



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI 0713646-3 A2**

(22) Data de Depósito: 14/06/2007
(43) Data da Publicação: 23/10/2012
(RPI 2181)



(51) *Int.Cl.:*
C07K 14/325
C12N 15/82

(54) Título: MOLÉCULA DE ÁCIDO NUCLEICO ISOLADA, CÉLULA HOSPEDEIRA, PLANTA E SEMENTE TRANSGÊNICAS, POLIPETÍDEO ISOLADO COM ATIVIDADE PESTICIDA E MÉTODO PARA SUA PRODUÇÃO, ANTICORPO, COMPOSIÇÃO E MÉTODOS PARA CONTROLE DE POPULAÇÃO DE PRAGAS, PARA EXTERMINAR UMA PRAGA DE LEPIDÓPTEROS OU COLEÓPTEROS, PARA PROTEGER UMA PLANTA DE UMA PRAGA E DE UMA PRAGA DE NEMATÓIDES

(30) Prioridade Unionista: 14/06/2006 US 60/813,774

(73) Titular(es): Athenix Corporation

(72) Inventor(es): Daniel John Tomso, Julia Thissen, Nadine Carozzi, Nalini Desai, Nicholas B. Duck, Rong Guo, Shruti Jain, Theodore Kahn

(74) Procurador(es): Claudia Christina Schulz

(86) Pedido Internacional: PCT US2007071173 de 14/06/2007

(87) Publicação Internacional: WO 2007/147029 de 21/12/2007

(57) Resumo: MOLÉCULA DE ÁCIDO NUCLEÍCO ISOLADA, VETOR, CÉLULA HOSPEDEIRA, PLANTA E SEMENTE TRANSGÊNICAS, POLIPEPTÍDEO ISOLADO COM ATIVIDADE PESTICIDA E MÉTODO PARA SUA PRODUÇÃO, ANTICORPO, COMPOSIÇÃO E MÉTODOS PARA CONTROLE DE POPULAÇÃO DE PRAGAS, PARA EXTERMINAR UM PRAGA DE LEPIDÓPTEROS OU COLEÓPTEROS, PARA PROTEGER UMA PLANTA DE UMA PRAGA E DE UMA PRAGA DE NEMATÓIDES. São fornecidas composições e métodos para conferir atividades pesticida a bactérias, plantas, células vegetais, tecidos e sementes. São fornecidas composições compreendendo uma sequência codificante de um polipeptídeo de delta-endotoxina. As sequências codificantes podem ser usadas em construções de DNA ou cassetes de expressão para transformação e expressão em plantas e bactérias. Composições também compreendem bactérias, plantas, células vegetais, tecidos e sementes transformadas. Em particular, são fornecidas moléculas de ácido nucleico de delta-endotoxinas isoladas. Adicionalmente, são englobadas sequências de aminoácidos. Em particular, a presente invenção proporciona moléculas de ácido nucleico compreendendo as sequências de nucleotídeos codificantes das sequências de aminoácidos mostradas em SE ID NO: 2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38, ou a sequência de nucleotídeos apresentada em SE ID NO: 1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37, bem como variantes e fragmentos destas.

RELATÓRIO DESCRITIVO

MOLÉCULA DE ÁCIDO NUCLEÍCO ISOLADA; VETOR; CÉLULA
HOSPEDEIRA; PLANTA E SEMENTE TRANSGÊNICAS; POLIPEPTÍDEO
5 ISOLADO COM ATIVIDADE PESTICIDA E MÉTODO PARA A SUA
PRODUÇÃO; ANTICORPO; COMPOSIÇÃO E MÉTODO PARA CONTROLE DE
POPULAÇÕES DE PRAGAS PARA EXTERMINAR UMA PRAGA DE
LEPIDÓPTEROS OU COLEÓPTEROS; PARA PROTEGER UMA PLANTA DE
UMA PRAGA DE UMA PRAGA DE NEMATÓIDE

10

CAMPO DA INVENÇÃO

Esta invenção diz respeito ao campo da biologia
molecular. São fornecidos novos genes que codificam
proteínas pesticidas. Estas proteínas e as sequências de
15 ácidos nucleicos que as codificam são úteis na preparação
de formulações pesticidas e na produção de plantas
transgênicas resistentes à praga.

FUNDAMENTOS DA INVENÇÃO

20 *Bacillus thuringiensis* é uma bactéria do solo Gram-
positiva formadora de esporos caracterizada pela sua
capacidade de produzir inclusões cristalinas que são
especificamente tóxicas para certas espécies e ordens de
insetos, mas são inofensivas para as plantas e outros
25 organismos não-alvo. Por esta razão, composições incluindo
cepas de *Bacillus thuringiensis* ou suas proteínas
inseticidas podem ser utilizadas como inseticidas
ambientalmente aceitáveis para controlar pragas agrícolas
ou insetos vetores para uma variedade de doenças humanas ou
30 animais.

Proteínas Crystal (Cry) (endotoxinas delta) de *Bacillus thuringiensis* têm potente atividade inseticida predominantemente contra lepidópteros, dípteros, e larvas coleópteras. Estas proteínas também têm demonstrado atividade contra pragas das ordens Hymenoptera, Homoptera, Phthiraptera, Mallophaga e Acari, bem como de outros invertebrados, como os das ordens Nemathelminthes, Platyhelminthes, e Sarcomastigophora (Feitelson (1993), The *Bacillus Thuringiensis* family tree. In *Advanced Engineered Pesticides*, Marcel Dekker, Inc., New York, N.Y.) Estas proteínas foram inicialmente classificadas como CryI a CryV baseado essencialmente em sua atividade inseticida. As principais classes foram Lepidoptera-específica (I), Lepidoptera e Diptera-específica (II), Coleoptera-específica (III), Díptera-específica (IV), e nematóide-específicas (V) e (VI). As proteínas foram ainda classificadas em subfamílias; para proteínas mais fortemente relacionadas dentro de cada família foram atribuídas letras divisórias, como *Cry1A*, *Cry1B*, *Cry1C*, etc. Para proteínas ainda mais estreitamente relacionadas dentro de cada divisão foram dados nomes como *Cry1C1*, *Cry1C2*, etc.

Uma nova nomenclatura foi descrita recentemente para os genes Cry com base em homologias de sequência de aminoácidos, em vez de especificidade de insetos-alvo (Crickmore et al. (1998) *Microbiol. Mol. Biol. Rev.* 62: 807-813). Na nova classificação, a cada toxina é atribuído um nome único incorporando uma categoria primária (um número arábico), uma categoria derivada (uma letra

maiúscula), uma terceira categoria (uma letra minúscula), e uma categoria quaternária (outro número arábico). Na nova classificação, os numerais romanos foram trocados por números arábicos na primeira categoria. Proteínas com menos
5 de 45% de identidade de sequência têm diferentes categorias primárias, e os critérios para as categorias secundária e terciária são 78% e 95%, respectivamente.

A proteína cristal não apresenta atividade inseticida
10 até ter sido ingerida e solubilizada no intestino médio do inseto. A protoxina ingerida é hidrolisada por proteases no trato digestivo do inseto a uma molécula ativa tóxica. (Höfte e Whiteley (1989) *Microbiol. Rev.* 53: 242-255). Esta toxina se liga a escova apical dos receptores de borda no
15 intestino médio das larvas alvo e se insere na membrana apical criando canais iônicos ou poros, resultando na morte larval.

Endotoxinas delta geralmente têm cinco domínios de
20 sequência conservados, e três domínios estruturais conservados (ver, por exemplo, de Maagd et al. (2001) *Trends Genetics* 17: 193-199). O primeiro domínio estrutural conservado consiste de sete alfa hélices e está envolvido na inserção da membrana e na formação dos poros. O domínio
25 II consiste de três lâminas beta dispostas em uma configuração GreeK key, e o domínio III consiste de duas lâminas beta antiparalelas em formato "jelly-roll" (de Maagd et al., 2001, supra). Os domínios II e III estão envolvidos no reconhecimento e ligação do receptor, e,

portanto, são considerados determinantes na especificidade da toxina.

Devido à devastação que os insetos podem conferir
5 existe uma constante necessidade de descobrir novas formas de endotoxinas delta de *Bacillus thuringiensis*.

RESUMO DA INVENÇÃO

Composições e métodos para conferir resistência a
10 pragas em bactérias, plantas, células vegetais, tecidos e sementes são fornecidos. Composições incluem moléculas de ácidos nucleicos codificando sequências para polipeptídeos endotoxina delta, vetores compreendendo essas moléculas de ácidos nucleicos, e células hospedeiras compreendendo os
15 vetores. Composições incluem também sequências de polipeptídeo da endotoxina, e anticorpos para esses polipeptídeos. As sequências nucleotídicas podem ser utilizadas em construções de DNA ou cassetes de expressão para transformação e expressão em organismos, incluindo
20 microorganismos e plantas. Os nucleotídeos ou sequências de aminoácidos podem ser sintéticos que tenham sido concebidos para a expressão em um organismo, incluindo, mas não se limitando a, um microorganismo ou uma planta. Composições também compreendem bactérias transformadas, plantas,
25 células vegetais, tecidos e sementes.

Em particular, moléculas isoladas de ácidos nucleicos correspondentes a sequências de ácido nucleico delta-endotoxina são fornecidas. Adicionalmente, sequências de
30 aminoácidos correspondentes aos polinucleotídeos são

englobadas. Em particular, a presente invenção fornece uma molécula do ácido nucléico isolada compreendendo uma seqüência nucleotídica que codifica a seqüência de aminoácidos mostrada em SEQ ID NO: 2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38, uma seqüência nucleotídica estabelecida em SEQ ID NO: 1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, ou 37, ou as seqüências nucleotídicas endotoxina delta depositadas nas bactérias hospedeiras com N° de Acesso B-30935, B-30936, B-30937, e B-50046, bem como as variantes e os fragmentos destas. Sequências nucleotídicas que são complementares a uma seqüência nucleotídica da invenção, ou que hibridiza para uma seqüência da invenção também são englobadas.

15

Métodos são fornecidos para produzir os polipeptídeos da invenção, e para utilizar esses polipeptídeos para controlar ou matar uma praga lepidóptera ou coleóptera. Métodos e kits para a detecção em uma amostra de ácidos nucléicos e polipeptídeos da invenção também estão incluídos.

20

Métodos para controlar ou matar uma população de praga nematóide são ainda fornecidos. Os métodos incluem introduzir em uma planta um polinucleotídeo codificando um polipeptídeo nematóide-ativo com um tamanho molecular superior a 22 kDa. Estes polipeptídeos nematóide-ativos são úteis para controlar ou matar nematodos parasitas de plantas, particularmente cistos de nematóides.

25

30

As composições e métodos da invenção são úteis para a produção de organismos com resistência aos pesticidas, especialmente bactérias e plantas. Estes organismos e composições derivados deles são desejáveis para fins agrícolas. As composições da invenção também são úteis para gerar proteínas endotoxina delta alteradas ou melhoradas que têm atividade pesticida, ou para detectar a presença de proteínas ou ácidos nucleicos endotoxina delta em produtos ou organismos.

10

DESCRIÇÃO DETALHADA DA INVENÇÃO

A presente invenção é projetada para composições e métodos para regulação de resistência a pragas nos organismos, particularmente plantas ou células vegetais. Os métodos envolvem organismos transformados com uma sequência nucleotídica que codifica uma proteína endotoxina delta da invenção. Em particular, as sequências nucleotídicas da invenção são úteis para preparar plantas e microorganismos que possuem atividade pesticida. Assim, bactérias transformadas, plantas, células vegetais, tecidos vegetais e sementes são fornecidos. Composições são ácidos nucleicos e proteínas endotoxina delta de *Bacillus thuringiensis*. As sequências encontram uso na construção de vetores de expressão para posterior transformação em organismos de interesse, como sondas para o isolamento de outros genes endotoxina delta, e para a geração de proteínas pesticidas alteradas por métodos conhecidos na arte, como a troca de domínio ou embaralhamento de DNA. As proteínas encontram utilização no controle ou morte de lepdópteros,

25

coleópteros, e populações de pragas nematóides e para a produção de composições com atividade pesticida.

Plasmídeos contendo a sequência nucleotídica da
5 invenção foram depositados na coleção permanente do
Agricultural Research Service Culture Collection, Northern
Regional Research Laboratory (NRRL), 1815 North University
Street, Peoria, Illinois 61604, Estados Unidos da América,
em

10 9 de junho de 2006, e atribuídos os Nos. de Acesso
NRRL B-30935 (para *axmi-031*), NRRL B-30936 (para *axmi-039*),
NRRL B-30937 (para *axmi-040*) e em 29 de maio de 2007 e
atribuído o No. de Acesso NRRL B-50046 (*axmi-049*). Estes
depósitos serão mantidos nos termos do Tratado de Budapeste
15 sobre o Reconhecimento Internacional do Depósito de
Microorganismos para fins de Processo de Patente. Estes
depósitos foram feitos apenas como uma conveniência para
aqueles habilitados na arte e não são uma admissão de que
um depósito seja necessário sob 35 USC §112.

20

Por "endotoxina delta" entende-se uma toxina de
Bacillus thuringiensis que possui atividade tóxica contra
uma ou várias pragas, incluindo mas não limitando-se a,
membros das ordens *Lepidoptera*, *Díptera* e *Coleoptera* e
25 membros do filo *Nematoda*, ou uma proteína que tenha
homologia com essa proteína. Em alguns casos, proteínas
endotoxina delta têm sido isoladas de outros organismos,
incluindo *Clostridium bifermentans* e *Paenibacillus*
popilliae. Proteínas endotoxina delta incluem sequências de
30 aminoácidos deduzidas de sequências nucleotídicas de

comprimento total divulgadas aqui, e sequências de aminoácidos que são mais curtos que as sequências de comprimento total, quer devido à utilização de um suplente a jusante no início do sítio, ou devido ao tratamento que produz uma proteína menor com atividade pesticida. O tratamento pode ocorrer no organismo da proteína que é expressa, ou na praga após a ingestão da proteína. Endotoxinas delta incluem proteínas identificadas como *cry1*, através de *cry43*, *cyt1* e *cyt2*, e toxina Cyt. Existem atualmente mais de 250 espécies conhecidas de endotoxinas delta com uma vasta gama de especificidades e toxicidades. Para uma lista extensa ver Crickmore et al. (1998), *Microbiol. Mol. Biol. Rev.* 62: 807-813, e para as atualizações regulares ver Crickmore et al. (2003) "*Bacillus thuringiensis* toxin nomenclature", no [www.biols.susx.ac.uk/ Home/Neil_Crickmore/Bt/index](http://www.biols.susx.ac.uk/Home/Neil_Crickmore/Bt/index).

Aqui são fornecidas novas sequências isoladas de nucleotídeos que conferem atividade pesticida. São fornecidas, também, as sequências de aminoácidos de proteínas endotoxina delta. A proteína resultante da tradução deste gene permite às células controlar ou matar pragas que a ingere.

25 Moléculas Isoladas de Ácido Nucléico, e variantes e fragmentos destas

Um aspecto da invenção pertence a moléculas de ácidos nucleicos isoladas ou recombinantes compreendendo sequências nucleotídicas codificando proteínas endotoxina delta e polipeptídeos ou partes biologicamente ativas dos

mesmos, bem como moléculas de ácidos nucléicos suficientes para utilização como sondas de hibridização para identificar endotoxinas delta codificando ácidos nucléicos. Conforme utilizado aqui, o termo "molécula de ácido nucléico"

destina-se a incluir moléculas de DNA (por exemplo, DNA recombinante, cDNA ou DNA genômico) e moléculas de RNA (por exemplo, mRNA) e análogos do DNA ou RNA gerados utilizando nucleotídeos análogos. A molécula de ácido nucléico pode ser de fita única ou de fita dupla, mas de preferência é um DNA de dupla fita.

