Pro Thr
 Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala Thr Gln
 Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr Leu Lys
 Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg Ser Gly
 Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys Thr Phe
 Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp Ser Ser
 Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu Phe Glu
 Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser Gln Glu
 Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu Phe Ser
 Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys
 Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys
 Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr Val
 Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val Pro
 Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu

625	Val	Thr	Val	Asn	Pro	630	Phe	Val	Ser	Val	635	Ala	Thr	Ala	Asn	Ala	Lys	640	Val
					645							650					655		
Leu	Ile	Glu	Leu	Glu	Pro	Pro	Phe	Gly	660	Asp	Ser	Tyr	Ile	Val	670	Val	Gly		
Arg	Gly	Glu	Gln	Gln	Ile	Asn	His	His	665	Trp	His	Lys	Ser	Gly	675	Ser	Ser		
Ile	Gly	Lys	Ala	Phe	Thr	Thr	680	Leu	Lys	Gly	Ala	Gln	Arg	Leu	685	Ala			
Ala	Leu	Gly	Asp	Thr	Ala	Trp	690	Asp	Phe	Gly	Ser	Val	Gly	Gly	700	Val	Phe		
705	Thr	Ser	Val	Gly	Lys	Ala	Val	His	Gln	Val	715	Phe	Gly	Gly	720	Ala	Phe		
					725												735	Arg	
Ser	Leu	Phe	Gly	Met	Ser	Trp	Ile	Thr	730	Gln	Gly	Leu	Leu	Gly	740	Ala			
Leu	Leu	Leu	Trp	Met	Gly	Ile	Asn	Ala	745	Arg	Asp	Arg	Ser	Ile	750	Ala	Leu		
Thr	Phe	Leu	Ala	Ala	Gly	Ile	760	Val	Thr	Leu	Tyr	Leu	Gly	Val	765	Met	Val		
Gln	Ala	Asp	Ser	Gly	Cys	Val	770	Val	Ser	Trp	Lys	Asn	Lys	Glu	775	Lys			
Cys	Gly	Ser	Gly	Ile	805	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	810	Thr	Glu		
Gln	Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	820	Ser	Lys	Leu	Ala	Ser	825	Ala	Ile		
Gln	Lys	Ala	His	Glu	Glu	Gly	Ile	Cys	830	Gly	Ile	Arg	Ser	Val	835	Thr	Arg		
Leu	Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	840	Thr	Pro	Glu	Leu	Asn	845	His	Ile		
Leu	Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	850	Ile	Met	Thr	Gly	Asp	855	Ile	Lys		
Gly	Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	860	Leu	Arg	Pro	Gln	Pro	865	Thr	Glu		
Leu	Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	870	Lys	Ala	Lys	Met	Leu	875	Ser	Thr		
Glu	Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	880	Asp	Gly	Pro	Glu	Val	885	Ala	Glu		
Cys	Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	890	Ser	Leu	Glu	Val	Glu	895	Asp	Tyr		
Gly	Phe	Gly	Val	Phe	Thr	930	Asn	Ile	935	Trp	Leu	Lys	Leu	Lys	940	Glu	Lys		
Gln	Asp	Val	Phe	Cys	Asp	945	Ser	Lys	950	Leu	Met	Ser	Ala	Ala	955	Ile	Asp		
Asn	Arg	Ala	Val	His	Ala	960	Asp	Met	965	Gly	Tyr	Trp	Ile	Glu	970	Ser	Ala		
Asn	Asp	Thr	Trp	Lys	Ile	980	Glu	Lys	985	Ala	Ser	Phe	Ile	Glu	990	Val	Lys		
Cys	His	Trp	Pro	Lys	Ser	995	His	Thr	1000	Leu	Trp	Ser	Asn	Gly	1005	Val	Leu		
Ser	Glu	Met	Ile	Ile	Pro	1010	Lys	Asn	1015	Leu	Ala	Gly	Pro	Val	1020	Ser	Gln		
Asn	Tyr	Arg	Pro	Gly	Tyr	1025	His	Thr	1030	Gln	Ile	Thr	Gly	Pro	1035	Trp	His		
Gly	Lys	Leu	Glu	Met	Asp	1040	Phe	Asp	1045	Phe	Cys	Asp	Gly	Thr	1050	Thr	Val		
Val	Thr	Glu	Asp	Cys	Gly	1055	Asn	Gly	1060	Pro	Ser	Leu	Arg	Thr	1065	Thr	Thr		
Ala	Ser	Gly	Lys	Leu	Ile	1070	Thr	Glu	1075	Trp	Cys	Cys	Arg	Ser	1080	Cys	Thr		
Pro	Pro	Leu	Arg	Tyr	Arg	1085	Gly	Glu	1090	Asp	Gly	Cys	Trp	Tyr	1095	Gly	Met		
Ile	Arg	Pro	Leu	Lys	Glu	1100	Lys	Glu	1105	Asn	Leu	Val	Asn	Ser	1110	Leu	Val		
Thr	Ala	Gly	His	Gly	Gln	1115	Val	Asp	1120	Asn	Phe	Ser	Leu	Gly	1125	Val	Gly		
Met	Ala	Leu	Phe	Leu	Glu	1130	Glu	Met	1135	Leu	Arg	Thr	Arg	Val	1140	Gly	Thr		
His	Ala	Ile	Leu	Leu	Val	1145	Ala	Val	1150	Ser	Phe	Val	Thr	Leu	1155	Ile	Thr		
						1160			1165						1170		Gly		
						1175									1180				

Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly Ala
 1185 1190 1195 1200
 Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala Leu Leu
 1205 1210 1215
 Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu Leu Leu Arg
 1220 1225 1230
 Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile Gly Ile Val Leu
 1235 1240 1245
 Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu Thr Asp Ala
 1250 1255 1260
 Leu Ala Leu Gly Met Met Val Leu Lys Met Val Arg Asn Met Glu Lys
 1265 1270 1275 1280
 Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu Cys Val Pro Asn Ala
 1285 1290 1295
 Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys Thr Ile Leu Ala Val
 1300 1305 1310
 Val Ser Val Ser Pro Leu Leu Leu Thr Ser Ser Gln Gln Lys Thr Asp
 1315 1320 1325
 Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile
 1330 1335 1340
 Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu
 1345 1350 1355 1360
 Asn Glu Ala Ile Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser
 1365 1370 1375
 Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly
 1380 1385 1390
 Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu
 1395 1400 1405
 Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
 1410 1415 1420
 Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met
 1425 1430 1435 1440
 Ser Ile Lys Asn Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg
 1445 1450 1455
 Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro Ile
 1460 1465 1470
 Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg Ala Gly
 1475 1480 1485
 Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys Ala Glu Leu
 1490 1495 1500
 Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile Leu Gly Tyr Ser
 1505 1510 1515 1520
 Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr Phe His Thr Met Trp
 1525 1530 1535
 His Val Thr Arg Gly Ala Val Leu Met His Lys Gly Lys Arg Ile Glu
 1540 1545 1550
 Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile Ser Tyr Gly Gly Gly
 1555 1560 1565
 Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu Glu Val Gln Val Leu
 1570 1575 1580
 Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly
 1585 1590 1595 1600
 Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe
 1605 1610 1615
 Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val
 1620 1625 1630
 Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val
 1635 1640 1645
 Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
 1650 1655 1660
 Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His
 1665 1670 1675 1680
 Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu
 1685 1690 1695
 Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg Val
 1700 1705 1710
 Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile Arg Tyr
 1715 1720 1725
 Gln Thr Pro Ala Ile Arg Ala Glu His Thr Gly Arg Glu Ile Val Asp

```

1730      1735      1740
Leu Met Cys His Ala Thr Phe Thr Met Arg Leu Leu Ser Pro Val Arg
1745      1750      1755      1760
Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu Ala His Phe Thr Asp
1765      1770      1775
Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg Val Glu Met
1780      1785      1790
Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr Pro Pro Gly Ser Arg
1795      1800      1805
Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu
1810      1815      1820
Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe
1825      1830      1835      1840
Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp
1845      1850      1855
Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser
1860      1865      1870
Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp
1875      1880      1885
Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1890      1895      1900
Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu
1905      1910      1915      1920
Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr
1925      1930      1935
His Ser Ser Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro Lys
1940      1945      1950
Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu Asn Asp
1955      1960      1965
Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu Asp Asn Ile
1970      1975      1980
Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu Pro Glu Arg Glu
1985      1990      1995      2000
Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu Arg Gly Glu Ala Arg
2005      2010      2015
Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp Leu Pro Val Trp Leu
2020      2025      2030
Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr Ala Asp Arg Arg Trp
2035      2040      2045
Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu Glu Glu Asn Val Glu
2050      2055      2060
Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg
2065      2070      2075      2080
Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe
2085      2090      2095
Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr
2100      2105      2110
Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala
2115      2120      2125
Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2130      2135      2140
Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu
2145      2150      2155      2160
Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2165      2170      2175
Met Ser Gly Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys Ile
2180      2185      2190
Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro His Trp
2195      2200      2205
Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val Leu Leu Ile
2210      2215      2220
Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn Gln Leu Thr Tyr
2225      2230      2235      2240
Val Val Ile Ala Ile Leu Thr Val Val Ala Ala Thr Met Ala Asn Glu
2245      2250      2255
Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu Gly Leu Gly Ser Ile
2260      2265      2270
Ala Thr Gln Pro Glu Ser Asn Ile Leu Asp Ile Asp Leu Arg Pro
2275      2280      2285

```

Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr Thr Phe Val Thr Pro
 2290 2295 2300
 Met Leu Arg His Ser Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr
 2305 2310 2315 2320
 Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp
 2325 2330 2335
 Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys
 2340 2345 2350
 Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu
 2355 2360 2365
 Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
 2370 2375 2380
 Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr
 2385 2390 2395 2400
 Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro
 2405 2410 2415
 Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys Val
 2420 2425 2430
 Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu Ala Leu
 2435 2440 2445
 Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly Asn Pro Gly
 2450 2455 2460
 Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala Asn Ile Phe Arg
 2465 2470 2475 2480
 Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe Ser Ile Met Lys Asn
 2485 2490 2495
 Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile Gly Glu Thr Leu Gly
 2500 2505 2510
 Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly Lys Ser Glu Phe Gln
 2515 2520 2525
 Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp Arg Thr Leu Ala Lys
 2530 2535 2540
 Glu Gly Ile Lys Arg Gly Glu Thr Asp His His Ala Val Ser Arg Gly
 2545 2550 2555 2560
 Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu
 2565 2570 2575 2580
 Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr
 2585 2590 2595
 Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly
 2600 2605
 Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
 2610 2615 2620
 Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu
 2625 2630 2635 2640
 Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro
 2645 2650 2655
 Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu Asn
 2660 2665 2670
 Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn Pro Tyr
 2675 2680 2685
 Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg Lys Tyr Gly
 2690 2695 2700
 Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser Thr His Glu Met
 2705 2710 2715 2720
 Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val Ser Ser Val Asn Met
 2725 2730 2735
 Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met Arg Tyr Lys Lys Ala
 2740 2745 2750
 Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly Thr Arg Asn Ile Gly
 2755 2760 2765
 Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu
 2770 2775 2780
 Lys Ile Lys Gln Glu His Glu Thr Ser Trp His Tyr Asp Gln Asp His
 2785 2790 2795 2800
 Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr
 2805 2810 2815
 Gly Ser Ala Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys
 2820 2825 2830
 Pro Trp Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr

2835 2840 2845
 Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
 2850 2855 2860
 Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala
 2865 2870 2875 2880
 Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys
 2885 2890 2895
 Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Ala Leu Gly
 2900 2905 2910
 Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu Ala Val
 2915 2920 2925
 Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg Asn Leu His
 2930 2935 2940
 Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met Met Gly Lys Arg
 2945 2950 2955 2960
 Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys Gly Ser Arg Ala Ile
 2965 2970 2975
 Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu Phe Glu Ala Leu Gly
 2980 2985 2990
 Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu Asn Ser Leu Ser Gly
 2995 3000 3005
 Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr Ile Leu Arg Asp Val
 3010 3015 3020
 Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp
 3025 3030 3035 3040
 Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr
 3045 3050 3055
 Asn His Met Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys
 3060 3065 3070
 Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg
 3075 3080 3085
 Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
 3090 3095 3100
 Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln
 3105 3110 3115 3120
 Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His
 3125 3130 3135
 Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg Val
 3140 3145 3150
 Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp Cys Val
 3155 3160 3165
 Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr Ala Leu Asn
 3170 3175 3180
 Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg
 3185 3190 3195 3200
 Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys Ser His His Phe His
 3205 3210 3215
 Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn
 3220 3225 3230
 Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp
 3235 3240 3245
 Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp
 3250 3255 3260
 Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala
 3265 3270 3275 3280
 Ile Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr
 3285 3290 3295
 Trp Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu
 3300 3305 3310
 Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp
 3315 3320 3325
 Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
 3330 3335 3340
 Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr
 3345 3350 3355 3360
 Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
 3365 3370 3375
 Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg
 3380 3385 3390

Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3395 3400

<210> SEQ ID NO 27

<211> COMPRIMENTO: 10723

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<220> CARACTERÍSTICA:

<221> NOME/CHAVE: CDS

<222> LOCALIZAÇÃO: (97)...(10272)

<400> SEQUÊNCIA 27

```

agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
gttctaacag tttttaatt agagagcaga tctctg atg aat aac caa cgg aaa 114
Met Asn Asn Gln Arg Lys
1
aag gcg aaa aac acg cct ttc aat atg ctg aaa cgc gag aga aac cgc 162
Lys Ala Lys Asn Thr Pro Phe Asn Met Leu Lys Arg Glu Arg Asn Arg
10
gtg tcg act gtg caa cag ctg aca aag aga ttc tca ctt gga atg ctg 210
Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu
25
cag gga cga gga cca tta aaa ctg ttc atg gcc ctg gtg gcg ttc ctt 258
Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu
40
cgt ttc cta aca atc cca cca aca gca ggg ata ttg aag aga tgg gga 306
Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly Ile Leu Lys Arg Trp Gly
55
aca att aaa aaa tca aaa gct att aat gtt ttg aga ggg ttc agg aaa 354
Thr Ile Lys Lys Ser Lys Ala Ile Asn Val Leu Arg Gly Phe Arg Lys
75
gag att gga agg atg ctg aac atc ttg aat agg aga cgc aga tct gca 402
Glu Ile Gly Arg Met Leu Asn Ile Leu Asn Arg Arg Arg Ser Ala
90
ggc atg atc att atg ctg att cca aca gtg atg gcg ttc cat tta acc 450
Gly Met Ile Ile Met Leu Ile Pro Thr Val Met Ala Phe His Leu Thr
105
acg cgt ggg gga gag ccg cat atg ata gtt agc aag cag gaa aga gga 498
Thr Arg Gly Gly Glu Pro His Met Ile Val Ser Lys Gln Glu Arg Gly
120
aag tca ctt ttg ttc aag acc tct gca ggt gtc aac atg tgc acc ctc 546
Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly Val Asn Met Cys Thr Leu
135
att gcg atg gat ttg gga gag ttg tgt gag gac acg atg acc tac aaa 594
Ile Ala Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Met Thr Tyr Lys
155
tgc ccc cgg atc act gag gcg gaa cca gat gac gtt gac tgt tgg tgc 642
Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp Asp Val Asp Cys Trp Cys
165

```

aat	gcc	acg	gac	aca	tgg	gtg	acc	tat	gga	acg	tgc	tct	caa	act	ggc	690
Asn	Ala	Thr	Asp	Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Ser	Gln	Thr	Gly	
		185					190					195				
gaa	cac	cga	cga	gac	aaa	cgt	tcc	gtc	gca	ttg	gcc	cca	cac	gtg	ggg	738
Glu	His	Arg	Arg	Asp	Lys	Arg	Ser	Val	Ala	Leu	Ala	Pro	His	Val	Gly	
		200				205					210					
ctt	ggc	cta	gaa	aca	aga	gcc	gaa	acg	tgg	atg	tcc	tct	gaa	ggt	gct	786
Leu	Gly	Leu	Glu	Thr	Arg	Ala	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	
		215			220					225					230	
tgg	aaa	cag	ata	caa	aaa	gta	gag	act	tgg	gct	ctg	aga	cat	cca	gga	834
Trp	Lys	Gln	Ile	Gln	Lys	Val	Glu	Thr	Trp	Ala	Leu	Arg	His	Pro	Gly	
				235					240					245		
ttc	acg	gtg	ata	gcc	ctt	ttt	cta	gca	cat	gcc	ata	gga	aca	tcc	atc	882
Phe	Thr	Val	Ile	Ala	Leu	Phe	Leu	Ala	His	Ala	Ile	Gly	Thr	Ser	Ile	
			250				255						260			
acc	cag	aaa	ggg	atc	att	ttc	att	ttg	ctg	atg	ctg	gta	aca	cca	tct	930
Thr	Gln	Lys	Gly	Ile	Ile	Phe	Ile	Leu	Leu	Met	Leu	Val	Thr	Pro	Ser	
		265					270					275				
atg	gcc	atg	cga	tgc	gtg	gga	ata	ggc	aac	aga	gac	ttc	gtg	gaa	gga	978
Met	Ala	Met	Arg	Cys	Val	Gly	Ile	Gly	Asn	Arg	Asp	Phe	Val	Glu	Gly	
		280				285					290					
ctg	tca	gga	gca	aca	tgg	gtg	gat	gtg	gta	ctg	gag	cat	gga	agt	tgc	1026
Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val	Leu	Glu	His	Gly	Ser	Cys	
		295			300				305						310	
gtc	acc	acc	atg	gca	aaa	aac	aaa	cca	aca	ctg	gac	att	gaa	ctc	ttg	1074
Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Ile	Glu	Leu	Leu	
				315					320					325		
aag	acg	gag	gtc	aca	aac	cct	gca	gtt	ctg	cgt	aaa	ttg	tgc	att	gaa	1122
Lys	Thr	Glu	Val	Thr	Asn	Pro	Ala	Val	Leu	Arg	Lys	Leu	Cys	Ile	Glu	
			330				335						340			
gct	aaa	ata	tca	aac	acc	acc	acc	gat	tgc	aga	tgt	cca	aca	caa	gga	1170
Ala	Lys	Ile	Ser	Asn	Thr	Thr	Thr	Asp	Ser	Arg	Cys	Pro	Thr	Gln	Gly	
		345					350					355				
gaa	gcc	aca	ctg	gtg	gaa	gaa	caa	gac	gcg	aac	ttt	gtg	tgc	cga	cga	1218
Glu	Ala	Thr	Leu	Val	Glu	Gln	Gln	Asp	Ala	Asn	Phe	Val	Cys	Arg	Arg	
		360			365						370					
acg	ttc	gtg	gac	aga	ggc	tgg	ggc	aat	ggc	tgt	ggg	cta	ttc	gga	aaa	1266
Thr	Phe	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	
		375			380					385					390	
ggg	agt	cta	ata	acg	tgt	gcc	aag	ttt	aag	tgt	gtg	aca	aaa	cta	gaa	1314
Gly	Ser	Leu	Ile	Thr	Cys	Ala	Lys	Phe	Lys	Cys	Val	Thr	Lys	Leu	Glu	
			395						400					405		
gga	aag	ata	gtt	caa	tat	gaa	aac	cta	aaa	tat	tca	gtg	ata	gtc	acc	1362
Gly	Lys	Ile	Val	Gln	Tyr	Glu	Asn	Leu	Lys	Tyr	Ser	Val	Ile	Val	Thr	
		410					415						420			
gtc	cac	act	gga	gat	cag	cac	cag	gtg	gga	aat	gag	act	aca	gaa	cat	1410
Val	His	Thr	Gly	Asp	Gln	His	Gln	Val	Gly	Asn	Glu	Thr	Thr	Glu	His	
		425					430					435				
gga	aca	act	gca	acc	ata	aca	cct	caa	gct	cct	acg	tgc	gaa	ata	cag	1458
Gly	Thr	Thr	Ala	Thr	Ile	Thr	Pro	Gln	Ala	Pro	Thr	Ser	Glu	Ile	Gln	
		440			445					450						
ctg	acc	gac	tac	gga	acc	ctt	aca	tta	gat	tgt	tca	cct	agg	aca	ggg	1506
Leu	Thr	Asp	Tyr	Gly	Leu	Leu	Thr	Leu	Asp	Cys	Ser	Pro	Arg	Thr	Gly	
		455			460				465						470	
cta	gat	ttt	aac	gag	atg	gtg	ttg	ctg	aca	atg	aaa	gaa	aga	tca	tgg	1554
Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Thr	Met	Lys	Glu	Arg	Ser	Trp	
				475					480					485		
ctt	gtc	cac	aaa	caa	tgg	ttc	cta	gac	tta	cca	ctg	cct	tgg	acc	tct	1602
Leu	Val	His	Lys	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Thr	Thr	Ser	
			490					495					500			
ggg	gct	tca	aca	tcc	caa	gag	act	tgg	aac	aga	caa	gat	tta	ctg	gtc	1650
Gly	Ala	Ser	Thr	Ser	Gln	Glu	Thr	Trp	Asn	Arg	Gln	Asp	Leu	Leu	Val	
		505					510					515				
aca	ttt	aag	aca	gct	cat	gca	aag	aag	cag	gaa	gta	gtc	gta	cta	gga	1698
Thr	Phe	Lys	Thr	Ala	His	Ala	Lys	Lys	Gln	Glu	Val	Val	Val	Leu	Gly	
		520				525					530					
tca	caa	gaa	gga	gca	atg	cac	act	gcg	ctg	act	gga	gcg	aca	gaa	atc	1746
Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile	

535	caa	acg	tca	gga	acg	aca	aca	att	ttc	gca	gga	cac	cta	aaa	tgc	aga	1794
	Gln	Thr	Ser	Gly	Thr	Thr	Thr	Ile	Phe	Ala	Gly	His	Leu	Lys	Cys	Arg	
				555						560					565		
	cta	aaa	atg	gac	aaa	cta	act	tta	aaa	ggg	atg	tca	tat	gtg	atg	tgc	1842
	Leu	Lys	Met	Asp	Lys	Leu	Thr	Leu	Lys	Gly	Met	Ser	Tyr	Val	Met	Cys	
				570					575					580			
	aca	ggc	tca	ttc	aag	tta	gag	aaa	gaa	gtg	gct	gag	acc	cag	cat	gga	1890
	Thr	Gly	Ser	Phe	Lys	Leu	Glu	Lys	Glu	Val	Ala	Glu	Thr	Gln	His	Gly	
			585					590				595					
	act	gtt	ctg	gtg	cag	gtt	aaa	tat	gaa	gga	aca	gac	gca	cca	tgc	aag	1938
	Thr	Val	Leu	Val	Gln	Val	Lys	Tyr	Glu	Gly	Thr	Asp	Ala	Pro	Cys	Lys	
		600					605					610					
	att	ccc	ttt	tcg	acc	caa	gat	gag	aaa	gga	gca	acc	cag	aat	ggg	aga	1986
	Ile	Pro	Phe	Ser	Thr	Gln	Asp	Glu	Lys	Gly	Ala	Thr	Gln	Asn	Gly	Arg	
		615				620					625				630		
	tta	ata	aca	gcc	aac	ccc	ata	gtc	act	gac	aaa	gaa	aaa	cca	gtc	aat	2034
	Leu	Ile	Thr	Ala	Asn	Pro	Ile	Val	Thr	Asp	Lys	Glu	Lys	Pro	Val	Asn	
				635						640					645		
	att	gag	gca	gaa	cca	ccc	ttt	ggt	gag	agc	tac	atc	gtg	gta	gga	gca	2082
	Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Glu	Ser	Tyr	Ile	Val	Val	Gly	Ala	
				650				655					660				
	ggt	gaa	aaa	gct	ttg	aaa	cta	agc	tgg	ttc	aag	aaa	gga	agc	ata		2130
	Gly	Glu	Lys	Ala	Leu	Lys	Leu	Ser	Trp	Phe	Lys	Lys	Gly	Ser	Ile		
			665					670					675				
	ggg	aaa	atg	ttt	gaa	gca	act	gcc	cga	gga	gca	cga	agg	atg	gcc	att	2178
	Gly	Lys	Met	Phe	Glu	Ala	Thr	Ala	Arg	Gly	Ala	Arg	Arg	Met	Ala	Ile	
			680				685					690					
	ctg	gga	gac	acc	gca	tgg	gac	ttc	ggt	tct	ata	gga	gga	gtg	ttc	acg	2226
	Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Ile	Gly	Gly	Val	Phe	Thr	
						700					705				710		
	tct	atg	gga	aaa	ctg	gta	cac	cag	gtt	ttt	gga	act	gca	tat	gga	gtt	2274
	Ser	Met	Gly	Lys	Leu	Val	His	Gln	Val	Phe	Gly	Thr	Ala	Tyr	Gly	Val	
					715					720					725		
	ttg	ttt	agc	gga	gtt	tct	tgg	acc	atg	aaa	ata	gga	ata	ggg	att	ctg	2322
	Leu	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys	Ile	Gly	Ile	Gly	Ile	Leu	
				730					735					740			
	ctg	aca	tgg	cta	gga	tta	aat	tca	agg	aac	acg	tcc	ctt	tcg	atg	atg	2370
	Leu	Thr	Trp	Leu	Gly	Leu	Asn	Ser	Arg	Asn	Thr	Ser	Leu	Ser	Met	Met	
			745					750					755				
	tgc	atc	gca	gcc	ggc	att	gtg	aca	ctg	tat	ttg	gga	gtc	atg	gtg	cag	2418
	Cys	Ile	Ala	Ala	Gly	Ile	Val	Thr	Leu	Tyr	Leu	Gly	Val	Met	Val	Gln	
			760				765					770					
	gcc	gat	agt	ggt	tgc	gtt	gtg	agc	tgg	aaa	aac	aaa	gaa	ctg	aaa	tgt	2466
	Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys	Asn	Lys	Glu	Leu	Lys	Cys	
				775		780					785				790		
	ggc	agt	ggg	att	ttc	aca	gac	aac	gtg	cac	aca	tgg	aca	gaa	caa		2514
	Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val	His	Thr	Trp	Thr	Glu	Gln	
				795					800					805			
	tac	aag	ttc	caa	cca	gaa	tcc	cct	tca	aaa	cta	gct	tca	gct	atc	cag	2562
	Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys	Leu	Ala	Ser	Ala	Ile	Gln	
				810					815					820			
	aaa	gcc	cat	gaa	gag	gac	att	tgt	gga	atc	cgc	tca	gta	aca	aga	ctg	2610
	Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile	Arg	Ser	Val	Thr	Arg	Leu	
				825				830					835				
	gag	aat	ctg	atg	tgg	aaa	caa	ata	aca	cca	gaa	ttg	aat	cac	att	cta	2658
	Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro	Glu	Leu	Asn	His	Ile	Leu	
				840			845					850					
	tca	gaa	aat	gag	gtg	aag	tta	act	att	atg	aca	gga	gac	atc	aaa	gga	2706
	Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met	Thr	Gly	Asp	Ile	Lys	Gly	
				855		860					865				870		
	atc	atg	cag	gca	gga	aaa	cga	tct	ctg	cgg	cct	cag	ccc	act	gag	ctg	2754
	Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg	Pro	Gln	Pro	Thr	Glu	Leu	
					875					880				885			
	aag	tat	tca	tgg	aaa	aca	tgg	ggc	aaa	gca	aaa	atg	ctc	tct	aca	gag	2802
	Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala	Lys	Met	Leu	Ser	Thr	Glu	
				890					895					900			
	tct	cat	aac	cag	acc	ttt	ctc	att	gat	ggc	ccc	gaa	aca	gca	gaa	tgc	2850
	Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly	Pro	Glu	Thr	Ala	Glu	Cys	

ccc	aac	aca	aat	aga	gct	tgg	aat	tgc	ttg	gaa	ggt	gaa	gac	tat	ggc	2898
Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu	Glu	Val	Glu	Asp	Tyr	Gly	
920						925					930					
ttt	gga	gta	ttc	acc	acc	aat	ata	tgg	cta	aaa	ttg	aaa	gaa	aaa	cag	2946
Phe	Gly	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu	Lys	Leu	Lys	Glu	Lys	Gln	
935					940					945					950	
gat	gta	ttc	tgc	gac	tca	aaa	ctc	atg	tca	gcg	gcc	ata	aaa	gac	aac	2994
Asp	Val	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser	Ala	Ala	Ile	Lys	Asp	Asn	
				955					960					965		
aga	gcc	gtc	cat	gcc	gat	atg	ggt	tat	tgg	ata	gaa	agt	gca	ctc	aat	3042
Arg	Ala	Val	His	Ala	Asp	Met	Gly	Tyr	Trp	Ile	Glu	Ser	Ala	Leu	Asn	
			970					975					980			
gac	aca	tgg	aag	ata	gag	aaa	gcc	tct	ttc	att	gaa	ggt	aaa	aac	tgc	3090
Asp	Thr	Trp	Lys	Ile	Glu	Lys	Ala	Ser	Phe	Ile	Glu	Val	Lys	Asn	Cys	
		985					990					995				
cac	tgg	cca	aaa	tca	cac	acc	ctc	tgg	agc	aat	gga	gtg	cta	gaa	agt	3138
His	Trp	Pro	Lys	Ser	His	Leu	Trp	Ser	Asn	Gly	Val	Leu	Glu	Ser		
1000					1005					1010						
gag	atg	ata	att	cca	aag	aat	ctc	gct	gga	cca	gtg	tct	caa	cac	aac	3186
Glu	Met	Ile	Ile	Pro	Lys	Asn	Leu	Ala	Gly	Pro	Val	Ser	Gln	His	Asn	
1015					1020					1025					1030	
tat	aga	cca	ggc	tac	cat	aca	caa	ata	aca	gga	cca	tgg	cat	cta	ggt	3234
Tyr	Arg	Pro	Gly	Tyr	His	Thr	Gln	Ile	Thr	Gly	Pro	Trp	His	Leu	Gly	
			1035					1040						1045		
aag	ctt	gag	atg	gac	ttt	gat	ttc	tgt	gat	gga	aca	aca	gtg	gta	gtg	3282
Lys	Leu	Glu	Met	Asp	Phe	Asp	Phe	Cys	Asp	Gly	Thr	Thr	Val	Val	Val	
			1050					1055					1060			
act	gag	gac	tgc	gga	aat	aga	gga	ccc	tct	ttg	aga	aca	acc	act	gcc	3330
Thr	Glu	Asp	Cys	Gly	Asn	Arg	Gly	Pro	Ser	Leu	Arg	Thr	Thr	Thr	Ala	
		1065					1070					1075				
tct	gga	aaa	ctc	ata	aca	gaa	tgg	tgc	tgc	cga	tct	tgc	aca	tta	cca	3378
Ser	Gly	Lys	Leu	Ile	Thr	Glu	Trp	Cys	Cys	Arg	Ser	Cys	Thr	Leu	Pro	
1080					1085					1090						
ccg	cta	aga	tac	aga	ggt	gag	gat	ggg	tgc	tgg	tac	ggg	atg	gaa	atc	3426
Pro	Leu	Arg	Tyr	Arg	Gly	Glu	Asp	Gly	Cys	Trp	Tyr	Gly	Met	Glu	Ile	
1095					1100					1105					1110	
aga	cca	ttg	aag	gag	aaa	gaa	gag	aat	ttg	gtc	aac	tcc	ttg	gtc	aca	3474
Arg	Pro	Leu	Lys	Glu	Lys	Glu	Glu	Asn	Leu	Val	Asn	Ser	Leu	Val	Thr	
			1115						1120					1125		
gct	gga	cat	ggg	cag	gtc	gac	aac	ttt	tca	cta	gga	gtc	ttg	gga	atg	3522
Ala	Gly	His	Gly	Gln	Val	Asp	Asn	Phe	Ser	Leu	Gly	Val	Leu	Gly	Met	
			1130					1135					1140			
gca	ttg	ttc	ctg	gag	gaa	atg	ctt	agg	acc	cga	gta	gga	acg	aaa	cat	3570
Ala	Leu	Phe	Leu	Glu	Glu	Met	Leu	Arg	Thr	Arg	Val	Gly	Thr	Lys	His	
		1145					1150					1155				
gca	ata	cta	cta	gtt	gca	gtt	tct	ttt	gtg	aca	ttg	atc	aca	ggg	aac	3618
Ala	Ile	Leu	Leu	Val	Ala	Val	Ser	Phe	Val	Thr	Leu	Ile	Thr	Gly	Asn	
		1160			1165					1170						
atg	tcc	ttt	aga	gac	ctg	gga	aga	gtg	atg	gtt	atg	gta	ggc	gcc	act	3666
Met	Ser	Phe	Arg	Asp	Leu	Gly	Arg	Val	Met	Val	Met	Val	Gly	Ala	Thr	
1175					1180					1185					1190	
atg	acg	gat	gac	ata	ggt	atg	ggc	gtg	act	tat	ctt	gcc	cta	cta	gca	3714
Met	Thr	Asp	Asp	Ile	Gly	Met	Gly	Val	Thr	Tyr	Leu	Ala	Leu	Leu	Ala	
			1195					1200					1205			
gcc	ttc	aaa	gtc	aga	cca	act	ttt	gca	gct	gga	cta	ctc	ttg	aga	aag	3762
Ala	Phe	Lys	Val	Arg	Pro	Thr	Phe	Ala	Ala	Gly	Leu	Leu	Leu	Arg	Lys	
			1210					1215					1220			
ctg	acc	tcc	aag	gaa	ttg	atg	atg	act	act	ata	gga	att	gta	ctc	ctc	3810
Leu	Thr	Ser	Lys	Glu	Leu	Met	Met	Thr	Thr	Ile	Gly	Ile	Val	Leu	Leu	
		1225					1230					1235				
tcc	cag	agc	acc	ata	cca	gag	acc	att	ctt	gag	ttg	act	gat	gcg	tta	3858
Ser	Gln	Ser	Thr	Ile	Pro	Glu	Thr	Ile	Leu	Glu	Leu	Thr	Asp	Ala	Leu	
		1240				1245						1250				
gcc	tta	ggc	atg	atg	gtc	ctc	aaa	atg	gtg	aga	aat	atg	gaa	aag	tat	3906
Ala	Leu	Gly	Met	Met	Val	Leu	Lys	Met	Val	Arg	Asn	Met	Glu	Lys	Tyr	
1255					1260					1265					1270	
caa	ttg	gca	gtg	act	atc	atg	gct	atc	ttg	tgc	gtc	cca	aac	gca	gtg	3954
Gln	Leu	Ala	Val	Thr	Ile	Met	Ala	Ile	Leu	Cys	Val	Pro	Asn	Ala	Val	