Uma molécula de ácido nucléico ou proteína "isolada" ou "purificada", ou uma porção biologicamente ativa dela é praticamente isenta de outros materiais celulares, ou meio de cultura quando produzida por técnicas recombinantes, ou substancialmente livres de precursores químicos ou outros produtos químicos quando quimicamente sintetizados. De preferência, um ácido nucléico "isolado" é livre de sequências (de preferência proteína codificando sequências) que naturalmente flanqueiam o ácido nucléico (ou seja, sequências localizadas nas extremidades 5' e 3' do ácido nucléico) no DNA genômico do organismo do qual o ácido nucléico é derivado. Para efeitos da invenção, "isoladas" quando usado para referir-se a moléculas de ácido nucléico exclui cromossomos isolados. Por exemplo, em vários experimentos, a endotoxina delta isolada codificando a molécula de ácido nucléico pode conter menos de cerca de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb, ou 0,1 kb de sequências nucleotídicas que naturalmente flanqueiam a

molécula de ácido nucléico no DNA genômico da célula a partir do qual o ácido nucléico é derivado. A proteína endotoxina delta que é substancialmente livre de materiais celulares inclui preparações de proteína com menos de cerca
5 de 30%, 20%, 10%, ou 5% (em peso seco) de proteína endotoxina não-delta (também aqui referida como uma "proteína contaminante").

Sequências nucleotídicas codificando as proteínas da
10 presente invenção incluem a sequência estabelecida na SEQ ID NO: 1, 3, 5, 34 e 37, as sequências nucleotídicas endotoxina delta depositadas em bactérias hospedeiras como No. De Acesso NRRL B-30935, B-30936, B-30937, e B-50046, e variantes, fragmentos e complementos das mesmas (por
15 exemplo, SEQ ID NO: 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, e 32). Por "complementar" entende-se uma sequência nucleotídica que seja suficientemente complementar a uma determinada sequência nucleotídica que possa hibridizar a dada sequência de nucleotídeos, para assim formar um duplex
20 estável. A sequência de aminoácidos correspondente a proteínas endotoxina delta codificada por esta sequência nucleotídica é estabelecida na SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 e 37.

Moléculas de ácidos nucléicos que são fragmentos
25 destas sequências nucleotídicas de codificação de endotoxina delta também são abrangidas pela presente invenção (por exemplo, SEQ ID NO: 9, 11, 14, 18, 20, 22 e 24). Por "fragmento" entende-se uma porção da sequência nucleotídica que codifica uma proteína endotoxina delta. Um
30 fragmento de uma sequência nucleotídica pode codificar uma

porção biologicamente ativa de uma proteína endotoxina delta, ou pode ser um fragmento que pode ser usado como uma sonda de hibridização ou iniciador de PCR utilizando métodos divulgados abaixo. Moléculas de ácidos nucleicos

5 que são fragmentos de uma sequência nucleotídica endotoxina delta compreende, pelo menos, cerca de 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2150, 2200,

10 2250, 2300, 2350, 2400, 2450, 2500, 2550, 2600, 2650, 2700, 2750, 2800, 2850, 2900, 2950, 3000, 3050, 3100, 3150, 3200, 3250, 3300, 3350 nucleotídeos contíguos, ou até o número de nucleotídeos presentes em uma sequência nucleotídica de codificação de endotoxina delta de comprimento total

15 divulgada aqui (por exemplo, 3558 nucleotídeos para SEQ ID NO: 1, 3984 nucleotídeos para SEQ ID NO: 3, 3720 nucleotídeos para SEQ ID NO: 5, e 3669 nucleotídeos para SEQ ID NO: 34), dependendo da uso pretendido. Por nucleotídeos "contíguos"

20 é entendido resíduo de nucleotídeos que são imediatamente adjacentes um ao outro. Fragmentos das sequências nucleotídicas da presente invenção irão codificar fragmentos de proteínas que mantêm a atividade biológica da proteína endotoxina delta e, conseqüentemente, manter a

25 atividade pesticida. Por "manter atividade" entende-se que o fragmento terá pelo menos cerca de 30%, pelo menos cerca de 50%, pelo menos cerca de 70%, 80%, 90%, 95% ou mais de atividade pesticida da proteína endotoxina delta. Métodos para a medição da atividade pesticida são bem conhecidos na

30 arte. Ver, por exemplo, Czapla e Lang (1990) *J. Econ.*

Entomol. 83: 2480-2485; Andrews et al. (1988) *Biochem. J.* 252: 199-206; Marrone et al. (1985) *J. of Economic Entomology* 78: 290-293; e Patente US No. 5.743.477, todos os quais aqui incorporados por referência na sua
5 totalidade.

Um fragmento de uma sequência nucleotídica codificando uma endotoxina delta que codifica uma porção biologicamente ativa de uma proteína da invenção vai codificar, pelo
10 menos, cerca de 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100 aminoácidos contíguos, ou até o número total de aminoácidos presentes em uma proteína endotoxina delta de comprimento total da invenção
15 (por exemplo, 1185 aminoácidos para SEQ ID NO: 2, 1327 aminoácidos para SEQ ID NO: 4, 1239 aminoácidos para SEQ ID NO: 6, e 1223 aminoácidos de SEQ ID NO: 35).

20 Proteínas endotoxina delta preferidas da presente invenção são codificadas por uma sequência nucleotídica suficientemente idêntica à sequência de nucleotídeos SEQ ID NO: 1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, ou 37. Por "suficientemente idêntica" se
25 entende um aminoácido ou sequência nucleotídica que tem pelo menos cerca de 60% ou 65% de identidade de sequência, cerca de 70% ou 75% de identidade de sequência, cerca de 80% ou 85% de identidade de sequência, cerca de 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou uma identidade de
30 sequência maior comparada com uma sequência de referência

usando um dos programas de alinhamento de sequência descritos aqui utilizando parâmetros padrões. Um especialista na arte vai reconhecer que estes valores podem ser devidamente adaptados para determinar identidades correspondentes das proteínas codificadas por duas sequências nucleotídicas, levando em conta a degeneração de códon, similaridade de aminoácidos, posicionamento do molde de leitura, e coisas do gênero.

Para determinar a identidade percentual de duas sequências de aminoácidos ou de dois ácidos nucléicos, as sequências são alinhadas para propósitos de uma comparação ótima. A identidade percentual entre as duas sequências é uma função do número de posições idênticas partilhadas pelas sequências (ou seja, identidade percentual = número de posições idênticas / número total de posições (por exemplo, sobreposição de posições) x 100). Em um experimento, as duas sequências têm o mesmo comprimento. A identidade percentual entre as duas sequências pode ser determinada utilizando técnicas semelhantes às descritas abaixo, com ou sem permitir espaços. No cálculo da identidade percentual, correspondências tipicamente exatas são contadas.

A determinação da identidade percentual entre duas sequências pode ser realizada utilizando um algoritmo matemático. Um exemplo não limitante de um algoritmo matemático utilizado para a comparação das duas sequências é o algoritmo de Karlin e Altschul (1990) *Proc. Natl. Acad. Sci. E.U.A.* 87: 2264, modificado como no Karlin e Altschul

(1993) *Proc. Natl. Acad. Sci. E.U.A.* 90: 5873-5877. Tal algoritmo é incorporado nos programas BLASTN e BLASTX de Altschul et al. (1990) *J. Mol. Biol.* 215: 403. Pesquisas de nucleotídeo BLAST podem ser realizadas com o programa

5 BLASTN, *escore* = 100, *comprimento de palavra* = 12, para obter sequências nucleotídicas homólogas de moléculas de ácido nucléico endotoxina delta da invenção. Pesquisas de proteína BLAST podem ser realizadas com o programa BLASTX, *escore* = 50, *comprimento de palavra* = 3, para obter

10 aminoácidos homólogos às moléculas de proteína endotoxina delta da invenção. Para obter alinhamentos com lacunas para fins de comparação, Gapped BLAST (em BLAST 2.0) pode ser utilizado tal como descrito em Altschul et al. (1997) *Nucleic Acids Res.* 25: 3389. Alternativamente, PSI-Blast

15 pode ser utilizado para a realização de uma pesquisa que detecta relações distantes entre as moléculas. Ver Altschul et al. (1997) supra. Quando se utilizar programas BLAST, Gapped BLAST, e PSI-Blast, os parâmetros padrão dos respectivos programas (por exemplo, BLASTX e BLASTN) podem

20 ser utilizados. O alinhamento também pode ser realizado manualmente, por inspeção.

Outro exemplo não limitante de um algoritmo matemático utilizado para a comparação de sequências é o algoritmo

25 ClustalW (Higgins et al. (1994) *Nucleic Acids Res.* 22: 4673-4680). ClustalW compara sequências, e alinha a totalidade dos aminoácidos ou sequência de DNA, e assim pode fornecer dados sobre a conservação de toda a sequência de aminoácidos. O algoritmo ClustalW é utilizado em vários

30 software de soluções comercialmente disponíveis de análise

de DNA/aminoácido, tais como o módulo ALIGNX do Vector NTI Program Suite (Invitrogen Corporation, Carlsbad, CA). Após o alinhamento das sequências de aminoácidos com ClustalW, a identidade percentual de aminoácidos pode ser avaliada. Um exemplo não-limitante de um programa de software útil para a análise de alinhamentos ClustalW é GENEDOC (TM). GENEDOC (TM) (Karl Nicholas) permite a avaliação de similaridade e de identidade de aminoácidos (ou DNA) entre várias proteínas. Outro exemplo não-limitante de algoritmo matemático utilizado para a comparação de sequências é o algoritmo de Myers e Miller (1988) *CABIOS* 4:11-17. Tal algoritmo é incorporado no programa ALIGN (versão 2.0), que faz parte do Pacote de Software GCG Wisconsin Genetics, Versão 10 (disponível a partir de Accelrys, Inc., 9685 Scranton Rd., San Diego, CA, E.U.A.). Quando se utilizar o programa ALIGN para comparar sequências de aminoácidos, uma mesa de resíduo PAM de peso 120, um comprimento de lacuna de 12, e um desnível de 4 podem ser usados.

Salvo indicação em contrário, o GAP Versão 10, que utiliza o algoritmo de Needleman e Wunsch (1970) *J. Mol. Biol.* 48 (3): 443-453, será utilizado para determinar a identidade de sequência ou similaridade com os seguintes parâmetros: identidade % e similaridade % para uma sequência nucleotídica usando peso de GAP de 50 e Comprimento de 3, e a matriz de pontuação nwsgapdna.cmp; identidade % e similaridade % de uma sequência de aminoácidos utilizando peso de GAP de 8 e comprimento de 2, e o programa de escore BLOSUM62. Programas equivalentes também podem ser utilizados. Por "programa equivalente"

entende-se qualquer programa de comparação de sequência que, para quaisquer duas seqüências em questão, gera um alinhamento com resíduos de resíduos de nucleotídeos idênticos e uma identidade percentual de sequência

5 idêntica, quando comparada com o alinhamento correspondente gerado pelo GAP Versão 10. A invenção também engloba moléculas variantes de ácidos nucleicos (por exemplo, SEQ ID NO: 7, 9, 11, 16, 18, 22, 24, 26, 28, 30 e 32). "Variantes" de seqüências nucleotídicas de codificação de

10 endotoxina delta incluem as seqüências que codificam as proteínas endotoxina delta divulgadas aqui, mas que diferem conservativamente devido à degeneração do código genético, bem como aqueles que são suficientemente idênticos como discutido acima. Naturalmente variantes alélicos podem ser

15 identificados com o uso das bem conhecidas técnicas de biologia molecular, tais como a reação em cadeia da polimerase (PCR) e técnicas de hibridização como descrito abaixo. Seqüências nucleotídicas variantes também incluem seqüências nucleotídicas derivadas sinteticamente que foram

20 geradas, por exemplo, utilizando mutagênese sítio-dirigida, mas que ainda codificam proteínas endotoxina delta divulgadas na presente invenção como discutido abaixo. Proteínas variantes englobadas pela presente invenção são biologicamente ativas, isto é, elas continuam a possuir a

25 atividade biológica desejada da proteína nativa, ou seja, mantendo atividade pesticida. Por "manter atividade" entende-se a variante que terá pelo menos cerca de 30%, pelo menos cerca de 50%, pelo menos cerca de 70%, ou pelo menos cerca de 80% da atividade pesticida da proteína

30 nativa. Métodos para a medição da atividade pesticida são

bem conhecidos na arte. Ver, por exemplo, Czapla e Lang (1990) *J. Econ. Entomol.* 83: 2480-2485; Andrews et al. (1988) *Biochem. J.* 252: 199-206; Marrone et al. (1985) *J. of Economic Entomology* 78: 290-293; e Patente US No. 5 5.743.477, todos os quais são aqui incorporados por referência na sua totalidade.

O especialista na arte irá apreciar ainda que as mudanças podem ser introduzidas por mutação das sequências nucleotídicas da invenção, assim, levando a mudanças na sequência de aminoácidos das proteínas endotoxina delta codificadas, sem alterar a atividade biológica das proteínas. Assim, variantes de moléculas de ácido nucléico isolado podem ser criadas através da introdução de um ou mais substituições de nucleotídeos, acréscimos ou supressões na sequência nucleotídica correspondente divulgada aqui, de tal forma que uma ou mais substituições de aminoácidos, acréscimos ou supressões são introduzidas na proteína codificada. As mutações podem ser introduzidas pelas técnicas convencionais, tais como mutagênese sítio-dirigida e mutagênese mediada por PCR. Essas sequências nucleotídicas variantes também são abrangidas pela presente invenção.

Por exemplo, substituições de aminoácidos conservativas podem ser feitas em um ou mais resíduos de aminoácidos não essenciais previsíveis. Resíduo de aminoácidos "não essenciais" é um resíduo que pode ser alterado a partir da sequência do tipo selvagem de uma proteína endotoxina delta, sem alterar a atividade

biológica, enquanto que um resíduo de aminoácido "essencial" é necessário para a atividade biológica. Uma "substituição de aminoácido conservativa" é um aminoácido em que o resíduo é substituído por um resíduo de aminoácido com uma cadeia lateral semelhante. Famílias de resíduos de aminoácidos com cadeias laterais semelhantes foram definidas na arte. Essas famílias incluem aminoácidos com cadeias laterais básicas (por exemplo, lisina, arginina, histidina), cadeias laterais ácidas (por exemplo, ácido aspártico, ácido glutâmico), cadeias laterais polares não modificadas (por exemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadeias laterais não polares (por exemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano), cadeias laterais beta-ramificadas (por exemplo, treonina, valina, isoleucina) e cadeias laterais aromáticas (por exemplo, tirosina, fenilalanina, triptofano, histidina).

Endotoxinas delta geralmente têm cinco domínios de sequência conservados, e três domínios estruturais conservados (ver, por exemplo, de Maagd *et al.* (2001) *Trends Genetics* 17: 193-199). O primeiro domínio estrutural conservado consiste de sete alfa hélices e está envolvido na inserção da membrana e na formação dos poros. O domínio II é constituído por três beta-lâminas dispostas em uma configuração Greek key, e o domínio III é constituído por duas beta-lâminas antiparalelas em forma de "jelly-roll" (de Maagd *et al.*, 2001, *supra*). Os domínios II e III estão envolvidos no reconhecimento e ligação de receptor, e,

portanto, são considerados determinantes da especificidade da toxina.

Substituições de aminoácidos podem ser feitas em
5 regiões não conservadas que mantêm função. Em geral, essas substituições não seriam feitas para resíduos de aminoácido conservados, ou para resíduos de aminoácidos residindo dentro de um motivo conservado, onde esses resíduos são essenciais para a atividade da proteína. Exemplos de
10 resíduos que são conservados e que podem ser essenciais para a atividade da proteína incluem, por exemplo, resíduos que são idênticos entre todas as proteínas contidas em um alinhamento das sequências de aminoácidos da presente invenção e sequências endotoxina delta conhecidas. Exemplos
15 de resíduos que são conservados, mas que podem permitir substituições conservativas de aminoácido e manter atividade incluem, por exemplo, resíduos que têm apenas substituições conservativas entre todas as proteínas contidas em um alinhamento das sequências de aminoácidos da
20 presente invenção e sequências endotoxina delta conhecidas. No entanto, um habilitado na arte irá compreender que variantes funcionais podem ter menos alterações conservadas ou não conservadas nos resíduos conservados.

25 Alternativamente, sequências nucleotídicas variantes podem ser feitas através da introdução de mutações aleatoriamente ao longo de todo ou parte do código da sequência, tais como por mutagênese de saturação, e os mutantes resultantes podem ser rastreados para conferir
30 capacidade de atividade endotoxina delta para identificar

mutantes que mantenham a atividade. Após a mutagênese, as proteínas codificadas podem ser expressas recombinantemente e, a atividade da proteína pode ser determinada usando técnicas padrão de ensaio.

5

Utilizando métodos como PCR, hibridização, e as sequências endotoxina delta correspondentes podem ser identificadas, tais sequências tendo identidade substancial para as sequências da invenção. Ver, por exemplo, Sambrook
10 e Russell (2001) *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) e Innis, et al. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, NY).

15 Em um método de hibridização, o todo ou parte da sequência nucleotídica endotoxina delta pode ser utilizado para rastrear cDNA ou bibliotecas genômicas. Métodos para a construção de tais cDNA e bibliotecas genômicas são geralmente conhecidos na arte e são divulgados em Sambrook
20 e Russell, 2001, supra. As chamadas sondas de hibridização podem ser fragmentos de DNA genômico, fragmentos de cDNA, fragmentos de RNA, ou outros oligonucleotídeos, e podem ser marcadas com um grupo detectável, como ^{32}P , ou qualquer outro marcador detectável, como outros radioisótopos, um
25 composto fluorescente, uma enzima, ou um co-fator enzimático. Sondas de hibridização podem ser feitas por marcação de oligonucleotídeos sintéticos baseados nas sequências nucleotídicas que codificam a endotoxina delta divulgada aqui. Iniciadores degenerados concebidos com base
30 nos nucleotídeos conservados ou resíduos de aminoácidos na

sequência nucleotídica ou sequência codificada de aminoácidos podem também ser utilizados. A sonda tipicamente inclui uma região de sequência nucleotídica que hibridiza sob rigorosas condições de, pelo menos, cerca de 12, pelo menos cerca de 25, pelo menos cerca de 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, ou 400 nucleotídeos consecutivos de sequência nucleotídica de codificação da endotoxina delta da invenção ou um fragmento ou variante deste. Métodos para a preparação de sondas para hibridização são geralmente conhecidos na arte e são divulgados no Sambrook e Russell, 2001, supra, aqui incorporado por referência.

Por exemplo, uma sequência endotoxina delta divulgada aqui, ou uma ou mais partes dos mesmos, pode ser usada como uma sonda capaz de hibridizar especificamente correspondentes sequências endotoxina delta e RNAs mensageiros. Para atingir hibridização específica em uma variedade de condições, tais sondas incluem sequências que são únicas e são de preferência pelo menos cerca de 10 nucleotídeos de comprimento, ou pelo menos cerca de 20 nucleotídeos de comprimento. Essas sondas podem ser usadas para amplificar sequências endotoxina delta de um organismo escolhido por PCR. Esta técnica pode ser usada para isolar sequências codificantes adicionais de um organismo desejado ou como um teste diagnóstico para determinar a presença de sequências codificantes em um organismo. Técnicas de hibridização incluem triagem de hibridização de bibliotecas de DNA plaqueadas (ou placas ou colônias, ver, por exemplo, Sambrook et al. (1989) *Molecular Cloning: A Laboratory*

Manual (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

A hibridação de tais seqüências pode ser realizada sob
5 condições severas. Por "condições severas" ou "condições
severas de hibridização" entende-se as condições em que uma
sonda irá hibridizar a sua seqüência alvo para um grau
maior detectável do que de outras seqüências (por exemplo,
pelo menos 2 vezes acima). Condições severas são
10 dependentes da seqüência e serão diferentes em diferentes
circunstâncias. Ao controlar o rigor da hibridização e/ou
condições de lavagem, as seqüências alvo que são 100%
complementar da sonda podem ser identificadas (sondagem
homóloga). Alternativamente, condições severas podem ser
15 ajustadas para permitir algumas inadequações nas seqüências
para que menores graus de similaridade sejam detectados
(sondagem heteróloga). Geralmente, uma sonda é inferior a
cerca de 1000 nucleotídeos de comprimento, de preferência
com menos de 500 nucleotídeos de comprimento.

20

Normalmente, condições severas serão aquelas em que a
concentração de sal é inferior a cerca de 1,5 M de íon Na,
tipicamente cerca de 0,01 a 1,0 M de concentração de íons
Na (ou outros sais) em pH 7,0 a 8,3 e a temperatura é de
25 pelo menos cerca de 30°C para sondas curtas (por exemplo, 10
a 50 nucleotídeos) e pelo menos cerca de 60°C para sondas
longas (como por exemplo, superior a 50 nucleotídeos).
Condições severas também podem ser alcançadas com a adição
de agentes desestabilizadores, como a formamida. Condições
30 de baixa severidade exemplares incluem hibridização com uma

solução-tampão de 30 a 35% de formamida, 1 M de NaCl, 1% SDS (dodecil sulfato de sódio) a 37°C, e uma lavagem em 1X a 2X de SSC (20X de SSC = 3,0 M NaCl/0.3 M de citrato trissódico) em 50 a 55°C. Condições de moderada severidade
 5 exemplares incluem hibridização em 40 a 45% de formamida, 1,0 M de NaCl, 1% de SDS a 37°C, e uma lavagem em 0.5X a 1X de SSC em 55 a 60°C. Condições de elevada severidade exemplar incluem hibridização em 50% de formamida, 1 M de NaCl, 1% de SDS a 37°C, e uma lavagem em 0.1X de SSC em 60
 10 a 65°C. Opcionalmente, tampões de lavagem podem incluir cerca de 0,1% a cerca de 1% de SDS. A duração da hibridização é geralmente inferior a cerca de 24 horas, normalmente cerca de 4 a cerca de 12 horas.

15 Especificidade é normalmente a função de lavagens pós-hibridização, os fatores críticos sendo a força iônica e a temperatura da solução final de lavagem. Para híbridos DNA-DNA, a T_m pode ser aproximada a partir da equação de Meinkoth e Wahl (1984) *Anal. Biochem.* 138: 267-284:

20 $T_m = 81,5^{\circ}\text{C} + 16,6 (\log M) + 0,41 (\% \text{ de GC}) - 0,61 (\% \text{ form}) - 500/L$, onde M é a molaridade de cátions monovalentes, % de GC é o percentual dos nucleotídeos guanossina e citosina no DNA, % form é a percentagem de formamida na solução de hibridização, e L é o comprimento da base, em pares mistos.

25 A T_m é a temperatura (abaixo força iônica definida e pH) em que 50% de uma sequência alvo complementar hibridiza para uma sonda marcada perfeitamente. A T_m é reduzida em cerca de 1°C para cada 1% de inadequações; assim, T_m , hibridação e/ou condições de lavagem podem ser adaptadas para
 30 hibridizar sequências de identidade desejada. Por exemplo,

se seqüências com identidade $\geq$ a 90% são procuradas, a T_m pode ser diminuída 10°C. Geralmente, as condições severas são selecionadas para serem cerca de 5°C mais baixas do que o ponto de fusão térmico (T_m) para a sequência específica e seu complemento em uma força iônica definida e pH. No entanto, condições muito severas podem utilizar uma hibridação e/ou lavagem em 1, 2, 3 ou 4°C mais baixa do que o ponto de fusão térmico (T_m); condições moderadamente severas podem utilizar uma hibridização e/ou lavagem a 6, 7, 8, 9 ou 10°C menor do que o mesmo ponto de fusão (T_m); condições de baixa severidade podem utilizar uma hibridização e/ou lavagem em 11, 12, 13, 14, 15 ou 20°C inferior ao ponto de fusão térmico (T_m). Usando a equação, hibridização e composições de lavagem, e T_m desejada, um especialista na arte irá entender que as variações na severidade da hibridização e/ou soluções de lavagem são inerentemente descritas. Se o grau desejado de inadequações resulta em um T_m de menos de 45°C (solução aquosa) ou 32°C (solução de formamida), é preferível aumentar a concentração de SSC de modo que uma temperatura mais alta possa ser utilizada. Um extenso guia para a hibridização de ácidos nucleicos é encontrado em Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes*, Parte I, Capítulo 2, (Elsevier, Nova Iorque) e Ausubel et al., Eds. (1995) *Current Protocols in Molecular Biology*, Capítulo 2 (Greene Publishing and Wiley-Interscience, Nova Iorque). Ver Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nova Iorque).

Proteínas Isoladas e Variantes e Fragmentos Desta

Proteínas endotoxina delta também são englobadas no âmbito da presente invenção. Por "proteína endotoxina delta" entende-se uma proteína com sequência de aminoácidos estabelecida na SEQ ID NO: 2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38. Fragmentos, biologicamente ativos porções, e variantes, são também fornecidos, e pode ser utilizada para praticar os métodos da presente invenção.

"Fragmentos" ou "partes biologicamente ativas" incluem fragmentos de polipeptídeo compreendendo sequências de aminoácidos suficientemente idênticas à sequência de aminoácidos estabelecida em SEQ ID NO: 2, 4, 6, 8, 10, 12, 16, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38 e que apresentam atividade pesticida (por exemplo, SEQ ID NO: 15, 19, 21, 23 ou 25). Uma porção biologicamente ativa de uma proteína endotoxina delta pode ser um polipeptídeo que é, por exemplo, 10, 25, 50, 100 ou mais aminoácidos em comprimento. Essas porções biologicamente ativas podem ser preparadas por técnicas recombinantes e avaliadas para atividade pesticida. Métodos de medição da atividade pesticida são bem conhecidos na arte. Veja, por exemplo, Czapla e Lang (1990) *J. Econ. Entomol.* 83: 2480-2485; Andrews et al. (1988) *Biochem. J.* 252: 199-206; Marrone et al. (1985) *J. of Economic Entomology* 78: 290-293; e Patente US No. 5.743.477, todos aqui incorporados por referência, na sua totalidade. Conforme utilizado aqui, um fragmento compreende pelo menos 8 aminoácidos contíguos da SEQ ID NO:

2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38. A invenção engloba outros fragmentos, porém, tal como qualquer fragmento de proteína maior do que cerca de 10, 20, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, ou 1300 aminoácidos.

Por "variantes" é entendido polipeptídeos ou proteínas tendo uma sequência de aminoácidos que é, pelo menos, cerca de 60%, 65%, cerca de 70%, 75%, cerca de 80%, 85%, cerca de 90%, 91%, 92%, 93 %, 94%, 95%, 96%, 97%, 98% ou 99% idêntica à sequência de aminoácidos da SEQ ID NO: 2, 4, 6, 8, 10, 12, 16, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38 (por exemplo, SEQ ID NO: 10, 12, 17, 19, 23, 25, 27, 29, 31, ou 33). Variantes incluem também polipeptídeos codificados por uma molécula de ácido nucléico que hibridiza a molécula de ácido nucléico da SEQ ID NO: 1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, ou 37, ou um complemento das mesmas, sob condições estritas.

Variantes incluem polipeptídeos que diferem na sequência de aminoácidos devido a mutagênese. Proteínas variantes englobadas pela presente invenção são biologicamente ativas, isto é, elas continuam a possuir a atividade biológica desejada da proteína nativa, ou seja, mantendo atividade pesticida. Métodos para a medição da atividade pesticida são bem conhecidos na arte. Ver, por exemplo, Czapla e Lang (1990) *J. Econ. Entomol* 83: 2480-2485; Andrews et al. (1988) *Biochem. J.* 252: 199-206; Marrone et al. (1985) *J. of Economic Entomology* 78: 290-293; e Patente

US No. 5.743.477, os quais aqui incorporados por referência na sua totalidade.

Genes bacterianos, tais como os genes *axmi-031*, *axmi-039*, *axmi-040*, e *axmi-049* desta invenção, muitas vezes
5 possuem códons de iniciação de metionina múltiplos início na proximidade do início da abertura do molde de leitura. Muitas vezes, o início da tradução a um ou mais destes códons iniciadores levará à geração de uma proteína funcional. Estes códons iniciadores podem incluir códons
10 ATG. No entanto, bactérias tais como bactérias *Bacillus* sp. também reconhecem o códon GTG como um códon de início, e as proteínas que iniciam a tradução de códons GTG contêm uma metionina no primeiro aminoácido. Além disso, não é muitas vezes determinada *a priori* qual destes códons são
15 utilizados naturalmente na bactéria. Assim, é compreensível que a utilização de um dos códons de metionina alternativos possam também levar à geração de proteínas endotoxina delta que codificam atividade pesticida. Estas proteínas endotoxina delta são englobadas na presente invenção e
20 podem ser utilizadas nos métodos da presente invenção.

Anticorpos contra os polipeptídeos da presente invenção, ou as variantes ou fragmentos destes, também estão englobados. Métodos de produção de anticorpos são bem
25 conhecidos na arte (ver, por exemplo, Harlow e Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Patente US No. 4.196.265).

Variantes Alteradas ou Aperfeiçoadas

É reconhecido que sequências de DNA de uma endotoxina delta podem ser alteradas por vários métodos, e que estas alterações podem resultar em sequências de DNA codificando proteínas com sequências de aminoácidos diferentes do que codificadas por uma endotoxina delta da presente invenção. Esta proteína pode ser alterada de várias maneiras, incluindo substituições de aminoácidos, supressões, truncações, e inserções de um ou mais aminoácidos da SEQ ID NO: 2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, ou 38, incluindo-se cerca de 2, cerca de 3, cerca de 4, cerca de 5, cerca de 6, cerca de 7, cerca de 8, cerca de 9, cerca de 10, cerca de 15, cerca de 20, cerca de 25, cerca de 30, cerca de 35, cerca de 40, cerca de 45, cerca de 50, cerca de 55, cerca de 60, cerca de 65, cerca de 70, cerca de 75, cerca de 80, cerca de 85, cerca de 90, cerca de 100, cerca de 105, cerca de 110, cerca de 115, cerca de 120, cerca de 125, cerca de 130 ou mais substituições de aminoácidos, supressões ou inserções.

20

Métodos para tais manipulações são geralmente conhecidos na arte. Por exemplo, variantes de sequência de aminoácido uma proteína endotoxina delta podem ser preparadas por mutações no DNA. Isso também pode ser realizado por uma das várias formas de mutagênese e/ou em evolução dirigida. Em alguns aspectos, as alterações codificadas na sequência de aminoácidos não irão afetar substancialmente a função da proteína. Essas variantes vão possuir a desejada atividade pesticida. No entanto, entende-se que a capacidade de uma endotoxina delta para

30

conferir atividade pesticida pode ser melhorada através da utilização de tais técnicas sobre as composições da presente invenção. Por exemplo, um pode expressar uma endotoxina delta em células hospedeiras que apresentam

5 elevadas taxas de incorporação de base durante a replicação do DNA, como o XL-I Red (Stratagene). Depois de propagação nessas linhagens, pode-se isolar o DNA endotoxina delta (por exemplo, através da preparação do DNA plasmidial, ou por amplificação por PCR e clonagem do fragmento resultante

10 da PCR em um vetor), cultivar as mutações endotoxina delta em uma estirpe não-mutagênica, e identificar genes endotoxina delta mutados com atividade pesticida, por exemplo, realizando um ensaio para testar a atividade pesticida. Geralmente, a proteína é misturada e utilizada

15 nos ensaios de alimentação. Ver, por exemplo, Marrone *et al.* (1985) *J. of Economic Entomology* 78: 290-293. Tais testes podem incluir contatar plantas com uma ou mais pragas e determinar a capacidade da planta para sobreviver e/ou causar a morte dos parasitas. Exemplos de mutações que

20 resultam em aumento da toxicidade são encontrados em Schnepf *et al.* (1998) *Microbiol. Mol. Biol. Rev.* 62: 775-806.

Alternativamente, alterações podem ser feitas para a

25 sequência da proteína de muitas proteínas no amino ou carboxi terminal sem afetar sensivelmente a atividade. Isto pode incluir inserções, deleções, ou alterações introduzidas pelos modernos métodos moleculares, como PCR, incluindo a amplificação por PCR que alteram ou estendem a

30 sequência codificante da proteína em virtude da inclusão de

aminoácidos codificando sequências nos oligonucleotídeos utilizados na amplificação por PCR. Alternativamente, as sequências protéicas acrescentadas podem incluir toda a sequência de codificação de proteínas, como as usadas
5 comumente na arte de produzir fusões de proteína. Essas proteínas fusionadas são frequentemente utilizadas para (1) aumentar a expressão de uma proteína de interesse (2) introduzir um domínio de ligação, atividade enzimática, ou epítopo para facilitar a purificação de proteína, detecção
10 de proteína, ou outras utilizações experimentais conhecidas na arte (3) secreção alvo ou tradução de uma proteína de uma organela subcelular, tais como o espaço periplasmático de bactérias gram-negativas, ou do retículo endoplasmático das células eucarióticas, o último dos quais muitas vezes
15 resulta em glicosilação da proteína.