ata	tta	caa	aac	gca	tgg	aaa	gtg	agt	tgc	aca	ata	ttg	gca	gtg	gtg	4002
Ile	Leu	Gln	Asn	Ala	Trp	Lys	Val	Ser	Cys	Thr	Ile	Leu	Ala	Val	Val	
			1290					1295					1300			
tcc	gtt	tcc	cca	ctg	ttc	tta	aca	tcc	tca	cag	caa	aaa	aca	gat	tgg	4050
Ser	Val	Ser	Pro	Leu	Phe	Leu	Thr	Ser	Ser	Gln	Gln	Lys	Thr	Asp	Trp	
		1305					1310					1315				
ata	cca	tta	gca	ttg	acg	atc	aaa	ggt	ctc	aat	cca	aca	gct	att	ttt	4098
Ile	Pro	Leu	Ala	Leu	Thr	Ile	Lys	Gly	Leu	Asn	Pro	Thr	Ala	Ile	Phe	
	1320					1325					1330					
cta	aca	acc	ctc	tca	aga	acc	agc	aag	aaa	agg	agc	tgg	cca	tta	aat	4146
Leu	Thr	Thr	Leu	Ser	Arg	Thr	Ser	Lys	Lys	Arg	Ser	Trp	Pro	Leu	Asn	
	1335				1340					1345					1350	
gag	gct	atc	atg	gca	gtc	ggg	atg	gtg	agc	att	tta	gcc	agt	tct	ctc	4194
Glu	Ala	Ile	Met	Ala	Val	Gly	Met	Val	Ser	Ile	Leu	Ala	Ser	Ser	Leu	
				1355					1360						1365	
cta	aaa	aat	gat	att	ccc	atg	aca	gga	cca	tta	gtg	gct	gga	ggg	ctc	4242
Leu	Lys	Asn	Asp	Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	
			1370					1375					1380			
ctc	act	gtg	tgc	tac	gtg	ctc	act	gga	cga	tcg	gcc	gat	ttg	gaa	ctg	4290
Leu	Thr	Val	Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu	
		1385					1390					1395				
gag	aga	gca	gcc	gat	gtc	aaa	tgg	gaa	gac	cag	gca	gag	ata	tca	gga	4338
Glu	Arg	Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	
	1400					1405					1410					
agc	agt	cca	atc	ctg	tca	ata	aca	ata	tca	gaa	gat	ggt	agc	atg	tcg	4386
Ser	Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser	
	1415				1420				1425						1430	
ata	aaa	aat	gaa	gag	gaa	gaa	caa	aca	ctg	acc	ata	ctc	att	aga	aca	4434
Ile	Lys	Asn	Glu	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg	Thr	
			1435						1440					1445		
gga	ttg	ctg	gtg	atc	tca	gga	ctt	ttt	cct	gta	tca	ata	cca	atc	acg	4482
Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro	Ile	Thr	
			1450				1455						1460			
gca	gca	gca	tgg	tac	ctg	tgg	gaa	gtg	aag	aaa	caa	cgg	gcc	gga	gta	4530
Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg	Ala	Gly	Val	
		1465				1470						1475				
ttg	tgg	gat	gtt	cct	tca	ccc	cca	ccc	atg	gga	aag	gct	gaa	ctg	gaa	4578
Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys	Ala	Glu	Leu	Glu	
	1480					1485					1490					
gat	gga	gcc	tat	aga	att	aag	caa	aaa	ggg	att	ctt	gga	tat	tcc	cag	4626
Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile	Leu	Gly	Tyr	Ser	Gln	
	1495				1500				1505						1510	
atc	gga	gcc	gga	gtt	tac	aaa	gaa	gga	aca	ttc	cat	aca	atg	tgg	cat	4674
Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr	Phe	His	Thr	Met	Trp	His	
			1515						1520					1525		
gtc	aca	cgt	ggc	gct	gtt	cta	atg	cat	aaa	gga	aag	agg	att	gaa	cca	4722
Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	
			1530				1535						1540			
tca	tgg	gcg	gac	gtc	aag	aaa	gac	cta	ata	tca	tat	gga	gga	ggc	tgg	4770
Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	
		1545					1550					1555				
aag	tta	gaa	gga	gaa	tgg	aag	gaa	gga	gaa	gaa	gtc	cag	gta	ttg	gca	4818
Lys	Leu	Glu	Gly	Glu	Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	
	1560					1565					1570					
ctg	gag	cct	gga	aaa	aat	cca	aga	gcc	gtc	caa	acg	aaa	cct	ggt	ctt	4866
Leu	Glu	Pro	Gly	Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	
	1575				1580					1585					1590	
ttc	aaa	acc	aac	gcc	gga	aca	ata	ggt	gct	gta	tct	ctg	gac	ttt	tct	4914
Phe	Lys	Thr	Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	
			1595						1600					1605		
cct	gga	acg	tca	gga	tct	cca	att	atc	gac	aaa	aaa	gga	aaa	gtt	gtg	4962
Pro	Gly	Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	
			1610				1615						1620			
ggt	ctt	tat	ggt	aat	ggt	gtt	gtt	aca	agg	agt	gga	gca	tat	gtg	agt	5010
Gly	Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser	
		1625				1630						1635				
gct	ata	gcc	cag	act	gaa	aaa	agc	att	gaa	gac	aac	cca	gag	atc	gaa	5058
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile	Glu	

1640	gat gac att ttc cga aag aga aga ctg acc atc atg gac ctc cac cca	1645	1650	5106
Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro	1655	1660	1665	1670
gga gcg gga aag acg aag aga tac ctt ccg gcc ata gtc aga gaa gct	5154	1675	1680	1685
Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu Ala	5202	1690	1695	1700
ata aaa cgg ggt ttg aga aca tta atc ttg gcc ccc act aga gtt gtg	5250	1705	1710	1715
Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg Val Val	5298	1720	1725	1730
gca gct gaa atg gag gaa gcc ctt aga gga ctt cca ata aga tac cag	5346	1735	1740	1745
Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile Arg Tyr Gln	5394	1750	1755	1760
acc cca gcc atc aga gct gtg cac acc ggg cgg gag att gtg gac cta	5442	1765	1770	1775
Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu Ile Val Asp Leu	5490	1775	1780	1785
atg tgt cat gcc aca ttt acc atg agg ctg cta tca cca gtt aga gtg	5538	1790	1795	1800
Met Cys His Ala Thr Phe Thr Met Arg Leu Leu Ser Pro Val Arg Val	5586	1805	1810	1815
cca aac tac aac ctg att atc atg gac gaa gcc cat ttc aca gac cca	5634	1820	1825	1830
Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu Ala His Phe Thr Asp Pro	5682	1835	1840	1845
gca agt ata gca gct aga gga tac atc tca act cga gtg gag atg ggt	5730	1850	1855	1860
Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Arg Val Glu Met Gly	5778	1865	1870	1875
gag gca gct ggg att ttt atg aca gcc act ccc ccg gga agc aga gac	5826	1880	1885	1890
Glu Ala Ala Gly Ile Phe Met Thr Ala Thr Pro Pro Gly Ser Arg Asp	5874	1895	1900	1905
cca ttt cct cag agc aat gca cca atc ata gat gaa gaa aga gaa atc	5922	1910	1915	1920
Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile	5970	1925	1930	1935
cct gaa cgc tgc tgg aat tcc gga cat gaa tgg gtc acg gat ttt aaa	6018	1940	1945	1950
Pro Glu Arg Ser Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys	6066	1955	1960	1965
ggg aag act gtt tgg ttc gtt cca agt ata aaa gca gga aat gat ata	6114	1970	1975	1980
Gly Lys Thr Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile	6162	1985	1990	1995
gca gct tgc ctg agg aaa aat gga aag aaa gtg ata caa ctc agt agg		2000	2005	
Ala Ala Cys Leu Arg Lys Asn Gly Lys Val Ile Gln Leu Ser Arg				
aag acc ttt gat tct gag tat gtc aag act aga acc aat gat tgg gac				
Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp				
ttc gtg gtt aca act gac att tca gaa atg ggt gcc aat ttc aag gct				
Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala				
gag agg gtt ata gac ccc aga cgc tgc atg aaa cca gtc ata cta aca				
Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu Thr				
gat ggt gaa gag cgg gtg att ctg gca gga cct atg cca gtg acc cac				
Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr His				
tct agt gca gca caa aga aga ggg aga ata gga aga aat cca aaa aat				
Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro Lys Asn				
gag aat gac cag tac ata tac atg ggg gaa cct ctg gaa aat gat gaa				
Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu Asn Asp Glu				
gac tgt gca cac tgg aaa gaa gct aaa atg ctc cta gat aac atc aac				
Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu Asp Asn Ile Asn				
acg cca gaa gga atc att cct agc atg ttc gaa cca gag cgt gaa aag				
Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu Pro Glu Arg Glu Lys				
gtg gat gcc att gat ggc gaa tac cgc ttg aga gga gaa gca agg aaa				
Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu Arg Gly Glu Ala Arg Lys				
acc ttt gta gac tta atg aga aga gga gac cta cca gtc tgg tgg gcc				
Thr Phe Val Asp Leu Met Arg Arg Gly Asp Leu Pro Val Trp Leu Ala				

2375	gat gga ata aca gtg att gac cta gat cca ata cct tat gat cca aag	2380	2385	2390	7314
Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro Lys	2395	2400	2405	2410	
ttt gaa aag cag ttg gga caa gta atg ctc cta gtc ctc tgc gtg act	2410	2415	2420	2425	7362
Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys Val Thr	2425	2430	2435	2440	
caa gta ttg atg atg agg act aca ttg gct ctg tgt gag gct tta acc	2440	2445	2450	2455	7410
Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu Ala Leu Thr	2455	2460	2465	2470	
tta gct acc ggg ccc atc tcc aca ttg ttg gaa gga aat cca ggg agg	2460	2465	2470	2475	7458
Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly Asn Pro Gly Arg	2475	2480	2485	2490	
ttt tgg aac act acc att gcg gtg tca atg gct aac att ttt aga ggg	2480	2485	2490	2495	7506
Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala Asn Ile Phe Arg Gly	2495	2500	2505	2510	
agt tac ttg gcc gga gct gga ctt ctc ttt tct att atg aag aac aca	2505	2510	2515	2520	7554
Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe Ser Ile Met Lys Asn Thr	2520	2525	2530	2535	
acc aac aca aga agg gga act ggc aac ata gga gag acg ctt gga gag	2535	2540	2545	2550	7602
Thr Asn Thr Arg Arg Gly Thr Gly Ile Gly Glu Thr Leu Gly Glu	2550	2555	2560	2565	
aaa tgg aaa agc cga ttg aac gca ttg gga aaa agt gaa ttc cag atc	2565	2570	2575	2580	7650
Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile	2580	2585	2590	2595	
tac aag aaa agt gga atc cag gaa gtg gat aga acc tta gca aaa gaa	2595	2600	2605	2610	7698
Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu	2610	2615	2620	2625	
ggc att aaa aga gga gaa acg gac cat cac gct gtg tcg cga ggc tca	2625	2630	2635	2640	7746
Gly Ile Lys Arg Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser	2640	2645	2650	2655	
gca aaa ctg aga tgg ttc gtt gag aga aac atg gtc aca cca gaa ggg	2655	2660	2665	2670	7794
Ala Lys Leu Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly	2670	2675	2680	2685	
aaa gta gtg gac ctc ggt tgt ggc aga gga ggc tgg tca tac tat tgt	2685	2690	2695	2700	7842
Lys Val Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys	2700	2705	2710	2715	
gga gga cta aag aat gta aga gaa gtc aaa ggc cta aca aaa gga gga	2715	2720	2725	2730	7890
Gly Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly	2730	2735	2740	2745	
cca gga cac gaa gaa ccc atc ccc atg tca aca tat ggg tgg aat cta	2745	2750	2755	2760	7938
Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu	2760	2765	2770	2775	
gtg cgt ctt caa agt gga gtt gac gtt ttc ttc atc ccg cca gaa aag	2775	2780	2785	2790	7986
Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys	2790	2795	2800	2805	
tgt gac aca tta ttg tgt gac ata ggg gag tca tca cca aat ccc aca	2805	2810	2815	2820	8034
Cys Asp Thr Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro Thr	2820	2825	2830	2835	
gtg gaa gca gga cga aca ctc aga gtc ctt aac tta gta gaa aat tgg	2835	2840	2845	2850	8082
Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu Asn Trp	2850	2855	2860	2865	
ttg aac aac aac act caa ttt tgc ata aag gtt ctc aac cca tat atg	2865	2870	2875	2880	8130
Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn Pro Tyr Met	2880	2885	2890	2895	
ccc tca gtc ata gaa aaa atg gaa gca cta caa agg aaa tat gga gga	2895	2900	2905	2910	8178
Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg Lys Tyr Gly Gly	2910	2915	2920	2925	
gcc tta gtg agg aat cca ctc tca cga aac tcc aca cat gag atg tac	2925	2930	2935	2940	8226
Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser Thr His Glu Met Tyr	2940	2945	2950	2955	
ttg gta tcc aat gct tcc ggg aac ata gtg tca tca gtg aac atg att	2955	2960	2965	2970	8274
Trp Val Ser Asn Ala Ser Gly Asn Ile Val Ser Ser Val Asn Met Ile	2970	2975	2980	2985	
tca agg atg ttg atc aac aga ttt aca atg aga tac aag aaa gcc act	2985	2990	2995	3000	8322
Ser Arg Met Leu Ile Asn Arg Phe Thr Met Arg Tyr Lys Lys Ala Thr	3000	3005	3010	3015	
tac gag ccg gat gtt gac ctc gga agc gga acc cgt aac atc ggg att	3015	3020	3025	3030	8370
Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile	3030	3035	3040	3045	


```

          3115          3120          3125
aca atc aca gaa gaa atc gct gtg caa aac tgg tta gca aga gtg ggg 9522
Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg Val Gly
          3130          3135          3140
cgc gaa agg tta tca aga atg gcc atc agt gga gat gat tgt gtt gtg 9570
Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp Cys Val Val
          3145          3150          3155
aaa cct tta gat gac agg ttc gca agc gct tta aca gct cta aat gac 9618
Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr Ala Leu Asn Asp
          3160          3165          3170
atg gga aag att agg aaa gac ata caa caa tgg gaa cct tca aga gga 9666
Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp Glu Pro Ser Arg Gly
          3175          3180          3185
tgg aat gat tgg aca caa gtg ccc ttc tgt tca cac cat ttc cat gag 9714
Trp Asn Asp Trp Thr Gln Val Pro Phe Cys Ser His His Phe His Glu
          3195          3200          3205
tta atc atg aaa gac ggt cgc gta ctc gtt gtt cca tgt aga aac caa 9762
Leu Ile Met Lys Asp Gly Arg Val Leu Val Val Pro Cys Arg Asn Gln
          3210          3215          3220
gat gaa ctg att ggc aga gcc cga atc tcc caa gga gca ggg tgg tct 9810
Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser
          3225          3230          3235
ttg cgg gag acg gcc tgt ttg ggg aag tct tac gcc caa atg tgg agc 9858
Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser
          3240          3245          3250
ttg atg tac ttc cac aga cgc gac ctc agg ctg gcg gca aat gct att 9906
Leu Met Tyr Phe His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile
          3255          3260          3265
tgc tcg gca gta cca tca cat tgg gtt cca aca agt cga aca acc tgg 9954
Cys Ser Ala Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Trp
          3275          3280          3285
tcc ata cat gct aaa cat gaa tgg atg aca acg gaa gac atg ctg aca 10002
Ser Ile His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr
          3290          3295          3300
gtc tgg aac agg gtg tgg att caa gaa aac cca tgg atg gaa gac aaa 10050
Val Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
          3305          3310          3315
act cca gtg gaa tca tgg gag gaa atc cca tac ttg ggg aaa aga gaa 10098
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu
          3320          3325          3330
gac caa tgg tgc ggc tca ttg att ggg tta aca agc agg gcc acc tgg 10146
Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
          3335          3340          3345
gca aag aac atc caa gca gca ata aat caa gtt aga tcc ctt ata ggc 10194
Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile Gly
          3355          3360          3365
aat gaa gaa tac aca gat tac atg cca tcc atg aaa aga ttc aga aga 10242
Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe Arg Arg
          3370          3375          3380
gaa gag gaa gaa gca gga gtt ctg tgg tag aaagcaaaac taacatgaaa 10292
Glu Glu Glu Glu Ala Gly Val Leu Trp *
          3385          3390
caaggctaga agtcaggctg gattaagcca tagtacggaa aaaactatgc tacctgtgag 10352
ccccgtccaa ggacgttaaa agaagtcagg ccatacataaa tgccatagct tgagtaaaact 10412
atgcagcctg tagctccacc tgagaagggtg taaaaaatcc gggaggccac aaaccatgga 10472
agctgtacgc atggcgtagt ggactagcgg ttagaggaga cccctccctt acaaatcgca 10532
gcaacaatgg gggcccaagg cgagatgaag ctgtagtctc gctggaagga ctagagggtta 10592
gaggagacc ccccgaaaca aaaaacagca tattgacgct gggaaagacc agagatcctg 10652
ctgtctcttc agcatcattc caggcacaga acgccagaaa atggaatggt gctgttgaat 10712
caacaggttc t 10723