Nucleotídeos variantes e aminoácidos da presente invenção também englobam sequências derivadas de procedimentos mutagênicos e recombinogênicos tais como
20 embaralhamento de DNA. Com tal procedimento, uma ou mais proteínas endotoxina delta diferentes codificando regiões de podem ser usadas para criar uma nova proteína endotoxina delta que possua as propriedades desejadas. Desta forma, bibliotecas de polinucleotídeos recombinantes são geradas a
25 partir de uma população de polinucleotídeos de sequências relacionadas compreendendo regiões da sequência que têm uma identidade de sequência substancial e que podem ser homologamente recombinadas *in vitro* ou *in vivo*. Por exemplo, usando esta abordagem, motivos de sequência
30 codificando um domínio de interesse podem ser embaralhados

entre um gene endotoxina delta da invenção e outros genes endotoxina delta conhecidos para obter um novo gene que codifica para uma proteína com uma melhoria da propriedade de interesse, tais como um aumento da atividade inseticida.

- 5 Estratégias para tal embaralhamento de DNA são conhecidas na arte. Ver, por exemplo, Stemmer (1994) *Proc. Natl. Acad. Sci. E.U.A.* 91: 10747-10751; Stemmer (1994) *Nature* 370: 389-391; Cramer et al. (1997) *Nature Biotech.* 15: 436-438; Moore et al. (1997) *J. Mol. Biol.* 272: 336-347; Zhang et al. (1997) *Proc. Natl. Acad. Sci. E.U.A.* 94: 4504-4509; 10 Cramer et al. (1998) *Nature* 391: 288-291, e Patentes US nº 5.605.793 e 5.837.458.

- A troca ou embaralhamento de domínio é outro mecanismo 15 de geração de proteínas endotoxina delta alteradas. Os domínios II e III podem ser trocados entre proteínas endotoxina delta, resultando em toxinas híbridas ou quiméricas com melhor atividade pesticida ou espectro alvo. Métodos para a geração de proteínas recombinantes e teste 20 da atividade pesticida são bem conhecidos na arte (ver, por exemplo, Naimov et al. (2001) *Appl. Environ. Microbiol.* 67: 5328-5330; de Maagd et al. (1996) *Appl. Environ. Microbiol.* 62: 1537-1543, Ge et al. (1991) *J. Biol. Chem.* 266: 17954-17958; Schnepf et al. (1990), *J. Biol. Chem.* 265: 20923-20930; Rang et al. (1999) *Appl. Environ. Microbiol.* 65: 2918-2925).

Vetores

- A sequência endotoxina delta da invenção pode ser 30 feita em um cassete de expressão para a expressão em uma

planta de interesse. Por "cassete de expressão vegetal" entende-se uma construção de DNA que é capaz de resultar na expressão de uma proteína a partir de um molde de leitura aberta na célula vegetal. Tipicamente, estes contêm um
5 promotor e uma sequência codificante. Muitas vezes, essas construções também irão conter uma região 3' não traduzida. Essas construções podem conter uma "sequência sinal" ou "sequência líder" (ou seja, SEQ ID NO: 9, 11, 28, 30 e 32) para facilitar o transporte co-traducional ou pós-
10 traducional do peptídeo a certas estruturas intracelulares, tais como o cloroplasto (ou outros plastídeos), retículo endoplasmático, ou aparelho de Golgi.

Por "sequência sinal" entende-se uma sequência que é
15 conhecida ou suspeita para resultar em transporte de peptídeo co-traducional ou pós-traducional através da membrana celular. Em eucariotos, este normalmente envolve secreção no aparelho de Golgi, com alguns resultando em glicosilação. Por "sequência líder" entende-se qualquer
20 sequência que, quando traduzida, resulta em uma sequência de aminoácidos (ou seja, SEQ ID NO: 10, 12, 29, 31 e 33) suficiente para acionar o transporte co-traducional da cadeia peptídica para uma organela sub-celular. Assim, este inclui sequências líder objetivando transporte e/ou
25 glicosilação através da passagem para o retículo endoplasmático, passagem de vacúolos, plastídeos incluindo cloroplastos, mitocôndrias, e coisas do gênero.

Por "vetor de transformação de plantas" entende-se uma
30 molécula de DNA que é necessária para transformação

eficiente de uma célula vegetal. Essa molécula pode ser constituída por um ou mais cassetes de expressão vegetais, e pode ser organizada em mais de uma molécula de DNA "vetor". Por exemplo, vetores binários são vetores de

5 transformação de plantas que utilizam dois vetores de DNA não contíguos para codificar todas as funções e operações cis- de transformação de células vegetais (Hellens e Mullineaux (2000) *Trends in Plant Science* 5: 446-451). "Vetor" refere-se a uma construção de ácido nucléico

10 concebido para o transporte entre as diferentes células hospedeiras. "Vetor de expressão" refere-se a um vetor que tem a capacidade de incorporar, integrar e expressar sequências de DNA heterólogas ou fragmentos de uma célula externa. O cassete incluirá sequências reguladoras 5' e 3' operacionalmente ligadas a uma sequência da invenção. Por

15 "operacionalmente ligadas" entende-se uma ligação funcional entre um promotor e uma segunda sequência, onde a sequência do promotor inicia e medeia a transcrição da sequência de DNA correspondente à segunda sequência. Geralmente,

20 operacionalmente ligado significa que as sequências de ácido nucléico estando ligadas são contíguas e, quando necessário juntam duas regiões codificadoras de proteínas, contíguas e no mesmo molde de leitura. O cassete pode também conter, pelo menos, um gene adicional a ser co-

25 transformado para o organismo. Alternativamente, o(s) gene(s) adicional(s) pode ser fornecido em múltiplos cassetes de expressão.

"Promotor" refere-se a uma sequência de ácidos

30 nucléicos que funciona para dirigir a transcrição a jusante

de uma sequência codificante. O promotor, juntamente com outras sequências de ácidos nucléicos reguladoras transcricionais e traducionais (também denominadas "sequências controle") são necessários para a expressão de
5 uma sequência de DNA de interesse.

Tal cassete de expressão é fornecido com uma pluralidade de restrições locais para a inserção de a sequência endotoxina delta estar sob regulação
10 transcricional das regiões reguladoras.

O cassete de expressão vai incluir na direção 5'-3' de transcrição, uma região de iniciação transcricional e traducional (ou seja, um promotor), uma sequência de DNA da
15 invenção, e uma região terminal traducional e transcricional (ou seja, região de terminação) funcional nas plantas. O promotor pode ser nativo ou análogo, ou externo ou heterólogo, para a planta hospedeira e/ou para a sequência de DNA da invenção. Adicionalmente, o promotor
20 pode ser a sequência natural ou, alternativamente, uma sequência sintética. Onde o promotor é "nativo" ou "homólogo" para a planta hospedeira, pretende-se que o promotor seja encontrado na planta nativa em que o promotor é introduzido. Onde o promotor é "externo" ou "heterólogo"
25 para a sequência de DNA da invenção, pretende-se que o promotor não seja o nativo ou de ocorrência natural para a sequência de DNA operacionalmente ligada da invenção.

A região terminal pode ser nativa com a região de
30 iniciação transcricional, pode ser nativo com a sequência

de DNA de interesse operacionalmente ligada, pode ser nativa com a planta hospedeira, ou pode ser derivada de uma outra fonte (ou seja, externa ou heteróloga para o promotor, a sequência de DNA de interesse, a planta hospedeira, ou qualquer combinação das mesmas). Regiões de terminação convenientes encontram-se disponíveis a partir do plasmídeo-Ti de *A. tumefaciens*, tais como as regiões de terminação octopina sintase e nopalina sintase. Veja também Guerineau et al. (1991) *Mol Gen. Genet.* 262: 141-144; Proudfoot (1991) *Cell* 64: 671-674; Sanfacon et al. (1991) *Genes Dev.* 5: 141-149; Mogm et al. (1990) *Plant Cell* 2: 1261-1272; Munroe et al. (1990) *Gene* 91: 51-158; Ballas et al. (1989) *Nucleic Acids Res.* 17: 7891-7903; e Joshi et al. (1987) *Nucleic Acid Res.* 15: 9627-9639.

15

Quando apropriado, o(s) gene(s) pode ser otimizado para uma maior expressão na célula hospedeira transformada. Ou seja, os genes podem ser sintetizados utilizando códons de células hospedeiras preferenciais para uma melhor expressão, ou podem ser sintetizados utilizando códons em um códon hospedeiro preferido de frequência útil. Geralmente, o conteúdo GC do gene será aumentado. Ver, por exemplo, Campbell e Gowri (1990) *Plant Physiol.* 92: 1-11 para uma discussão de uso de códon hospedeiro preferido. Métodos estão disponíveis na arte para síntese de genes vegetais preferidos. Ver, por exemplo, Patente US nº 5.380.831 e 5.436.391, e Murray et al. (1989) *Nucleic Acids Res.* 17: 477-498, aqui incorporados por referência.

25

Em um experimento, a endotoxina delta é orientada para o cloroplasto para expressão. Desta forma, quando a endotoxina delta não está diretamente inserida no cloroplasto, o cassete de expressão irá conter
5 adicionalmente um ácido nucléico codificando um peptídeo de trânsito para orientar a endotoxina delta aos cloroplastos. Esses peptídeos de trânsito são conhecidos na arte. Ver, por exemplo, Von Heijne et al. (1991) *Plant Mol. Biol. Rep.* 9: 104-126; Clark et al. (1989) *J. Biol. Chem.* 264: 17544-
10 17550; Della-Cioppa et al. (1987) *Plant Physiol.* 84:965-968; Romer et al. (1993) *Biochem. Biophys. Res. Commun.* 196: 1414-1421; e Shah et al. (1986) *Science* 233: 478-481.

O gene endotoxina delta a ser orientado para a
15 cloroplasto pode ser otimizado para a expressão em cloroplasto para levar em conta as diferenças do códon utilizado entre o núcleo da planta e esta organela. Desta forma, os ácidos nucléicos de interesse podem ser sintetizados utilizando códons de cloroplastos preferidos.
20 Ver, por exemplo, Patente US No. 5.380.831, aqui incorporada por referência.

Transformação de Plantas

Métodos da invenção envolvem a introdução de uma
25 construção de nucleotídeo em uma planta. Por "introdução" entende-se apresentar à planta a construção de nucleotídeo de tal forma que o constructo tenha acesso ao interior de uma célula da planta. Os métodos da invenção não exigem que um determinado método para a introdução de uma construção
30 de nucleotídeo em uma planta seja utilizado, só que a

construção de nucleotídeo ganhe acesso ao interior de, pelo menos, uma célula da planta. Métodos para introduzir construções de nucleotídeos em plantas são conhecidos na arte, incluindo, mas não se limitando a, métodos de
5 transformação estável, métodos de transformação transientes, e métodos mediados por vírus.

Por "planta" entende-se a planta inteira, órgãos vegetais (por exemplo, folhas, caules, raízes, etc.),
10 sementes, células vegetais, propágulos, embriões e descendentes do mesmo. Células vegetais podem ser diferenciadas ou indiferenciadas (por exemplo, calos, cultura de células em suspensão, protoplastos, células foliares, células de raiz, células de floema, pólen).

15 "Plantas transgênicas" ou "plantas transformadas" ou plantas "estavelmente transformadas" ou células ou tecidos se referem às plantas que têm incorporado ou integrado sequências de ácido nucléico exógena ou fragmentos de DNA na célula vegetal. Estas sequências de ácidos nucléicos
20 incluem as que são exógenas, ou não presentes na célula vegetal não transformada, bem como aquelas que podem ser endógenas, ou presentes na célula vegetal não transformada. "Heterólogas" geralmente se refere às sequências de ácido nucléico que não são endógenas para a célula ou parte do
25 genoma nativo em que estão presentes, e foram adicionados à célula por infecção, transfecção, microinjeção, eletroporação, microprojeção, ou coisas do gênero.

Transformação de células vegetais pode ser realizada
30 por uma das várias técnicas conhecidas na arte. O gene

endotoxina delta da invenção pode ser modificado para obter ou melhorar a expressão em células vegetais. Normalmente, um construto que expressa uma proteína irá conter um promotor para conduzir uma transcrição do gene, assim como
5 uma região 3' não traduzida para permitir o término da transcrição e poliadenilação. A organização de tais construções é bem conhecida na arte. Em alguns casos, pode ser útil para a engenharia o gene, tal que o peptídeo resultante seja secretado, ou outra forma segmentada dentro
10 da célula vegetal. Por exemplo, o gene pode ser projetado para conter um peptídeo sinal para facilitar a transferência do peptídeo para o retículo endoplasmático. Também pode ser preferível a engenharia do cassete de expressão vegetal para conter um íntron, de tal forma que o
15 processamento do mRNA no íntron seja necessário para a expressão.

Normalmente este "cassete de expressão vegetal" será inserido em um "vetor de transformação vegetal". Este vetor
20 de transformação vegetal pode ser composto apenas por um ou mais vetores de DNA necessários para a realização da transformação de plantas. Por exemplo, é uma prática comum na arte utilizar vetores de transformação vegetal que são compostos por mais de um segmento de DNA contíguos. Esses
25 vetores são muitas vezes referidos na arte como "vetores binários". Vetores binários, bem como vetores com plasmídios de socorro são mais frequentemente utilizados para transformação mediada por *Agrobacterium*, onde a dimensão e complexidade dos segmentos de DNA necessários
30 para alcançar a eficiente transformação é muito grande, e é

vantajoso para separar funções nas moléculas de DNA separadas. Vetores binários normalmente contêm um vetor plasmidial que contém sequências de ação cis exigidas para a transferência do T-DNA (tal como borda esquerda e borda
5 direita), um marcador selecionável que está projetado para ser capaz de expressão em uma célula vegetal, e um "gene de interesse" (um gene projetado para ser capaz de expressão em uma célula vegetal para geração de plantas transgênicas é desejado). Também presentes neste vetor plasmidial estão
10 sequências necessárias para a replicação bacteriana. As sequências de ação cis- são dispostas em uma forma eficiente para permitir a transferência em células vegetais e expressão. Por exemplo, o gene marcador selecionável e a endotoxina delta estão localizados entre a borda esquerda e
15 a direita. Muitas vezes um segundo vetor plasmidial contém os fatores de ação trans- que medeiam a transferência de T-DNA de *Agrobacterium* para as células vegetais. Este plasmídeo frequentemente contém as funções de virulência (genes Vir) que permitem a infecção das células vegetais
20 por *Agrobacterium*, e a transferência de DNA pela clivagem das sequências das bordas e transferência de DNA mediada por vir, como é entendida na arte (Hellens e Mullineaux (2000) *Trends in Plant Science* 5: 446-451). Vários tipos de cepas de *Agrobacterium* (por exemplo, LBA4404, GV3101,
25 EHA101, EHA105, etc.) podem ser utilizados para transformação de plantas. O segundo vetor plasmidial não é necessário para transformar as plantas através de outros métodos, tais como microprojeção, microinjeção, eletroporação, polietileno glicol, etc.

Em geral, métodos de transformação vegetal envolvem transferência de DNA heterólogo em células vegetais alvo (por exemplo, embriões imaturos ou maduros, culturas em suspensão, calos indiferenciados, protoplastos, etc.), seguida pela aplicação de um limiar máximo de nível de seleção adequado (em função do gene marcador selecionável) para recuperar as células vegetais transformadas a partir de um grupo de massa de células não transformadas. Explantes normalmente são transferidos para um suprimento fresco do mesmo meio e cultivados rotineiramente. Posteriormente, as células transformadas são diferenciadas em brotações após a colocação no meio de regeneração suplementado com um limite máximo de nível de agente de seleção. Os brotos são então transferidos para um meio seletivo de enraizamento para recuperar a raiz ou broto. Os brotos transgênicos então crescem em uma planta madura e produzem sementes férteis (por exemplo, Hiei et al. (1994) *The Plant Journal* 6: 271-282; Ishida et al. (1996) *Nature Biotechnology* 14: 745-750). Explantes normalmente são transferidos para um suprimento fresco de um mesmo meio e cultivados rotineiramente. Uma descrição geral das técnicas e métodos para a geração de plantas transgênicas é encontrada em Ayres e Park (1994) *Critical Reviews in Plant Science* 13: 219-239 e Bommineni e Jauhar (1997) *Maydica* 42: 107-120. Uma vez que o material transformado contém muitas células, ambas as células transformadas e não transformadas estão presentes em qualquer parte do calo objeto alvo ou tecido ou grupo de células. A capacidade de matar células não transformadas e permitir células transformadas de multiplicar resulta em culturas de plantas transformadas.

Muitas vezes, a capacidade de eliminar células não transformadas é uma limitação a uma cobertura rápida de células vegetais transformadas e de geração bem sucedida de plantas transgênicas.

5

Protocolos de transformação, bem como protocolos para a introdução de sequências nucleotídicas em plantas podem variar dependendo do tipo de planta ou de células vegetais, ou seja, monocotiledôneas ou dicotiledôneas, orientadas para a transformação. A geração de plantas transgênicas
10 pode ser realizada por um de vários métodos, incluindo, mas não limitados a, microinjeção, eletroporação, transferência direta de genes, introdução de DNA heterólogo por *Agrobacterium* em células vegetais (transformação mediada
15 por *Agrobacterium*), bombardeamento de células vegetais com DNA externo heterólogo aderido às partículas, aceleração de partícula balística, transformação por feixe de aerossóis (Pedido US Publicado No. 20010026941; Patente US No. 4.945.050; Internacional Publication No. WO 91/00915,
20 Pedido US Publicado No. 2002015066), transformação Led, e diversos outros métodos mediados diretamente por não partículas de transferência de DNA.

Métodos de transformação de cloroplastos são
25 conhecidos na arte. Ver, por exemplo, Svab et al. (1990) *Proc. Natl. Acad. Sci USA* 87: 8526-8530; Svab e Maliga (1993) *Proc. Natl. Acad. Sci USA* 90: 913-917; Svab e Maliga (1993) *EMBO J.* 12: 601-606. O método baseia-se na entrega de partícula de DNA que contém um marcador selecionável e
30 segmentação do DNA para o genoma plastidial através de

recombinação homóloga. Além disso, a transformação plastidial pode ser realizada por transativação de um transgene plastidial silencioso através da expressão do tecido preferido do núcleo codificado e da RNA polimerase dirigida pelo plastídio. Esse sistema tem sido relatado em McBride *et al.* (1994) *Proc. Natl. Acad. Sci USA* 91: 7301-7305.

Seguindo a integração do DNA externo heterólogo em células vegetais, aplica-se um limite máximo de nível de seleção adequado a meio para matar as células não transformadas e separar e proliferar as células putativamente transformadas que sobrevivem a este tratamento de seleção regularmente pela transferência regular para um meio fresco. Por passagem e troca contínua com a seleção adequada, uma identifica e prolifera as células que são transformadas com o vetor plasmidial. Métodos moleculares e bioquímicos podem então ser utilizados para confirmar a presença do gene heterólogo integrado de interesse para o genoma da planta transgênica.

As células que foram transformadas podem ser cultivadas em plantas de acordo com as formas convencionais. Ver, por exemplo, McCormick *et al.* (1986) *Plant Cell Reports* 5: 81-84. Estas plantas podem então ser cultivadas, e polinizadas com a mesma linhagem transformada ou linhagens diferentes, e o híbrido resultante tendo expressão constitutiva da característica fenotípica pretendida identificada. Duas ou mais gerações podem ser cultivadas para garantir que a expressão das

características fenotípicas desejadas sejam mantidas estáveis e herdáveis e, em seguida, as sementes colhidas a fim de garantir que a expressão das características fenotípicas pretendidas tenha sido alcançada. Desta forma, a presente invenção prevê sementes transformadas (também referidas como "sementes transgênicas"), com um constructo de nucleotídeo do invento, por exemplo, um cassete de expressão da invenção, estavelmente incorporado em seu genoma.

Avaliação da Transformação Vegetal

Após a introdução do DNA heterólogo externo em células vegetais, a transformação ou integração de genes heterólogos no genoma vegetal é confirmada por vários métodos, como a análise de ácidos nucleicos, proteínas e metabólitos associados com o gene integrado.

Análise por PCR é um método rápido para escanear células transformadas, tecidos ou ramo para a presença de genes incorporados na fase anterior, antes de transplantar para o solo (Sambrook e Russell (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). A PCR é realizada utilizando iniciadores oligonucleotídeos específicos para o gene de interesse ou vetor *Agrobacterium*, etc.

A transformação vegetal pode ser confirmada por análise de Southern blot do DNA genômico (Sambrook e Russell, 2001, supra). Em geral, o DNA total é extraído do transformante, digerido com enzimas de restrição

apropriadas, fracionado em um gel de agarose e transferido para uma membrana de nitrocelulose ou nylon. A membrana ou "blot" é, então, analisada, por exemplo, com fragmento de DNA alvo radiomarcado com ^{32}P para confirmar a integração do gene introduzido no genoma vegetal, de acordo com as técnicas padrão (Sambrook e Russell, 2001, supra).

Na análise de Northern blot, o RNA é isolado a partir de tecidos específicos de transformantes, fracionado em um gel de agarose de formaldeído, e repassado para um filtro de nylon, de acordo com procedimentos padrão que são rotineiramente utilizados na arte (Sambrook e Russell, 2001, supra). A expressão de RNA codificada pela endotoxina delta é então testada por hibridização do filtro para uma sonda radioativa proveniente de uma endotoxina delta, através de métodos conhecidos na arte (Sambrook e Russell, 2001, supra).

Western Blot, ensaios bioquímicos e semelhantes podem ser realizados sobre as plantas transgênicas para confirmar a presença da proteína codificada pelo gene endotoxina delta por procedimentos padrão (Sambrook e Russell, 2001, supra), utilizando anticorpos que se ligam a um ou mais epítomos presentes na proteína endotoxina delta.

Atividade Pesticida em Plantas

Em outro aspecto da invenção, um pode gerar plantas transgênicas expressando uma endotoxina delta que tenha atividade pesticida. Métodos descritos acima, a título de exemplo podem ser utilizados para gerar plantas

transgênicas, mas a maneira em que as células da planta transgênica são geradas não é crítica para esta invenção. Métodos conhecidos ou descritos na arte, como transformação mediada por *Agrobacterium*, transformação por biobalística, e métodos mediados por não-partículas podem ser utilizados a critério do experimentador. Plantas expressando uma endotoxina delta podem ser isoladas por métodos comuns descritos na arte, por exemplo, por transformação de calos, seleção de calos transformados, e regeneração de plantas férteis de tais calos transgênicos. Nesse processo, pode utilizar qualquer gene como um marcador selecionável tão longo quanto a sua expressão nas células vegetais conferindo a capacidade de identificar e selecionar para células transformadas.

Um certo número de marcadores tem sido desenvolvido para uso com células vegetais, tais como resistência ao cloranfenicol, a aminoglicosídeos G418, higromicina, ou coisas do gênero. Outros genes que codificam um produto envolvido no metabolismo do cloroplasto também podem ser utilizados como marcadores selecionáveis. Por exemplo, genes que oferecem resistência a herbicidas como o glifosato, bromoxinil, ou imidazolinona podem encontrar uso particular. Esses genes têm sido relatados (Stalker *et al.* (1985) *J. Biol. Chem.* 263: 6310-6314 (gene nitrilase com resistência ao bromoxinil) e Sathasivan *et al.* (1990) *Nucl. Acids Res.* 18: 2188 (gene com resistência a imidazolinona AHAS). Adicionalmente, os genes divulgados aqui são úteis como indicadores para avaliar a transformação de bactérias ou células vegetais. Métodos para detectar a presença de um

transgene em uma planta, órgão vegetal (por exemplo, folhas, caules, raízes, etc.), sementes, células vegetais, propágulo, embrião ou descendentes dos mesmos são bem conhecidos na arte. Em um experimento, a presença do
5 transgene é detectada pelo teste da atividade pesticida.

Plantas férteis expressando uma endotoxina delta podem ser testadas para atividade pesticida, e as plantas que apresentam ótima atividade selecionadas para posterior
10 reprodução. Métodos estão disponíveis na arte para ensaio da atividade de pragas. Geralmente, a proteína é misturada e utilizada nos ensaios de alimentação. Ver, por exemplo Marrone et al. (1985) *J. of Economic Entomology* 78: 290-293.

15

A presente invenção pode ser utilizada para transformação de qualquer espécie vegetal, incluindo mas não limitada a, monocotiledôneas e dicotiledôneas. Exemplos de plantas de interesse incluem, mas não estão limitados a,
20 cereais (milho), sorgo, trigo, girassol, tomate, crucíferas, pimentão, batata, algodão, arroz, soja, sacarina de beterraba, cana-de-açúcar, tabaco, cevada e colza, *Brassica* sp., alfafa, centeio, gramínea, cártamo, amendoim, batata doce, mandioca, café, coco, abacaxi,
25 citros, cacau, chá, banana, abacate, figo, goiaba, manga, azeitona, mamão, caju, macadâmia, amêndoa, aveia, legumes, plantas ornamentais, e coníferas.

Legumes incluem, mas não estão limitados a, tomate,
30 alface, feijão verde, feijões, ervilhas, e os membros do

gênero *Curcumis* tais como pepino, melão, melancia e almíscar. Plantas ornamentais incluem, mas não estão limitadas a, azaléia, hortências, hibisco, rosas, tulipas, narcisos, petúnias, cravo, poinsettia e crisântemo.

5 Preferencialmente, as plantas da presente invenção são espécies vegetais (por exemplo, milho, sorgo, trigo, girassol, tomate, crucíferas, pimentão, batata, algodão, arroz, soja, sacarina de beterraba, cana-de-açúcar, tabaco, cevada, colza, etc.).

10

Utilização em Controle de Peste

Métodos gerais para empregar linhagens compreendendo uma sequência de nucleotídeo da presente invenção, ou uma variante da mesma, em controle pesticida ou em engenharia
15 de outros organismos como agentes pesticidas são conhecidos na arte. Ver, por exemplo, Patente US No. 5.039.523 e EP 0480762A2.

As linhagens de *Bacillus* contendo uma sequência de
20 nucleotídeo da presente invenção, ou uma variante, ou os microorganismos que foram geneticamente alterados para conter um gene pesticida e proteína podem ser utilizados para proteger culturas agrícolas e produtos de pragas. Em um aspecto da invenção, células inteiras, ou seja, não
25 lisadas, de um organismo produzindo uma toxina (pesticidas) são tratadas com reagentes que prolongam a atividade da toxina produzida na célula quando a célula é aplicada para o ambiente das pragas-alvo(s).

Alternativamente, o pesticida é produzido através da introdução de um gene endotoxina delta em um hospedeiro celular. Expressão do gene endotoxina delta resulta, direta ou indiretamente, na produção e manutenção intracelular do pesticida. Em um aspecto da invenção, estas células são, em seguida, tratadas em condições que prolongam a atividade da toxina produzida na célula quando a célula é aplicada no ambiente das pragas-alvo. O produto resultante mantém a toxicidade da toxina. Estes pesticidas naturalmente encapsulados podem então ser formulados de acordo com técnicas convencionais de aplicação no ambiente das pragas hospedeiras alvo, por exemplo, solo, água e folhagem das plantas. Ver, por exemplo EPA 0192319, e as referências aí citadas. Alternativamente, pode formular as células para expressar o gene da presente invenção, como para permitir a aplicação do material resultante como um pesticida.

Controle de nematóide

Nematóides parasitas de plantas, incluindo cisto de nematóides, nematóides de nó de raiz, e outros nematóides, causam bilhões de dólares em danos para as culturas anuais. Nematóides parasitas de plantas furam paredes de células vegetais com seu estilete, que é formado por algumas das partes da boca e esôfago, e uma bomba até a célula vegetal em seu sistema digestivo. Numerosas publicações na arte concluem que nematóides parasitas de plantas sedentários não ingerem moléculas maiores que cerca de 23-28 kDa in vivo. Isto levou à crença generalizada de que o tubo de alimentação produzido por estes nematóides age como uma peneira molecular, restringindo o tamanho das moléculas que

podem ser ingeridas. Böckenhoff e Grundler ((1994) *Parasitology* 109: 249-254) injetaram dextranos fluorescentemente etiquetados em um sincício que estava sendo alimentado por *Heterodera schachtii*. Eles descobriram
5 que os nematóides ingeriram 22 kDa de dextranos, mas não 40 kDa de dextranos. Urwin et al. ((1998) *Plant* 204:472-479) descobriram que uma proteína de fusão de 23 kDa expressa em *Arabidopsis* transgênica não pode ser ingerida por *H. schachtii*, apesar de uma proteína de cerca de 11 kDa poder
10 ser ingerida. Do mesmo modo, Unwin et al. ((1997) *Molecular Plant-Microbe Interactions* 10: 394-400) constataram que uma proteína verde fluorescente (GFP) de 28 kDa, expressa em *Arabidopsis* transgênica não pode ser ingerida por *H. schachtii*. Assim, não foi anteriormente reconhecido ou
15 demonstrado que nematóides podem ser controlados pela expressão de uma proteína maior do que 22 kDa em uma planta, uma vez que foi entendido que nematóides não ingerem essa proteína.

20 São fornecidos aqui métodos e composições para conferir resistência à nematóide em plantas. Os métodos incluem introdução, em uma planta de pelo menos uma sequência nucleotídica que codifica um polipeptídeo nematóide-ativo e cultivo da planta em um campo contendo
25 nematóides. Por "polipeptídeo nematóide-ativo" entende-se um polipeptídeo que, quando ingerido pelo nematóide, resulta na morte ou nanismo de crescimento ou proliferação de pelo menos um nematóide. Em um experimento, o polipeptídeo nematóide-ativo tem um peso molecular superior
30 a cerca de 22 kDa, cerca de 23 kDa, cerca de 24, cerca de

25, cerca de 26, cerca de 27, cerca de 28, cerca de 29, cerca de 30, cerca de 31, cerca de 32, cerca de 33, cerca de 34, cerca de 35, cerca de 36, cerca de 37, cerca de 38, cerca de 39, cerca de 40, cerca de 41, cerca de 42, cerca de 43, cerca de 44, cerca de 45, cerca de 46, cerca de 47, cerca de 48, cerca de 49 , cerca de 50, cerca de 51, cerca de 52, cerca de 53, cerca de 54, cerca de 55, cerca de 56, cerca de 57, cerca de 58, cerca de 59, cerca de 60, cerca de 61, cerca de 62, cerca de 63, cerca de 64, cerca de 65, cerca de 66, cerca de 67, cerca de 68, cerca de 69, cerca de 70, cerca de 71, cerca de 72, cerca de 73, cerca de 74, cerca de 75, cerca de 76, cerca de 77, cerca de 78, cerca de 79, cerca de 80 kDa, ou maior.

Em outra incorporação, o polipeptídeo nematóide-ativo tem atividade contra nematóides parasitas de plantas. A expressão do polipeptídeo nematóide-ativo em uma planta reduz a capacidade dos nematóides para infestar ou alimentar raízes da planta. Exemplos de nematóides parasitas de plantas sensíveis às composições da presente invenção incluem nematóides de nó de raiz (*Meloidogyne* spp.), nematóide dublê (*Tylenchorhynchus* sp.), nematóide lança (*Hoplolaimus* sp.), nematóide em espiral (*Helicotylenchus* sp.), nematóide de lesão (*Pratylenchus* sp.), cisto de nematóide (*Heterodera* sp. e *Globodera* sp.), e nematóide anel (*Criconea* sp.).