```

<210> SEQ ID NO 28

<211> COMPRIMENTO: 3391

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

EP2278012B1

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 28

Met	Asn	Asn	Gln	Arg	Lys	Lys	Ala	Lys	Asn	Thr	Pro	Phe	Asn	Met	Leu
1				5					10					15	
Lys	Arg	Glu	Arg	Asn	Arg	Val	Ser	Thr	Val	Gln	Gln	Leu	Thr	Lys	Arg
			20					25					30		
Phe	Ser	Leu	Gly	Met	Leu	Gln	Gly	Arg	Gly	Pro	Leu	Lys	Leu	Phe	Met
		35					40					45			
Ala	Leu	Val	Ala	Phe	Leu	Arg	Phe	Leu	Thr	Ile	Pro	Pro	Thr	Ala	Gly
	50					55					60				
Ile	Leu	Lys	Arg	Trp	Gly	Thr	Ile	Lys	Lys	Ser	Lys	Ala	Ile	Asn	Val
65					70					75				80	
Leu	Arg	Gly	Phe	Arg	Lys	Glu	Ile	Gly	Arg	Met	Leu	Asn	Ile	Leu	Asn
				85					90					95	
Arg	Arg	Arg	Arg	Ser	Ala	Gly	Met	Ile	Ile	Met	Leu	Ile	Pro	Thr	Val
			100					105					110		
Met	Ala	Phe	His	Leu	Thr	Thr	Arg	Gly	Gly	Glu	Pro	His	Met	Ile	Val
		115					120					125			
Ser	Lys	Gln	Glu	Arg	Gly	Lys	Ser	Leu	Leu	Phe	Lys	Thr	Ser	Ala	Gly
	130					135					140				
Val	Asn	Met	Cys	Thr	Leu	Ile	Ala	Met	Asp	Leu	Gly	Glu	Leu	Cys	Glu
145					150					155				160	
Asp	Thr	Met	Thr	Tyr	Lys	Cys	Pro	Arg	Ile	Thr	Glu	Ala	Glu	Pro	Asp
				165					170					175	
Asp	Val	Asp	Cys	Trp	Cys	Asn	Ala	Thr	Asp	Thr	Trp	Val	Thr	Tyr	Gly
		180						185					190		
Thr	Cys	Ser	Gln	Thr	Gly	Glu	His	Arg	Arg	Asp	Lys	Arg	Ser	Val	Ala
	195						200					205			
Leu	Ala	Pro	His	Val	Gly	Leu	Gly	Leu	Glu	Thr	Arg	Ala	Glu	Thr	Trp
	210					215					220				
Met	Ser	Ser	Glu	Gly	Ala	Trp	Lys	Gln	Ile	Gln	Lys	Val	Glu	Thr	Trp
225					230					235				240	
Ala	Leu	Arg	His	Pro	Gly	Phe	Thr	Val	Ile	Ala	Leu	Phe	Leu	Ala	His
				245					250					255	
Ala	Ile	Gly	Thr	Ser	Ile	Thr	Gln	Lys	Gly	Ile	Ile	Phe	Ile	Leu	Leu
		260						265					270		
Met	Leu	Val	Thr	Pro	Ser	Met	Ala	Met	Arg	Cys	Val	Gly	Ile	Gly	Asn
	275						280					285			
Arg	Asp	Phe	Val	Glu	Gly	Leu	Ser	Gly	Ala	Thr	Trp	Val	Asp	Val	Val
	290					295					300				
Leu	Glu	His	Gly	Ser	Cys	Val	Thr	Thr	Met	Ala	Lys	Asn	Lys	Pro	Thr
305					310					315				320	
Leu	Asp	Ile	Glu	Leu	Lys	Thr	Glu	Val	Thr	Asn	Pro	Ala	Val	Leu	
				325				330					335		
Arg	Lys	Leu	Cys	Ile	Glu	Ala	Lys	Ile	Ser	Asn	Thr	Thr	Thr	Asp	Ser
			340					345					350		
Arg	Cys	Pro	Thr	Gln	Gly	Glu	Ala	Thr	Leu	Val	Glu	Glu	Gln	Asp	Ala
	355						360					365			
Asn	Phe	Val	Cys	Arg	Arg	Thr	Phe	Val	Asp	Arg	Gly	Trp	Gly	Asn	Gly
	370					375					380				
Cys	Gly	Leu	Phe	Gly	Lys	Gly	Ser	Leu	Ile	Thr	Cys	Ala	Lys	Phe	Lys
385					390					395				400	
Cys	Val	Thr	Lys	Leu	Glu	Gly	Lys	Ile	Val	Gln	Tyr	Glu	Asn	Leu	Lys
				405					410					415	
Tyr	Ser	Val	Ile	Val	Thr	Val	His	Thr	Gly	Asp	Gln	His	Gln	Val	Gly
		420						425					430		
Asn	Glu	Thr	Thr	Glu	His	Gly	Thr	Thr	Ala	Thr	Ile	Thr	Pro	Gln	Ala
	435						440					445			
Pro	Thr	Ser	Glu	Ile	Gln	Leu	Thr	Asp	Tyr	Gly	Thr	Leu	Thr	Leu	Asp
	450					455				460					
Cys	Ser	Pro	Arg	Thr	Gly	Leu	Asp	Phe	Asn	Glu	Met	Val	Leu	Leu	Thr
465					470					475				480	
Met	Lys	Glu	Arg	Ser	Trp	Leu	Val	His	Lys	Gln	Trp	Phe	Leu	Asp	Leu
				485					490					495	
Pro	Leu	Pro	Trp	Thr	Ser	Gly	Ala	Ser	Thr	Ser	Gln	Glu	Thr	Trp	Asn
		500						505					510		
Arg	Gln	Asp	Leu	Leu	Val	Thr	Phe	Lys	Thr	Ala	His	Ala	Lys	Lys	Gln
	515						520					525			
Glu	Val	Val	Val	Leu	Gly	Ser	Gln	Glu	Gly	Ala	Met	His	Thr	Ala	Leu
	530					535					540				
Thr	Gly	Ala	Thr	Glu	Ile	Gln	Thr	Ser	Gly	Thr	Thr	Thr	Ile	Phe	Ala

545 Gly His Leu Lys Cys 550 Arg Leu Lys Met Asp 555 Lys Leu Thr Leu Lys 560 Gly
 Met Ser Tyr Val 565 Met Cys Thr Gly Ser 570 Phe Lys Leu Glu Lys 575 Glu Val
 Ala Glu Thr 580 Gln His Gly Thr Val 585 Leu Val Gln Val Lys 590 Tyr Glu Gly
 Thr Asp 595 Pro Cys Lys Ile 600 Phe Ser Thr Gln 605 Asp Glu Lys Gly
 610 Ala Thr Gln Asn Gly Arg 615 Leu Ile Thr Ala Asn 620 Pro Ile Val Thr Asp
 625 Lys Glu Lys Pro Val 630 Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
 Tyr Ile Val 645 Val Gly Ala Gly Glu Lys 650 Ala Leu Lys Leu Ser 655 Trp Phe
 Lys Lys Gly 660 Ser Ser Ile Gly Lys 665 Met Phe Glu Ala Thr 670 Ala Arg Gly
 Ala Arg 675 Met Ala Ile Leu 680 Gly Asp Thr Ala Trp 685 Asp Phe Gly Ser
 690 Ile Gly Gly Val Phe Thr 695 Ser Met Gly Lys Leu Val His Gln Val Phe
 705 Gly Thr Ala Tyr Gly 710 Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
 Ile Gly Ile Gly 725 Ile Leu Leu Thr Trp 730 Leu Gly Leu Asn Ser 735 Arg Asn
 Thr Ser Leu 740 Ser Met Met Cys Ile 745 Ala Ala Gly Ile Val Thr Leu Tyr
 Leu Gly Val 755 Met Val Gln Ala 760 Asp Ser Gly Cys Val 765 Val Ser Trp Lys
 770 Asn Lys Glu Leu Lys Cys 775 Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 785 His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 Leu Ala Ser 805 Ile Gln Lys Ala His 810 Glu Glu Asp Ile Cys 815 Gly Ile
 Arg Ser Val 820 Thr Arg Leu Glu Asn 825 Leu Met Trp Lys Gln Ile Thr Pro
 Glu Leu Asn 835 His Ile Leu Ser 840 Glu Asn Glu Val Lys 845 Leu Thr Ile Met
 850 Thr Gly Asp Ile Lys Gly 855 Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 Lys Met Leu 885 Ser Thr Glu Ser His Asn 890 Gln Thr Phe Leu Ile Asp Gly
 900 Pro Glu Thr Ala Glu Cys Pro Asn 905 Thr Asn Arg Ala Trp Asn Ser Leu
 Glu Val Glu Asp Tyr Gly Phe 920 Val Phe Thr Thr Asn Ile Trp Leu
 930 Lys Leu Lys Glu Lys Gln 935 Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 Ala Ala Ile Lys Asp Asn Arg Ala Val His 955 Ala Asp Met Gly Tyr Trp
 965 Ile Glu Ser Ala Leu Asn Asp Thr Trp 970 Lys Ile Glu Lys Ala Ser Phe
 980 Ile Glu Val Lys Asn Cys His Trp 985 Pro Lys Ser His Thr Leu Trp Ser
 995 Asn Gly Val Leu Glu Ser Glu 1000 Met Ile Ile Pro Lys Asn Leu Ala Gly
 1010 Pro Val Ser Gln His Asn 1015 Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys Asp
 1045 Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly Pro Ser
 1060 Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu Trp Cys Cys
 1075 Arg Ser Cys Thr Leu Pro Pro 1080 Leu Arg Tyr Arg Gly 1085 Glu Asp Gly Cys
 1090 1095 1100

Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu Lys Glu Glu Asn Leu
 1105 1110 1115 1120
 Val Asn Ser Leu Val Thr Ala Gly His Gly Gln Val Asp Asn Phe Ser
 1125 1130 1135
 Leu Gly Val Leu Gly Met Ala Leu Phe Leu Glu Glu Met Leu Arg Thr
 1140 1145 1150
 Arg Val Gly Thr Lys His Ala Ile Leu Leu Val Ala Val Ser Phe Val
 1155 1160 1165
 Thr Leu Ile Thr Gly Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met
 1170 1175 1180
 Val Met Val Gly Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr
 1185 1190 1195 1200
 Tyr Leu Ala Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala
 1205 1210 1215
 Gly Leu Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr
 1220 1225 1230
 Ile Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
 1235 1240 1245
 Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val
 1250 1255 1260
 Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
 1265 1270 1275 1280
 Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser Cys
 1285 1290 1295
 Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr Ser Ser
 1300 1305 1310
 Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile Lys Gly Leu
 1315 1320 1325
 Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg Thr Ser Lys Lys
 1330 1335 1340
 Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala Val Gly Met Val Ser
 1345 1350 1355 1360
 Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp Ile Pro Met Thr Gly Pro
 1365 1370 1375
 Leu Val Ala Gly Gly Leu Leu Thr Val Cys Tyr Val Leu Thr Gly Arg
 1380 1385 1390
 Ser Ala Asp Leu Glu Leu Glu Arg Ala Ala Asp Val Lys Trp Glu Asp
 1395 1400 1405
 Gln Ala Glu Ile Ser Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser
 1410 1415 1420
 Glu Asp Gly Ser Met Ser Ile Lys Asn Glu Glu Glu Gln Thr Leu
 1425 1430 1435 1440
 Thr Ile Leu Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro
 1445 1450 1455
 Val Ser Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys
 1460 1465 1470
 Lys Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Met
 1475 1480 1485
 Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly
 1490 1495 1500
 Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
 1505 1510 1515 1520
 Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His Lys
 1525 1530 1535
 Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp Leu Ile
 1540 1545 1550
 Ser Tyr Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys Glu Gly Glu
 1555 1560 1565
 Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn Pro Arg Ala Val
 1570 1575 1580
 Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala Gly Thr Ile Gly Ala
 1585 1590 1595 1600
 Val Ser Leu Asp Phe Ser Pro Gly Thr Ser Gly Ser Pro Ile Ile Asp
 1605 1610 1615
 Lys Lys Gly Lys Val Val Gly Leu Tyr Gly Asn Gly Val Val Thr Arg
 1620 1625 1630
 Ser Gly Ala Tyr Val Ser Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu
 1635 1640 1645
 Asp Asn Pro Glu Ile Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr

1650 1655 1660
 Ile Met Asp Leu His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro
 1665 1670 1675 1680
 Ala Ile Val Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu
 1685 1690 1695
 Ala Pro Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly
 1700 1705 1710
 Leu Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
 1715 1720 1725
 Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu
 1730 1735 1740
 Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
 1745 1750 1755 1760
 Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser
 1765 1770 1775
 Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr Ala Thr
 1780 1785 1790
 Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala Pro Ile Ile
 1795 1800 1805
 Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn Ser Gly His Glu
 1810 1815 1820
 Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp Phe Val Pro Ser Ile
 1825 1830 1835 1840
 Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu Arg Lys Asn Gly Lys Lys
 1845 1850 1855
 Val Ile Gln Leu Ser Arg Lys Thr Phe Asp Ser Glu Tyr Val Lys Thr
 1860 1865 1870
 Arg Thr Asn Asp Trp Asp Phe Val Val Thr Thr Asp Ile Ser Glu Met
 1875 1880 1885
 Gly Ala Asn Phe Lys Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met
 1890 1895 1900
 Lys Pro Val Ile Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly
 1905 1910 1915 1920
 Pro Met Pro Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile
 1925 1930 1935
 Gly Arg Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu
 1940 1945 1950
 Pro Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
 1955 1960 1965
 Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe
 1970 1975 1980
 Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
 1985 1990 1995 2000
 Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly Asp
 2005 2010 2015
 Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile Asn Tyr
 2020 2025 2030
 Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn Gln Ile Leu
 2035 2040 2045
 Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu Gly Glu Arg Lys
 2050 2055 2060
 Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile Tyr Ser Asp Pro Leu
 2065 2070 2075 2080
 Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala Gly Arg Lys Ser Leu Thr
 2085 2090 2095
 Leu Asn Leu Ile Thr Glu Met Gly Arg Leu Pro Thr Phe Met Thr Gln
 2100 2105 2110
 Lys Ala Arg Asp Ala Leu Asp Asn Leu Ala Val Leu His Thr Ala Glu
 2115 2120 2125
 Ala Gly Gly Arg Ala Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr
 2130 2135 2140
 Leu Glu Thr Leu Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly
 2145 2150 2155 2160
 Ile Phe Leu Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu
 2165 2170 2175
 Gly Met Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln
 2180 2185 2190
 Ile Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
 2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp
 2210 2215 2220
 Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
 2225 2230 2235 2240
 Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp Leu
 2245 2250 2255
 Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile Leu Asp
 2260 2265 2270
 Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala Val Ala Thr
 2275 2280 2285
 Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu Asn Ser Ser Val
 2290 2295 2300
 Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala Thr Val Leu Met Gly
 2305 2310 2315 2320
 Leu Gly Lys Gly Trp Pro Leu Ser Lys Met Asp Ile Gly Val Pro Leu
 2325 2330 2335
 Leu Ala Ile Gly Cys Tyr Ser Gln Val Asn Pro Ile Thr Leu Thr Ala
 2340 2345 2350
 Ala Leu Phe Leu Leu Val Ala His Tyr Ala Ile Ile Gly Pro Gly Leu
 2355 2360 2365
 Gln Ala Lys Ala Thr Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile
 2370 2375 2380
 Met Lys Asn Pro Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro
 2385 2390 2395 2400
 Ile Pro Tyr Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu
 2405 2410 2415
 Leu Val Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala
 2420 2425 2430
 Leu Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
 2435 2440 2445
 Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met
 2450 2455 2460
 Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
 2465 2470 2475 2480
 Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn Ile
 2485 2490 2495
 Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala Leu Gly
 2500 2505 2510
 Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln Glu Val Asp
 2515 2520 2525
 Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu Thr Asp His His
 2530 2535 2540
 Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp Phe Val Glu Arg Asn
 2545 2550 2555 2560
 Met Val Thr Pro Glu Gly Lys Val Val Asp Leu Gly Cys Gly Arg Gly
 2565 2570 2575
 Gly Trp Ser Tyr Tyr Cys Gly Gly Leu Lys Asn Val Arg Glu Val Lys
 2580 2585 2590
 Gly Leu Thr Lys Gly Gly Pro Gly His Glu Glu Pro Ile Pro Met Ser
 2595 2600 2605
 Thr Tyr Gly Trp Asn Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe
 2610 2615 2620
 Phe Ile Pro Pro Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu
 2625 2630 2635 2640
 Ser Ser Pro Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu
 2645 2650 2655
 Asn Leu Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys
 2660 2665 2670
 Val Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
 2675 2680 2685
 Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn
 2690 2695 2700
 Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
 2705 2710 2715 2720
 Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr Met
 2725 2730 2735
 Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly Ser Gly
 2740 2745 2750
 Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu Asp Ile Ile

2755 2760 2765
 Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu Thr Ser Trp His
 2770 2775 2780
 Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala Tyr His Gly Ser Tyr
 2785 2790 2795 2800
 Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser Met Val Asn Gly Val Val
 2805 2810 2815
 Arg Leu Leu Thr Lys Pro Trp Asp Val Val Pro Met Val Thr Gln Met
 2820 2825 2830
 Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys Glu
 2835 2840 2845
 Lys Val Asp Thr Arg Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu
 2850 2855 2860
 Met Lys Ile Thr Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys
 2865 2870 2875 2880
 Thr Pro Arg Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser
 2885 2890 2895
 Asn Ala Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser
 2900 2905 2910
 Ala Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
 2915 2920 2925
 Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn
 2930 2935 2940
 Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
 2945 2950 2955 2960
 Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu Glu
 2965 2970 2975
 Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser Arg Glu
 2980 2985 2990
 Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys Leu Gly Tyr
 2995 3000 3005
 Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala Met Tyr Ala Asp
 3010 3015 3020
 Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu Glu Asp Leu Lys Asn
 3025 3030 3035 3040
 Glu Glu Met Val Thr Asn His Met Glu Gly Glu His Lys Lys Leu Ala
 3045 3050 3055
 Glu Ala Ile Phe Lys Leu Thr Tyr Gln Asn Lys Val Val Arg Val Gln
 3060 3065 3070
 Arg Pro Thr Pro Arg Gly Thr Val Met Asp Ile Ile Ser Arg Arg Asp
 3075 3080 3085
 Gln Arg Gly Ser Gly Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr
 3090 3095 3100
 Asn Met Glu Ala Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe
 3105 3110 3115 3120
 Lys Ser Ile Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn
 3125 3130 3135
 Trp Leu Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser
 3140 3145 3150
 Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln
 3170 3175 3180
 Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195 3200
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu Val
 3205 3210 3215
 Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg Ile Ser
 3220 3225 3230
 Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu Gly Lys Ser
 3235 3240 3245
 Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg Arg Asp Leu Arg
 3250 3255 3260
 Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro Ser His Trp Val Pro
 3265 3270 3275 3280
 Thr Ser Arg Thr Thr Trp Ser Ile His Ala Lys His Glu Trp Met Thr
 3285 3290 3295
 Thr Glu Asp Met Leu Thr Val Trp Asn Arg Val Trp Ile Gln Glu Asn
 3300 3305 3310

Pro Trp Met Glu Asp Lys Thr Pro Val Glu Ser Trp Glu Glu Ile Pro
 3315 3320 3325
 Tyr Leu Gly Lys Arg Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu
 3330 3335 3340
 Thr Ser Arg Ala Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln
 3345 3350 3355 3360
 Val Arg Ser Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser
 3365 3370 3375
 Met Lys Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

<210> SEQ ID NO 31

<211> COMPRIMENTO: 17

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 31.

tagagagcag atctctg 17

<210> SEQ ID NO 32

<211> COMPRIMENTO: 32

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 32

tgtgaccatg ccggctgcga tgcactcacc ga 32

<210> SEQ ID NO 33

<211> COMPRIMENTO: 38

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 33

tgctggccac ttaactacgc gtgatggaga gccgcgca 38

<210> SEQ ID NO 34

<211> COMPRIMENTO: 33

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 34

tgtaatgatg ccggccgcga tgcataaaaa tga 33

<210> SEQ ID NO 35

<211> COMPRIMENTO: 38

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 35

ggcgtttcac ttgtcaacgc gtgatggcga acccctca 38

<210> SEQ ID NO 36

<211> COMPRIMENTO: 36

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 36

agtgattccg ccggcagcta tgcacgtcat agccat 36

<210> SEQ ID NO 37

<211> COMPRIMENTO: 39

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 37

ggagcagtta ccctctctac gcgtcaaggg aaggtgatg 39

<210> SEQ ID NO 38

<211> COMPRIMENTO: 44

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 38

gaagagcaga actccgccgg ctgcgagaaa cgtgagagct atgg 44

<210> SEQ ID NO 39

<211> COMPRIMENTO: 28

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 39

```

Ser1 Ala Gly Met Ile5 Ile Met Leu Ile10 Pro Thr Val Met Ala Phe His15
Leu Thr Thr Arg Gln Gly Lys Val20 Met25 Met Thr Val

```

<210> SEQ ID NO 40

<211> COMPRIMENTO: 8

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 40

```

Gln1 Gly Lys Val5 Met Met Thr Val

```

<210> SEQ ID NO 41

<211> COMPRIMENTO: 56

<212> TIPO: ADN

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 41

aattcaacgc gtacatccgc gggcaccgga attgcagtca tgattggcct gatggc 56

<210> SEQ ID NO 42

<211> COMPRIMENTO: 32

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 42

Ser	Ala	Gly	Thr	Gly	Ile	Ala	Val	Met	Ile	Gly	Leu	Ile	Ala	Ser	Val
1				5					10					15	
Gly	Ala	Val	Thr	Leu	Ser	Asn	Phe	Gln	Gly	Lys	Val	Met	Met	Thr	Val
			20					25				30			

<210> SEQ ID NO 43

<211> COMPRIMENTO: 29

<212> TIPO: PRT

<213> ORGANISMO: Sequência artificial

<220> CARACTERÍSTICA:

<223> OUTRAS INFORMAÇÕES: Descrição da sequência artificial:/Nota = construção sintética

<400> SEQUÊNCIA: 43

Thr	Gly	Ile	Ala	Val	Met	Ile	Gly	Leu	Ile	Ala	Ser	Val	Gly	Ala	Val
1				5					10					15	
Thr	Leu	Ser	Asn	Phe	Gln	Gly	Lys	Val	Met	Met	Thr	Val			
			20					25							

DOCUMENTOS REFERIDOS NA DESCRIÇÃO

Esta lista de documentos referidos pelo autor do presente pedido de patente foi elaborada apenas para informação do leitor. Não é parte integrante do documento de patente europeia. Não obstante o cuidado na sua elaboração, o IEP não assume qualquer responsabilidade por eventuais erros ou omissões.