Composições pesticidas

Os ingredientes ativos da presente invenção são normalmente aplicados na forma de composições e podem ser

aplicados para a área ou planta a ser tratada, simultaneamente ou em sucessão, com outros compostos. Estes compostos podem ser adubos, matadores de ervas daninhas, crioprotetores, tensoativos, detergentes, sabonetes
5 pesticidas, óleos latentes, polímeros, e/ou formulações transportadoras biodegradáveis e com tempo de liberação que permitam dosagem a longo prazo de uma área-alvo seguida de uma única aplicação da formulação. Eles também podem ser herbicidas seletivos, inseticidas químicos, virucidas,
10 microbicidas, amebicidas, pesticidas, fungicidas, bacteriocidas, nematocidas, moluscicidas ou misturas de várias destas preparações, se o desejar, juntamente com outras veículos agronomicamente aceitáveis, tensoativos ou adjuvantes de promoção de aplicação habitualmente
15 empregados na arte da formulação. Veículos adequados e adjuvantes podem ser sólidos ou líquidos e correspondem às substâncias comumente empregadas na tecnologia da formulação, por exemplo, substâncias minerais naturais ou regeneradas, solventes, dispersantes, agentes molhantes,
20 tackifiers, adesivos ou fertilizantes. As formulações também podem ser preparadas em "iscas" comestíveis ou em "armadilhas" de pragas para permitir a alimentação ou ingestão por uma praga alvo da formulação pesticida.

25 Métodos de aplicação de um ingrediente ativo da presente invenção ou uma composição agroquímica da presente invenção que contêm pelo menos uma das proteínas pesticidas produzidas pelas cepas bacterianas da presente invenção incluem aplicação foliar, aplicação nas sementes do solo e
30 aplicação de revestimento. O número de aplicações e a taxa

de aplicação dependem da intensidade de infestação pela praga correspondente.

A composição pode ser formulada como um pó, poeira, 5 sedimento, grânulo, spray, emulsão, colóide, solução, ou afins, e pode ser preparada por meios convencionais, tais como desidratação, liofilização, homogeneização, extração, filtração, centrifugação, decantação, ou concentração de uma cultura de células compreendendo o polipeptídeo. Em 10 todas essas composições que contenham pelo menos um desses polipeptídeos pesticidas, o polipeptídeo pode estar presente em uma concentração de cerca de 1% para cerca de 99% em peso.

15 Parasitas lepidópteros, coleópteros, ou nematóides podem ser mortos ou reduzidos em número em uma dada área através dos métodos da invenção, ou pode ser aplicada profilaticamente a uma área ambiental para evitar a infestação de pragas suscetíveis. De preferência a praga 20 ingere, ou é contatada com uma quantidade pesticidamente efetiva do polipeptídeo. Por "quantidade pesticidamente efetiva" entende-se uma quantidade de pesticida que é capaz de levar a morte a pelo menos uma praga, ou reduzir visivelmente o crescimento de pragas, alimentação, ou 25 desenvolvimento fisiológico normal. Essa quantidade irá variar dependendo de fatores como, por exemplo, as pragas-alvo específicas a serem controladas, o ambiente específico, localização, planta, cultura, ou local agrícola a ser tratado, as condições ambientais, bem como o método, 30 ritmo, concentração, estabilidade, e quantidade de

aplicação da composição do polipeptídeo pesticidamente efetivo. As formulações também podem variar em relação às condições climáticas, ambientais, e/ou a frequência de aplicação e/ou severidade da infestação por pragas.

5

As composições pesticidas descritas podem ser feitas pela formulação tanto pela célula bacteriana, cristal e/ou suspensão de esporos, ou componente protéicos isolados com o veículo agronomicamente aceitável desejado. As composições podem ser formuladas antes da administração, em um meio adequado, como liofilizado, congelado a seco, desidratado, ou em um veículo aquoso, meio ou diluente adequado, como o soro fisiológico ou outros tampões. As composições podem ser formuladas sob a forma de um pó granular ou material, ou uma suspensão em óleo (vegetal ou mineral), ou água ou emulsões óleo/água, ou como um pó molhável, ou em combinação com qualquer outro material veículo adequado para aplicação agrícola. Veículos agrícolas adequados podem ser sólidos ou líquidos e são bem conhecidos na arte. O termo "veículo agronomicamente aceitável" abrange todos os adjuvantes, componentes inertes, dispersantes, agentes tensoativos, tackifiers, capas, etc., que são normalmente utilizados na tecnologia de formulação pesticida; estes são bem conhecidos para aqueles especialistas na formulação pesticida. As formulações podem ser misturadas com um ou mais adjuvantes sólidos ou líquidos e preparadas por diferentes meios, por exemplo, por misturar homogeneamente, misturar e/ou triturar a composição pesticidal com adjuvantes adequados utilizando técnicas convencionais de formulação.

30

Formulações adequadas e métodos de aplicação são descritos na Patente US No. 6.468.523, aqui incorporada por referência.

- 5 "Pragas" incluem, mas não está limitado a, insetos, fungos, bactérias, nematóides, ácaros, carrapatos, e coisas do gênero. Insetos pragas incluem insetos selecionados a partir das ordens *Coleoptera*, *Diptera*, *Hymenoptera*, *Lepidoptera*, *Mallophaga*, *Homoptera*, *Hemiptera*, *Orthoptera*,
 10 *Thysanoptera*, *Dermaptera*, *Isoptera*, *Anoplura*, *Siphonaptera*, *Trichoptera*, etc, especialmente *Coleoptera*, *Lepidoptera* e *Diptera*.

- A ordem *Coleoptera* inclui as subordens *Adephaga* e
 15 *Polyphaga*. A subordem *Adephaga* inclui as superfamílias *Caraboidea* e *Gyrinoidea*, enquanto a subordem *Polyphaga* inclui as superfamílias *Hydrophiloidea*, *Staphylinoidea*, *Cantharoidea*, *Cleroidea*, *Elateroidea*, *Dascilloidea*, *Dryopoidea*, *Byrrhoidea*, *Cucujoidea*, *Meloidea*, *Mordelloidea*,
 20 *Tenebrionoidea*, *Bostrichoidea*, *Scarabaeoidea*, *Cerambycoidea*, *Chrysomeloidea*, e *Curculionoidea*. A superfamília *Caraboidea* compreende as famílias *Cicindelidae*, *Carabidae* e *Dytiscidae*. A superfamília *Gyrinoidea* inclui a família *Gyrinidae*. A superfamília
 25 *Hydrophiloidea* inclui a família *Hydrophilidae*. A superfamília *Staphylinoidea* inclui as famílias *Silphidae* e *Staphylinidae*. A superfamília *Cantharoidea* inclui as famílias *Cantharidae* e *Lampyridae*. A superfamília *Cleroidea* inclui as famílias *Cleridae* e *Dermestidae*. A superfamília
 30 *Elateroidea* inclui as famílias *Elateridae* e *Buprestidae*. A

superfamília *Cucujoidea* inclui a família *Coccinellidae*. A superfamília *Meloidea* inclui a família *Meloidae*. A superfamília *Tenebrionoidea* inclui a família *Tenebrionidae*. A superfamília *Scarabaeoidea* inclui as famílias *Passalidae* e *Scarabaeidae*. A superfamília *Cerambycoidea* inclui a família *Cerambycidae*. A superfamília *Chrysomeloidea* inclui a família *Chrysomelidae*. A superfamília *Curculionoidea* inclui as famílias *Curculionidae* e *Scolytidae*.

10 A ordem *Diptera* inclui as subordens *Nematocera*, *Brachycera*, e *Cyclorrhapha*. A subordem *Nematocera* inclui as famílias *Tipulidae*, *Psychodidae*, *Culicidae*, *Ceratopogonidae*, *Chironomidae*, *Simuliidae*, *Bibionidae*, e *Cecidomyiidae*. A subordem *Brachycera* inclui as famílias
15 *Stratiomyidae*, *Tabanidae*, *Therevidae*, *Asilidae*, *Mydidae*, *Bombyliidae*, e *Dolichopodidae*. A subordem *Cyclorrhapha* inclui as divisões *Aschiza* e *Aschiza*. A divisão *Aschiza* inclui a famílias *Phoridae*, *Syrphidae* e *Conopidae*. A divisão *Aschiza* inclui as seções *Acalyptratae* e
20 *Calypttratae*. A seção *Acalyptratae* inclui as famílias *Otitidae*, *Tephritidae*, *Agromyzidae*, e *Drosophilidae*. A seção *Calypttratae* inclui as famílias *Hippoboscidae*, *Oestridae*, *Tachinidae*, *Anthomyiidae*, *Muscidae*, *Calliphoridae* e *Sarcophagidae*.

25

A ordem *Lepidoptera* compreende as famílias *Papilionidae*, *Pieridae*, *Lycaenidae*, *Nymphalidae*, *Danaidae*, *Satyridae*, *Hesperiidae*, *Sphingidae*, *Saturniidae*, *Geometridae*, *Arctiidae*, *Noctuidae*, *Lymantriidae*, *Sesiidae*,
30 e *Tineidae*.

Os insetos-praga da invenção para as principais culturas são: Milho: *Ostrinia nubilalis*, broca do milho europeu; *Agrotis ipsilon*, lagarta-rosca; *Helicoverpa zea*,
5 lagarta da espiga; *Spodoptera frugiperda*, lagarta militar; *Diatraea grandiosella*, lagarta do cartucho; *Elasmopalpus lignosellus*, broca do colo, *Diatraea saccharalis*, broca da cana-de-açúcar; *Diabrotica virgifera*, lagarta da raiz do milho; *Diabrotica longicornis barberi*, vaquinha do norte;
10 *Diabrotica undecimpunctata howardi*, vaquinha do sul; *Melanotus* spp. Larva arame; *Cyclocephala borealis*, besouro mascarado do norte (grub branco); *Cyclocephala immaculata*, besouro mascarados do sul (grub branco); *Popillia japonica*, besouro japonês; *Chaetocnema pulicaria*, pulguinha do milho e do arroz;
15 *Sphenophorus maidis*, broca do colmo; *Rhopalosiphum maidis*, pulgão das folhas do milho; *Anuraphis maidiradicis*, afídeo da raiz do milho; *Blissus leucopterus leucopterus*, percevejo das gramíneas; *Melanoplus femurrubrum*, gafanhoto; *Melanoplus sanguinipes*, gafanhoto migratório;
20 *Hylemya platura*, verme das sementes do milho; *Agromyza parvicornis*, traça do milho; *Anaphothrips obscurus*, tripes de erva; *Solenopsis milesta*, formigas ladrão; *Tetranychus urticae*, ácaro rajado; Sorgo: *Chilo partellus*, broca do sorgo; *Spodoptera frugiperda*, lagarta militar, *Helicoverpa zea*, lagarta da espiga; *Elasmopalpus lignosellus*, broca do colo; *Feltia subterranea*, vermes granulados; *Phyllophaga crinita*, grub branco; *Eleodes*, *Conoderus*, e *Aeolus* spp. vermes fio; *Oulema melanopus*, besouro da folha de cereais; *Chaetocnema pulicaria*,
25 pulguinha do milho e do arroz; *Sphenophorus maidis*, broca

do colmo; *Rhopalosiphum maidis*; pulgão das folhas do milho; *Siphaflava*, afídeo amarelo da cana; *Blissus leucopterus leucopterus*, percevejo; *Contarinia sorghicola*, sorgo anão; *Tetranychus cinnabarinus*, ácaro vermelho; *Tetranychus*
5 *urticae*, ácaro rajado; Trigo: *Pseudaletia unipunctata*, verme exército; *Spodoptera frugiperda*, lagarta militar, *Elasmopalpus lignosellus*, broca do colo; *Agrotis orthogonia*, lagarta do oeste; *Elasmopalpus lignosellus*, broca do colo; *Oulema melanopus*, besouro de cereais; *Hypera*
10 *punctata*, gorgulho da folha do trevo; *Diabrotica undecimpunctata howardi*, vaquinha do sul; afídeo russo do trigo; *Schizaphis graminum*, pulgão; *Macrosiphum avenae*, afídeo inglês de grão; *Melanoplus femurrubrum*, gafanhoto vermelho; *Melanoplus differentialis*, gafanhoto diferencial;
15 *Melanoplus sanguinipes*, gafanhoto migratório; *Mayetiola destructor*, Mosca Hesse; *Sitodiplosis mosellana*, trigo anão; *Meromyza americana*, verme do caule do trigo; *Hylemya coarctata*, mosca do bulbo do trigo; *Frankliniella fusca*, tripses do tabaco; *Cephus cinctus*, mosca do caule do trigo;
20 *Aceria tulipae*, ácaro do trigo; Girassol: *Suleima helianthana*, mariposa das gemas do girassol; *Homoeosoma electellum*, mariposa do girassol; *Zygogramma exclamationis*, besouro do girassol; *Bothyrus gibbosus*, besouro cenoura; *Neolasioptera murtfeldtiana*, praga da semente de girassol;
25 Algodão: *Heliothis virescens*, lagarta do algodão, *Helicoverpa zea*, lagarta do algodão; *Spodoptera exigua*, lagarta da beterraba; *Pectinophora gossypiella*, lagarta rosada; *Anthonomus grandis*, bicudo; *Aphis gossypii*, afídeo do algodão; *Pseudatomoscelis seriatus*, praga do algodoeiro;
30 *Trialeurodes abutilonea*, mosca branca da asa com banda;

Lygus lineolaris, inseto manchado de planta; *Melanoplus femurrubrum*, gafanhoto vermelho; *Melanoplus differentialis*, gafanhoto diferencial, *Thrips tabaci*, tripes da cebola; *Frankliniella fusca*, tripés do tabaco; *Tetranychus*
 5 *cinnabarinus*, ácaro carmim; *Tetranychus urticae*, ácaro rajado; Arroz: *Diatraea saccharalis*, broca da cana-de-açúcar; *Spodoptera frugiperda*, lagarta militar, *Helicoverpa*
zea, lagarta da espiga; *Colaspis brunnea*, colaspis da uva; *Lissorhoptrus oryzophilus*, gorgulho aquático do arroz,
 10 *Sitophilus oryzae*, gorgulho do arroz; *Nephotettix nigropictus*, cigarrinha do arroz; *Blissus leucopterus leucopterus*, percevejo; *Acrosternum hilare*, percevejo verde; Soja: *Pseudoplusia includens*, lagarta mede-palmo da soja, *Anticarsia gemmatilis*, parasitóide de ovos;
 15 *Plathypena scabra*, verme verde do trevo; *Ostrinia nubilalis*, broca do milho europeu; *Agrotis ipsilon*, lagarta rosca; *Spodoptera exigua*, lagarta da beterraba; *Heliothis virescens*, lagarta das maçãs; *Helicoverpa zea*, lagarta do algodão; *Epilachna varivestis*, besouro do feijão mexicano;
 ~' 20 *Myzus persicae*, pulgão; *Empoasca fabae*, cigarrinha da batata; *Acrosternum hilare*, percevejo verde; *Melanoplus femurrubrum*, gafanhoto vermelho; *Melanoplus differentialis*, gafanhoto diferencial; *Hylemya platura*, verme das sementes do milho; *Sericothrips variabilis*, tripes da soja, *Thrips*
 25 *tabaci*, tripes da cebola; *Tetranychus turkestani*, ácaro do morango; *Tetranychus urticae*, ácaro rajado; Cevada: *Ostrinia nubilalis*, broca do milho europeu; *Agrotis ipsilon*, lagarta rosca; *Schizaphis graminum*, pulgão ; *Blissus leucopterus leucopterus*, percevejo; *Acrosternum*
 30 *hilare*, percevejo verde; *Euschistus servus*, percevejo

marrom; *Delia platura*, mosca da semente do milho; *Mayetiola destructor*, Mosca Hesse; *Petrobia latens*, ácaro marrom do trigo; Semente de colza: *Brevicoryne brassicae*, afídeo do besouro; *Phyllotreta cruciferae*, besouro; *Mamestra configurata*, lagarta Bertha; *Plutella xylostella*, traça das crucíferas; *Delia* spp., vermes da raiz.

Nematóides incluem nematóides parasitas, como vermes de raiz, cisto, e nematóides de lesão, incluindo *Heterodera* spp., *Meloidogyne* spp., e *Globodera* spp.; em particular os membros dos nematóides do cisto, incluindo mas não limitados a, *Heterodera glycines* (nematóide de cisto da soja); *Heterodera schachtii* (nematóide de cisto da beterraba); *Heterodera avenae* (nematóide de cisto de cereais) e *Globodera rostochiensis* e *Globodera pailida* (nematóide de cisto da batateira). Nematóides de lesão incluem *Pratylenchus* spp.

Métodos para aumento da produção vegetal

Métodos para aumentar as características de produção são fornecidos. Os métodos incluem introdução em uma planta ou em uma célula vegetal um polinucleotídeo compreendendo uma sequência pesticida aqui divulgada. Tal como definido aqui, o "rendimento" da planta se refere à qualidade e/ou a quantidade de biomassa produzida pelas plantas. Por "biomassa" entende-se qualquer medida do produto vegetal. Um aumento na produção de biomassa é uma melhora no rendimento medido do produto vegetal. Aumentar a produção vegetal tem várias aplicações comerciais. Por exemplo, o aumento da biomassa vegetal da folha pode aumentar o rendimento das hortaliças folhosas para consumo humano ou

animal. Adicionalmente, o aumento da biomassa foliar pode ser usado para aumentar a produção de produtos farmacêuticos ou produtos industriais derivados de plantas. Um aumento no rendimento pode incluir qualquer aumento estatisticamente significativo, incluindo, mas não se limitando a, pelo menos um aumento de 1%, pelo menos um 3% de aumento, pelo menos um aumento de 5%, pelo menos um aumento de 10%, pelo menos um aumento de 20%, pelo menos 30%, pelo menos 50%, pelo menos 70%, pelo menos 100%, ou um maior aumento na produtividade em comparação com uma planta não expressando a sequência pesticida.

Nos métodos específicos, o rendimento vegetal é aumentado como resultado de uma melhoria da resistência a pragas em uma planta expressando uma proteína pesticida aqui divulgada. Expressão da proteína pesticida resulta em uma redução da capacidade de uma praga para infestar ou se alimentar da planta, assim, melhorando as características de produção.

Os seguintes exemplos são oferecidos a título de ilustração e não por meio de limitação.

EXPERIMENTAL

Exemplo 1. Crescimento de ATX9387, e preparação dos extratos.

A estirpe ATX9387, identificada como um membro do grupo de *Bacillus cereus*/*Bacillus thuringiensis* pela análise de MIDI, foi cultivada em meio T3 a 30 graus por

períodos variando de 16 horas a 5 dias. As culturas foram centrifugadas e os sobrenadantes foram passados através de filtros de 0,2 micron, resultando em sobrenadantes estéreis

5 Exemplo 2. Bioensaio com *C. elegans*.

Caenorhabditis elegans ("*C. elegans*") hermafroditas foram criados como conhecido na arte, para gerar populações de animais saudáveis para bioensaio. Procedimentos gerais para o crescimento, a colheita, e a manipulação genética de *C. elegans*, incluindo o crescimento em meio, etc podem ser encontrados na arte, por exemplo, em Wood, ed. (1988) *A Nematode Caenorhabditis elegans* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Sobrenadantes estéreis da estirpe ATX9387 foram testados para a atividade em *C. elegans*. Os bioensaios foram realizados em 96 placas. Cinco a dez nematóides foram adicionados aos 80 µl de meio S (Wood, supra) e foram misturadas com 20 µl de sobrenadante estéril, 0.5 µl de HB101 concentrado (preparado conforme descrito em Wood, supra) e rifampicina (concentração final de 0,1 µg/µl). Os ensaios foram autorizados a proceder em temperatura ambiente durante 3 dias e nematóides foram quantificados. Amostras de controle negativo (meio T3 estéril ou sobrenadantes inativos de cepas) continham centenas de nematóides ativos, enquanto amostras (contendo sobrenadantes ATX9387) continham 5 a 10 nematóides que estavam letárgicos ou mortos. Os resultados do bioensaio de extratos de ATX9387 em *C. elegans* são apresentados na Tabela 1.

Tabela 1. Atividade de extratos de ATX9387 em *C. elegans*

Tempo de crescimento	ATX9387	Controle
16 horas	-	-
1 dia	-	-
2 dias	-	-
3 dias	++	-
4 dias	++	-
5 dias	++	-

5 Exemplo 3. Atividade de ATX9387 sobre nematóides do cisto da soja.

Um sobrenadante estéril de uma cultura de 5 dias de ATX9387 foi concentrado em 40 vezes e alimentado à nematóides SCN J2. Nematóides que se alimentaram do sobrenadante estéril foram reprodutivelmente observados por estarem lentos, e apresentaram maior motilidade que nematóides alimentados com o extrato de um controle negativo concentrado na mesma medida.

15 Exemplo 4. Atividade anti-nematóide a partir de ATX9387 é conferida por uma proteína.

Para identificar a atividade anti-nematóide em ATX9387, os seguintes testes foram realizados. Em primeiro lugar, a capacidade da atividade a ser destruída pelo calor foi testada por aquecimento a partir de amostras de sobrenadante estéril de ATX9387 a 100°C durante 10 minutos e, em seguida, o material avaliado foi aquecido em um

bioensaio com nematóide. O tratamento térmico de sobrenadantes estéreis de ATX9387 resultou em uma perda da atividade anti-nematóide. Em seguida, as amostras ativas foram tratadas com pronase para degradar proteínas. Este
5 tratamento resultou em perda da atividade. Sobrenadantes estéreis de ATX9387 foram passados através de 3 kDa de peso molecular de corte ("MWCO") de unidade de concentração. O ingrediente ativo foi mantido pelo filtro de 3 kDa de MWCO enquanto o que passou pelo fluxo não mostrou nenhuma
10 atividade. Isso indica que a molécula ativa foi maior do que 3 kDa em tamanho.

Exemplo 5. Fracionamento de atividade a partir de ATX9387

O sobrenadante estéril de uma cultura de 4 dias de ATX9387 foi fracionado por cromatografia líquida em uma
15 coluna de troca aniônica em 20 mM de Tris a pH 8, utilizando um gradiente de 0 a 1 mM de NaCl. Várias frações consecutivas foram ativadas. A fração mais ativa foi submetida a SDS-PAGE, e duas bandas proeminentes de aproximadamente 130kd e aproximadamente 70kd foram
20 observadas.

Em outro experimento a cepa ATX9387 foi cultivada durante 5 dias em meio T3 em 30°C. A cultura foi centrifugada a 8000 x g por 10 minutos, e o sobrenadante
25 foi passado através de um filtro de 0,2 µm. O sobrenadante filtrado foi dializado contra 20 mM de Tris em pH 8 usando Spectra/Por 1 tubulação de diálise (6-8000 MWCO), e foi então fracionado em uma coluna de troca de ânion (Mono Q), utilizando um gradiente de 0 a 1 M de NaCl mais de 20
30 volumes. As frações foram dializadas contra 20 mM de Tris

em pH 8 em slides -A-Lyzer (Pierce Biotechnology, Rockford, IL) unidades de mini diálise (7000 MWCO), e bioensaiado em *C. elegans* utilizando 20µl da amostra em um 100µl de volume de bioensaio, conforme descrito a seguir. Frações 12 e 13 foram encontradas para ser ativas, e, em seguida, foram reunidas e concentradas 7 vezes usando uma Amicon concentrador Ultra-4 5000 MWCO. O material concentrado foi fracionado em uma coluna de filtração em gel (Superdex 200) em 50 mM de fosfato de sódio, 150 mM de NaCl, em pH 7. Frações foram concentradas 10 vezes utilizando concentradores Centricon YM-3, e foram bioensaiados em *C. elegans* como acima. As frações 5 e 6 foram as mais ativas, e a fração 7 foi um pouco ativa. O SDS-PAGE das frações revelou que frações 5-7 partilharam várias proteínas: uma proteína com cerca de 130 kDa, um gibão com cerca de 75 kDa, e uma proteína com cerca de 53 kDa. As bandas mais proeminentes foram submetidas a sequenciamento da proteína N-terminal por degradação de Edman. A proteína de 130 kDa e a proteína de 70 kDa tinham muitas sequências N-terminal semelhantes (a cisteína poderia não ser detectada pelo método utilizado), e, portanto, são susceptíveis de resultar da mesma proteína inicial. Esta proteína foi designada AXMI-031 (SEQ ID NO: 2).

Tabela 2. Sequência N-terminal da proteína de aproximadamente 130kDa de ATX9387

1	A, S+S', M
2	D, Q
30	3 ?, P, N

	4	N
	5	L
	6	Q
	7	S
5	8	Q
	9	?, Q
	10	N
	11	I
	12	P
10	13	Y
	14	N
	15	V

Tabela 3. Sequência N-terminal da proteína de cerca de
15 70kDa de ATX9387

	1	A, G, S + S'
	2	F
	3	P
20	4	N
	5	L
	6	Q
	7	V?
	8	Q
25	9	?
	10	N?, V?
	11	I
	12	P
	13	Y, Q
30	14	?, N

15 ?, V

Uma pesquisa das bases de dados de proteína com as sequências N-terminal da proteína de ~130kDa e da proteína
 5 de ~70kDa demonstrou que os aminoácidos N-terminal de proteínas AXMI-031 têm grande semelhança com o N-terminal da endotoxina Cry 14Aa (SEQ ID NO: 13).

A sequência N-terminal de Cry14A: MDCNLQSQQNIPYNV (resíduos de aminoácido de 1 a 15 da SEQ ID NO: 13).

10

Exemplo 6. Clonagem da região codificadora axmi-031 de ATX9387.

Uma biblioteca de fragmento aleatório de ATX9387 foi gerada, e um clone de DNA (pAX031) que contém a sequência
 15 de DNA codificando o N-terminal de axmi-031 foi identificado. Um segundo clone, pAX032, foi identificado como contendo o C-terminal de axmi-031. Clones pAX031 e pAX032 sobrepuseram substancialmente, de tal forma que é evidente para um técnico no assunto que os dois clones
 20 juntos compreendem toda a região codificadora axmi-031.

Para confirmar a natureza de AXMI-031, um clone genômico foi amplificado usando uma polimerase PFU ULTRA (TM) (Stratagene) de alta fidelidade do DNA total usando
 25 iniciadores desenhados para a extremidade de axmi-031 do molde de leitura aberta. O produto resultante da PCR foi clonado no vetor PCR-TOPOII Blunt (Invitrogen) para criar pAX980. A sequência de DNA de pAX980 foi determinada e verificou-se que contém o mesmo molde de leitura aberta
 30 como na biblioteca de fragmento aleatório de clones de

pAX031 e pAX032. A tradução deste molde de leitura aberta gera uma sequência de proteína coerente com a sequência N-terminal obtida a partir da proteína purificada AXMI-031. Assim, este molde de leitura aberta foi designado axmi-031 (SEQ ID NO: 1). O clone plasmidial pAX2515, contendo axmi-031 foi depositado em 9 de junho de 2006 e atribuído o número de acesso NRRL B-30935.

Exemplo 7. Comparação de AXMI-031 com outras endotoxinas conhecidas

Buscas no banco de dados usando sequências de proteínas AXMI-031 demonstraram que AXMI-031 é um membro da classe de proteínas inseticidas endotoxina delta. AXMI-031 é muito semelhante a endotoxina cryl4Aa1. A sequência de aminoácidos AXMI-031 é 86,6% idêntica à CRY14Aa1 (SEQ ID NO: 13).

Exemplo 8. Expressão do polipeptídeo AXMI-031 em *E. coli*

Para expressão solúvel em *E. coli*, iniciadores foram desenhados a fim de incluir o início da tradução e os códons de parada do ORF de axmi-031. Os iniciadores foram acrescentados em um RBS ótimo e em sítios de recombinação Gateway attB. A polimerase I Stratagene PFU foi utilizada para amplificar consistentemente o ORF de pAX980, e recombinada com pDONR221 (Invitrogen, Carlsbad, CA) para criar o vetor de entrada pAX2515 (como por protocolos da Invitrogen). O clone foi sequenciado para verificação. Uma nova recombinação foi realizada para introduzir o ORF em pDEST17 para ser expresso pelo promotor T7, rendendo pAX2530. A presença e orientação do fragmento inserido de

axmi-031 foram verificadas pela digestão de restrição, e transformadas em células de *E. coli* BL21 (DE3). Este vetor produz uma fusão traducional de 26 aminoácidos (3,17 kDa), incluindo um His tag de 6x, sobre o N-terminal de AXMI-031.

5 O His-AXMI-031 foi expresso a partir de pAX2530 em células BL21*(DE3) como se segue. Uma cultura de início de pAX2530 foi cultivado por toda a noite a 37°C em LB com 50 µg/ml de carbenicilina. No dia seguinte a cultura saturada foi diluída em meio fresco 1:100, e as novas culturas
10 cultivadas a 37° para um OD de 0,4, no qual o tempo de cultura foi colocada em temperatura ambiente com agitação durante a noite. Como controle, um vetor de expressão axmi-004 portador do gene (pAX2530) foi construído como para axmi-031, e as proteínas AXMI-004 (Depósito de Patente US
15 No. 10/782.020) foram expressas a partir de pAX2504 na mesma *E. coli* hospedeira. As culturas inteiras foram analisadas por SDS-PAGE. As culturas contendo pAX2530 produziram uma banda proeminente da proteína em aproximadamente 135 kDa, o tamanho esperado para a proteína
20 His-AXMI-031, enquanto culturas contendo pAX2504 não.

Exemplo 9. Anticorpos contra AXMI-031

Para gerar anticorpos anti-AXMI-031, afinidade purificada de AXMI-031 (SEQ ID NO: 2) a proteína foi
25 inoculada em coelhos e, antisoros para AXMI-031 isolados e titulados tal como é conhecido na arte. Os antisoros selecionados foram encontrados para também reagir com proteínas SYNAXMI-031 (SEQ ID NO: 8).

30 Exemplo 10. Atividade de AXMI-031 em *C. elegans*.

Culturas de *E. coli* contendo pAX2530, foram preparadas, e encontradas para produzir uma proteína 135kDa não presente nas cepas controle, sugerindo a expressão de AXMI-031 em pAX2530 contendo as cepas. Essas culturas foram

5 testadas para a atividade em *C. elegans* e foram ativadas contra *C. elegans*, enquanto culturas controle não mostraram nenhuma atividade contra *C. elegans*.

Exemplo 11. Expressão de AXMI A-031 em *Bacillus*

10 O gene inseticida axmi-031 é amplificado por PCR a partir de pAX980, e o produto da PCR é clonado no vetor de expressão pAX916 de *Bacillus*, ou outro vetor adequado, através de métodos bem conhecidos na arte. A estirpe resultante de *Bacillus*, contendo o vetor com axmi-031 é

15 cultivado em meio de crescimento convencional, como o meio CYS (10 g/l de Bacto-casitone; 3 g/l de extrato de levedura; 6 g/l de KH_2PO_4 ; 14 g/l de K_2HPO_4 , 0,5 mM de MgSO_4 ; 0,05 mM de MnCl_2 ; 0,05 mM de FeSO_4), até esporulação ser evidente pelo exame microscópico. As amostras são

20 preparadas, e a proteína AXMI-031 foi testada para a atividade em bioensaios contra *C. elegans*, e verificou-se que tinham atividade.

Exemplo 12. *synaxmi-031*, *synaxmi-031 (apo)* e *synaxmi-031 (ER)*

25 Em um aspecto da invenção, sequências sintéticas de axmi-031 são geradas, por exemplo, *synaxmi-031* (SEQ ID NO: 7). Estas sequências sintéticas têm uma sequência de DNA modificada em relação à sequência axmi-031, e codifica uma

30 proteína que é colinear com a proteína original AXMI-031,

mas falta o "domínio cristal" C-terminal presente em AXMI-031. A sequência do gene *synaxmi-031* codifica proteínas SYNAXMI-031 (SEQ ID NO: 8), que compreende os primeiros 685 aminoácidos da proteína AXMI-031.