Documentos de patente referidos na descrição

- WO 9306214 A [0009]
- WO 9837911 A [0009]
- EP 01910882 A [0162]
- US 0105142 W [0162]
- US 60182829 B [0162]

Documentos de não patente citados na descrição

- **CHAMBERS et al.** *J Virol*, 1999, vol. 73 (4), 3095-3101 [0009]
- **KINNEY et al.** *Virology*, 1993, vol. 67, 1269-1277 [0055]
- **SAMBROOK ; FRITSCH ; MANIATIS.** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, 1989 [0087]
- **TYAGI; KRAMER.** *Mature Biotechnology*, 1996, vol. 14, 303-308 [0094]
- **YOKSAN et al.** Dengue virus vaccine development: study on biological markers of uncloned dengue 1-4 viruses serially passaged in primary kidney cells. *Arbovirus research in Australia, Proceedings of the Fourth Symposium. CSIRO/QIMR, Brisbane, Australia*, 1986, 35-38 [0096]
- **BHAMARAPRAVATI ; SUTEE.** *Vaccine*, 2000, vol. 18, 44-47 [0096]
- **MILLER et al.** *Am. J. Trop. Med. & Hyg.*, 1986, vol. 35, 1302-1309 [0097] [0141]
- **KINNEY et al.** *Virology*, 1997, vol. 230, 300-308 [0098] [0142]

- **FU et al.** *Virology*, 1992, vol. 88, 953-958 **[0103]**
- **MULLER et al.** *Science*, 1994, vol. 264, 1918-1921 **[0106]**
- **KINNEY et al.** *J. Virol*, 1997, vol. 230, 300-308 **[0126]**
- **LILJESTROM et al.** *J. Virology*, 1991, vol. 63, 4107-4113 **[0143]**
- List includes the currently classified members of the Flavivirus genus in the International Catalogue of Arboviruses. Including Certain Other Viruses of Vertebrates. The American Society of Tropical Medicine and Hygiene, 1985 **[0161]**

Lisboa, 16 de Novembro de 2015

REIVINDICAÇÕES

1. Uma composição de flavivírus imunogénico que compreende

(i) uma quimera de ácido nucleico que compreende uma primeira sequência de nucleótidos que codifica para proteínas não estruturais e uma proteína da cápside (C) de um vírus do dengue-2, em que a proteína C não inclui o péptido de sinal do terminal-C da proteína do dengue-2 C, e uma segunda sequência de nucleótidos que codifica para o péptido de sinal do terminal-C da proteína C, uma proteína de pré-membrana (prM) e uma proteína do invólucro (E) de um vírus do Nilo Ocidental, ou;

(ii) um primeiro ácido nucleico que compreende uma sequência de nucleótidos que codifica para uma estirpe PDK53 do vírus do dengue-2 atenuado; um segundo ácido nucleico que compreende uma sequência de nucleótidos que codifica para proteínas não estruturais e uma proteína da cápside (C) de uma estirpe PDK53 do vírus do dengue-2 atenuado e uma sequência de nucleótidos que codifica para proteínas da pré-membrana (prM) e do invólucro (E) do vírus do dengue-1, ou que compreende uma sequência de nucleótidos que codifica para proteínas não estruturais de uma estirpe PDK53 do vírus do dengue-2 atenuado e uma sequência de nucleótidos que codifica para as proteínas C, prM e E do vírus do dengue-1;

um terceiro ácido nucleico que compreende uma sequência de nucleótidos que codifica para proteínas não estruturais e uma proteína C de uma estirpe PDK53 do vírus do dengue-2 atenuado e uma sequência de nucleótidos que codifica para as proteínas prM e E do vírus do dengue-3, em que a proteína E contém uma leucina na posição do aminoácido 345; e

um quarto ácido nucleico que compreende uma sequência de nucleótidos que codifica para proteínas não estruturais e

uma proteína C de uma estirpe PDK53 do vírus do dengue-2 atenuado e uma sequência de nucleótidos que codifica para as proteínas prM e E do vírus do dengue-4, em que a proteína E contém uma leucina na posição do aminoácido 447 e em que a proteína C contém uma serina na posição 100; em que a composição imunogénica compreende um transportador ou veículo farmacêuticamente aceitável, ou um adjuvante.

2. A composição da reivindicação 1, parte (i), em que o vírus do dengue-2 é o vírus do dengue-2 16681.

3. A composição da reivindicação 1, parte (i) que codifica para a sequência de aminoácidos da SEQ ID NO: 18.

4. A composição da reivindicação 3, que compreende a sequência de ácido nucleico da SEQ ID NO: 17.

5. A composição da reivindicação 1, parte (i), em que o vírus do dengue-2 é um vírus do dengue-2 atenuado.

6. A composição da reivindicação 5, em que o vírus do dengue-2 atenuado é a estirpe de vacina PDK-53.

7. A composição da reivindicação 1, parte (ii), em que a primeira molécula de ácido nucleico codifica para a sequência de aminoácidos da SEQ ID NO: 16, e/ou a segunda molécula de ácido nucleico codifica para a sequência de aminoácidos de qualquer uma das SEQ ID NO: 6, 8, e 28, e/ou a terceira molécula de ácido nucleico codifica para a sequência de aminoácidos da SEQ ID NO: 10, e/ou a quarta molécula de ácido nucleico codifica para a sequência de aminoácidos da SEQ ID NO: 12.

8. A composição da reivindicação 7, em que a primeira

molécula de ácido nucleico compreende a sequência de ácido nucleico da SEQ ID NO: 15, e/ou a segunda molécula de ácido nucleico compreende a sequência de ácido nucleico de qualquer uma das SEQ ID NO: 5, 7, e 27, e/ou a terceira molécula de ácido nucleico compreende a sequência de ácido nucleico da SEQ ID NO: 9, e/ou a quarta molécula de ácido nucleico compreende a sequência de ácido nucleico da SEQ ID NO: 11.

9. A composição de qualquer uma das reivindicações de 1 a 8, que compreende ainda uma composição imunizante selecionada a partir do grupo que consiste de uma vacina do vírus do dengue, uma vacina do vírus da febre amarela, uma vacina do vírus da encefalite transmitida por carraças, uma vacina do vírus da encefalite japonesa, uma vacina do vírus do Nilo Ocidental, uma vacina do vírus da hepatite C, e qualquer combinação das mesmas.

10. Uma composição de acordo com qualquer uma das reivindicações de 1 a 9 para utilização num método para a indução de uma resposta imunológica num sujeito, em que o método compreende a administração de uma quantidade eficaz da composição de qualquer uma das reivindicações de 1 a 9 ao sujeito.

11. Utilização da composição de qualquer uma das reivindicações de 1 a 9 no fabrico de uma vacina para a indução de uma resposta imunológica a um flavivírus.

Lisboa, 16 de Novembro de 2015

FIGURA 1

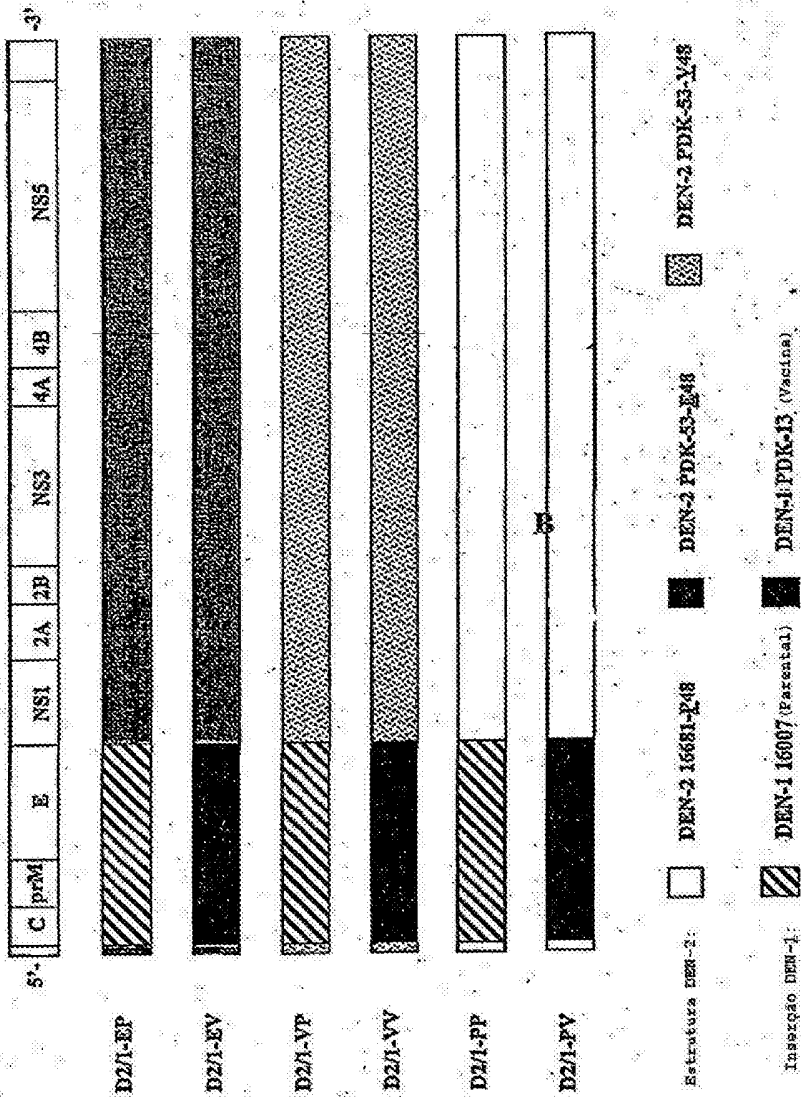


FIGURA 2

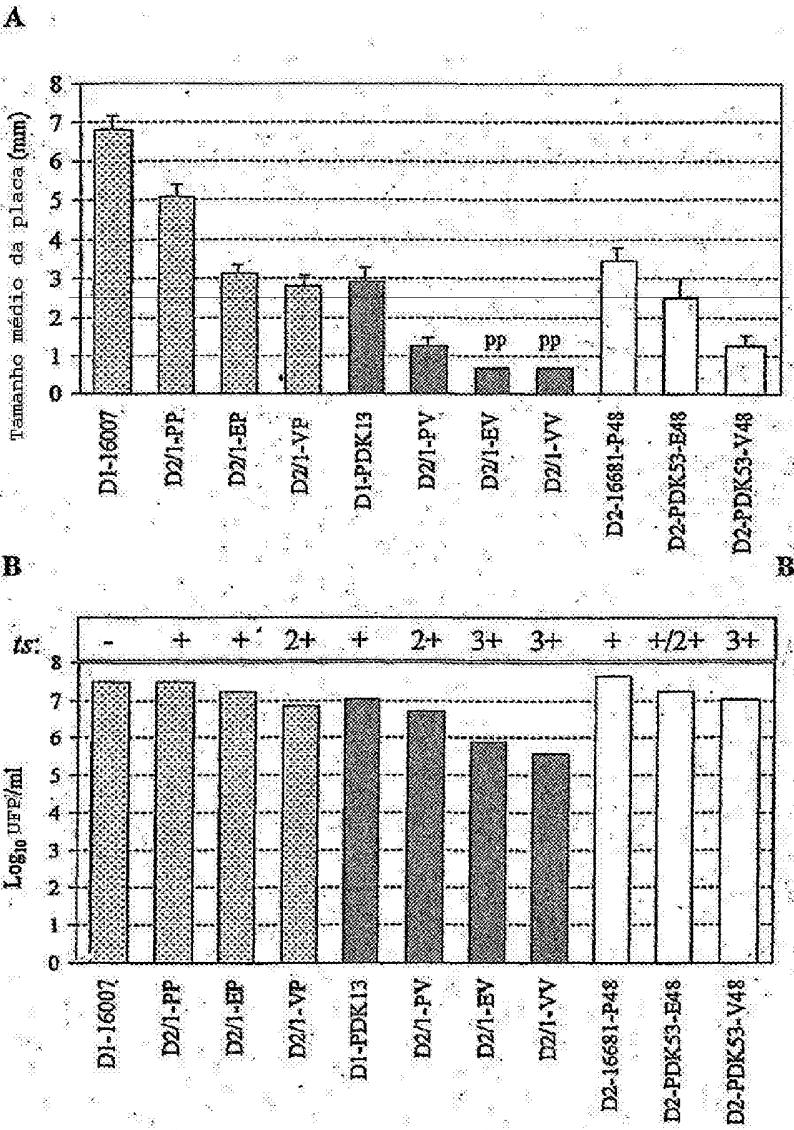


FIGURA 3