5

Em outro aspecto da invenção, versões modificadas do *synaxmi-031* são projetadas tais que o peptídeo resultante é direcionado para uma organela vegetal, como o retículo endoplasmático ou o apoplasto. Sequências peptídicas conhecidas para resultar no direcionamento da fusão de proteínas para organelas vegetais são conhecidas na arte. Por exemplo, a região N-terminal do gene da fosfatase ácida de White Lupin *Lupinus albus* (Genebank ID GI: 14276838, Miller et al. (2001) *Plant Physiology* 127: 594-606) é conhecida na arte de conduzir o retículo endoplasmático direcionando para proteínas heterólogas. Se a proteína de fusão resultante também contém uma sequência de retenção endoplasmática compreendendo o peptídeo N-terminal lisina-ácido aspártico-ácido glutâmico-leucina (isto é, o motivo "KDEL" (SEQ ID NO: 36), o C-terminal, a proteína de fusão será orientada para o retículo endoplasmático. Se a proteína de fusão não tem um retículo endoplasmático direcionando a sequência para o C-terminal, a proteína será orientada para o retículo endoplasmático, mas acabará por ser seqüestrado no apoplasto.

25

Assim, o gene *synaxmi-03 IER* (SEQ ID NO: 11) codifica uma proteína de fusão que contém trinta e um aminoácidos N-terminais do gene da fosfatase ácida de White Lupin *Lupinus albus* fundido para o N-terminal de SYNAXMI - 031, bem como

30

a sequência KDEL no C-terminal. Assim, a proteína resultante AXMI-03 IER (SEQ ID NO: 12) está prevista para ser direcionada para o retículo endoplasmático vegetal mediante a expressão em uma célula vegetal.

5

O gene *synaxmi-031 (apo)* (SEQ ID NO: 9) codifica uma proteína de fusão que contém trinta e um aminoácidos N-terminais do gene da fosfatase ácida de White Lupin *Lupinus albus* fundido para o N-terminal de SYNAXMI - 031, mas
 10 carece da sequência KDEL no C-terminal. Assim, a proteína resultante AXMI-031 (APO) (SEQ ID NO: 10) está prevista para ser direcionada para o apoplasto vegetal após expressão em uma célula vegetal.

15 Exemplo 13. Truncações de *synaxmi-031* para o rendimento alternativo de proteínas AXMI-031

Construtos de DNA que resultaram na expressão de proteínas variantes AXMI-031 foram desenvolvidos e expressos, além de sequências sintéticas de codificação de
 20 AXMI-031 e variantes e fragmentos destas (SEQ ID NO: 15-27). Um subconjunto destes genes foi testado para atividade nematóide in vitro.

Tabela 4. Atividade nematicida de variantes de AXMI-031 in vitro

Proteína	Nucleotídeos Sequência ID NO:	Aminoácidos Sequência ID NO:	Ativo em <i>C. elegans</i> ?
Expressão bacteriana			
AXMI-031-truncada	14	15	Sim
AXMI-031 (m1)	16	17	Sim
AXMI-031 (m1)-truncada	18	19	Sim
AXMI-031 (A-D)	20	21	Sim
SYNAXMI-031 (A-D)	26	27	NT
AXMI-031 (B-C)	22	23	Sim
AXMI-031 (B-D)	24	25	Sim

NT = Não testado

Exemplo 14. Trunicações e adição de domínio(s) celulares

5 direcionando para expressão vegetal

Em outro aspecto da invenção, versões modificadas das sequências *synaxmi-031* são projetadas tais que o peptídeo resultante é direcionado para uma organela vegetal, como o retículo endoplasmático ou o apoplasto.

Em outro aspecto da invenção, os genes são truncados, tais que o peptídeo resultante é uma versão truncada do AXMI-031, que pode ou não ser modificado para direcionamento para organelas vegetais, tais como o
 5 apoplasto ou o retículo endoplasmático.

Em outro aspecto da invenção, versões modificadas são desenvolvidas, que resultam na expressão de variantes truncadas que contêm domínios concebidos para orientar as
 10 proteínas resultantes para organelas vegetais.

As seguintes sequências nucleotídicas variantes foram concebidas:

15 *aposynaxmi-031* (A-D) (SEQ ID NO: 28) codifica a proteína APOAXMI-031 (A-D) (SEQ ID NO: 29). *apoSyn2axmi-031* (A-D) (SEQ ID NO: 30) codifica a proteína APOAXMI-031 (A-D) (SEQ ID NO: 31). *aposynaxmi-031 (fl)* (SEQ ID NO: 37) codifica a proteína APOSYNAXMI-031 (FL) (SEQ ID NO: 38).
 20 *Synaxmi031(fl)-ER* (SEQ ID NO: 32) codifica a proteína SYNAXMI-031 (FL)-ER (SEQ ID NO: 33).

Exemplo 15. Extração de DNA plasmidial de cepas ATX16538, ATX16093 e ATX21049

25 Cepas ATX16538, ATX16093, e ATX21049 foram selecionadas para análise. Culturas puras de cada linhagem foram cultivadas em grandes quantidades de meio rico. As culturas foram centrifugadas para colheita de células do sedimento. O sedimento celular foi então preparado por
 30 tratamento com EDS por métodos conhecidos na arte,

resultando na quebra da parede celular e libertação de DNA. Proteínas e grandes DNAs genômicos foram então precipitados por uma elevada concentração de sal. O DNA plasmidial foi então precipitado com etanol. Em vários casos, o DNA
5 plasmidial foi separado dos demais DNAs cromossômicos por centrifugação de alta velocidade através de um gradiente de cloreto de cério. Alternativamente, o DNA plasmidial foi purificado através da ligação a uma resina, tal como é conhecido na arte. Para cada cepa, a qualidade do DNA foi
10 verificada pela visualização de um gel de agarose por métodos conhecidos na arte.

Exemplo 16. Clonagem de genes das cepas ATX 16538, ATX 16093, e ATX21049

15 Bibliotecas de DNA plasmidial foram preparadas a partir do DNA de cada cepa. Isto pode ser alcançado de várias maneiras como conhecido na arte. Por exemplo, o DNA plasmidial purificado pode ser cisalhado em fragmentos de 5-10 kb e os fragmentos 5' e 3' de fita única reparados
20 usando DNA polimerase T4 e fragmento de Klenow na presença de todos os quatro dNTPs, tal como é conhecido na arte. Fosfatos podem então ser anexados a extremidade 5' por tratamento com polinucleotídeo quinase T4, tal como é conhecido na arte. A reparação dos fragmentos de DNA podem
25 então ser ligada por toda a noite em um vetor de cópia de padrão elevado (ou seja, pBLUESCRIPT®SK+), devidamente preparado para aceitar as inserções tal como é conhecido na arte (por exemplo, digestão com uma enzima de restrição produzindo extremidades blunt).

A qualidade do DNA resultando bibliotecas digestão foi analisado por um subconjunto de clones com uma enzima de restrição conhecida por ter uma clivagem sítios flanqueiam a clonagem sítios. Uma alta porcentagem de clones foram
5 determinados para conter inserções, geralmente com uma média inserir tamanho de 5-6 kb.

Exemplo 17. Sequenciamento através de Biblioteca de Placas

Uma vez que a qualidade da biblioteca de DNA foi
10 verificada e confirmada, colônias foram cultivadas em um caldo rico em 2ml de 96 blocos por toda a noite a 37°C, normalmente a uma velocidade de agitação de 350 rpm. Os blocos foram centrifugados para coletar as células, na parte inferior do bloco. Os blocos foram então preparados
15 por lise alcalina padrão em um formato de alto padrão.

As sequências finais dos clones desta biblioteca foram, então, determinadas para um grande número de clones a partir de cada bloco da seguinte maneira: A sequência do
20 DNA de cada clone escolhido para análise foi determinada utilizando a técnica de sequenciamento com corante fluorescente terminal (Applied Biosystems), por métodos conhecidos na arte utilizando uma máquina automática de sequenciamento de DNA, e oligonucleotídeos iniciadores
25 padrão que anelam o vetor plasmidial na região de flanqueamento do inserto.

Exemplo 18. Montagem e Análise de Dados do Sequenciamento

Sequências de DNA obtidas foram compiladas em um
30 projeto conjunto e alinhadas para formar contigs. Isto pode

ser feito de forma eficiente usando um programa de computador, tais como Vector NTI, ou, em alternativa, utilizando os programas Pred/Phrap de alinhamento e análise de DNA, tal como descrito aqui. Estes contigs, juntamente com qualquer leitura individual que pode não ter sido adicionada a um contig, foram comparados com uma compilação de dados de todas as classes de genes pesticidas conhecidos. Contigs ou leituras individuais identificadas como tendo uma identidade para uma endotoxina ou gene pesticida conhecido foram analisados mais detalhadamente.

Da estirpe ATX1 6538, pAX2579 foi encontrado para conter uma um molde de leitura aberta com homologia para endotoxinas delta do tipo "cry". Este molde de leitura aberta (ORF) foi designado como axmi-039 (SEQ ID NO: 3), e a proteína codificada foi designada AXMI-039 (SEQ ID NO: 5). O ORF axmi-039 e a sequência de acompanhamento (151 pb a montante do códon de início e 29 pb a jusante do códon de parada) foi amplificado por PCR, clonado em pRSFIB e sequenciado para gerar pAX2579. pAX2579 foi depositado junto ao ARS Patent Strain Collection em 9 de junho de 2006, e atribuído o nº NRRL B-30936. AXMI-039 é 43,9% de identidade de sequência de aminoácidos para Cry5Ba1, que é o homólogo mais próximo identificado.

25

Da estirpe ATX 16093, pAX4313 foi encontrado para conter um molde de leitura aberto com homologia com a endotoxina delta do tipo "cry". Este molde de leitura aberto (ORF) foi designado axmi-040 (SEQ ID NO: 7), e a proteína codificada foi designada AXMI-040 (SEQ ID NO: 8).

30

pAX4313 foi depositada junto do ARS Patent Strain Collection em 9 de junho de 2006, e atribuído o nº NRRL B-30937. AXMI-040 é 42,9% de identidade de sequência de aminoácidos para Cry21Bal, que é o homólogo mais próximo
5 identificado.

Da estirpe ATX21049, um molde de leitura aberto foi identificado que apresentou homologia com endotoxina delta do tipo "cry". Este molde de leitura aberto foi designado
10 axmi-049 (SEQ ID NO: 34), e a proteína codificada foi designada AXMI-049 (SEQ ID NO: 35). Este molde de leitura aberto foi amplificado por PCR e clonado em um vetor de rendimento pAX5039. pAX5039 foi depositado junto do ARS Patent Strain Collection em 29 de maio de 2007, e atribuído
15 o nº NRRL B-50046. AXMI-049 tem 46,1% de identidade de sequência de aminoácidos para Cry21Bal, que é o homólogo mais próximo identificado.

Exemplo 19. Vetorização dos genes pesticidas da invenção 20 para expressão vegetal

Cada uma das regiões codificantes dos genes da invenção estão conectadas independentemente com o promotor adequado e as sequências terminadoras de expressão em plantas. Tais sequências são bem conhecidas na arte e podem
25 incluir o promotor actina do arroz ou o promotor ubiquitina do milho para expressão em monocotiledôneas, um promotor Arabidopsis UBQ3 ou promotor CaMV 35S de expressão em dicotiledôneas, e os terminadores NSA ou PinII. Técnicas para produzir e confirmar construções promotor-gene-
30 terminador também são bem conhecidas na arte.

Exemplo 20. Transformação dos genes da invenção em células vegetais por Transformação mediada por *Agrobacterium*

Espigas foram colhidas 8-12 dias após a polinização.

- 5 Embriões foram isolados das espigas, e estes embriões de 0,8-1,5 mm de tamanho foram utilizados para a transformação. Embriões foram plaqueados com o escutelo para cima sobre meios de incubação adequados, e incubados por toda a noite a 25°C no escuro. No entanto, não é
- 10 necessário, por si só para incubar os embriões durante a noite. Embriões são contatados com uma estirpe de *Agrobacterium* contendo os vetores adequados para transferência mediada por plasmídeo Ti por 5-10 minutos e, em seguida, plaqueados em meio de co-cultura durante 3 dias
- 15 (25°C no escuro). Depois do co-cultivo, os explantes são transferidos para o meio de recuperação por cinco dias (a 25°C no escuro). Explantes são incubados em meios de seleção por até oito semanas, dependendo da natureza e característica de cada seleção utilizada. Após o período de
- 20 seleção, o calo resultante é transferido para o meio de maturação de embrião, até a formação de embriões somáticos maduros ser observada. Os embriões somáticos maduros resultantes são então colocados sob baixa luminosidade, e o processo de regeneração é iniciado como conhecido na arte.
- 25 Os ramos resultantes são possibilitados de enraizar sobre o meio de enraizamento, e as plantas resultantes são transferidas para os potes creche e propagadas como plantas transgênicas.

Exemplo 21. Plantas transgênicas expressando AXMI-031 e variantes.

Os cassetes de expressão vegetal descritos aqui são combinados com um marcador vegetal selecionável adequado para auxiliar nas seleções de células e tecidos transformados, e ligados em vetores de transformação vegetal. Estes podem incluir vetores binários de transformação mediada por *Agrobacterium* ou vetores plasmidiais simples para aerossol ou transformação por biobalística, *synaxmi-031 (apo)* foi clonado em um vetor de expressão vegetal, e esse vetor foi introduzido em *Agrobacterium tumefaciens* tal como é conhecido na arte. A região de codificação *synaxmi-031 (apo)* sob controle do promotor Arabidopsis UBQ3 (Norris et al. (1993) *J. Plant Mol. Biol.* 21: 895-906) foi introduzida em *Arabidopsis thaliana* por métodos florais conhecidos na arte, e plantas transgênicas foram selecionadas. A presença de *synaxmi-031 (apo)* na população transgênica T1 foi confirmada por análise por PCR tal como é conhecido na arte, utilizando oligonucleotídeos derivados da sequência de *synaxmi-031*, que também anelam para *synaxmi-031 (apo)*.

Tecido foliar e raiz de *synaxmi-031 (apo)* transgênicos contendo plantas foram preparados e separados em gel de poliacrilamida por controles não-transgênicos e diluições de proteínas AXMI-031, e transferidos para uma membrana como conhecido na arte. As amostras foram testadas para a presença do produto da expressão de *synaxmi-031 (apo)* por análise de Western Blot como conhecido na arte, utilizando anticorpos anti-AXMI-031 descritos a seguir. O produto de

expressão de *synaxmi-031 (apo)* de tamanho esperado foi detectado em amostras de ambos os tecidos das folhas e raízes para plantas contendo *synaxmi-031 (apo)* transgênicos, mas não foi detectado nos controles não-
 5 transgênicos.

Tabela 5. Detecção de anticorpos SYNAXMI-031 (APO) em plantas transgênicas

Amostra de tecido	Detecção de SYNAXMI -031 (APO)
Controle não transformado	-
<i>synaxmi-031 (apo)</i> contendo planta 'A'	+++
<i>synaxmi-031 (apo)</i> contendo planta 'B'	+++

10

Exemplo 22. Formação de cisto reduzido por plantas expressando AXMI-031.

Plantas transgênicas expressando *synaxmi-031 (apo)* foram testadas para a capacidade de resistir a infestação de *Heterodera schachtii*, em comparação com plantas controle. Plantas transgênicas, bem como plantas controle transformadas com vetor sozinho, foram infestadas com aproximadamente 100 hatchlings J2, e o número de nematóides entrando nas raízes, bem como o número de cistos formados,
 15 foram medidos. Plantas transgênicas expressando SYNAXMI-031 (APO) foram encontradas para ter reduzido consistentemente
 20

o número de nematóides que entram nas raízes, e reduzido o número de cistos formados em relação aos controles.

5 Tabela 6. Formação de cisto reduzido em plantas expressando AXMI-031

	Formação de cisto
Planta controle	++
Plantas expressando AXMI-031 (APO)	+

Exemplo 23. Ensaios adicionais para atividade pesticida

10 A capacidade de uma proteína pesticida para agir como um pesticida sobre uma praga é frequentemente avaliada de várias maneiras. Uma forma bem conhecida na arte é realizar um ensaio de alimentação. Nesse ensaio de alimentação, um expõe as pragas de uma amostra contendo compostos para serem testados, ou amostras controle. Frequentemente este é

15 feito colocando o material a ser ensaiado, ou uma diluição adequada de tais materiais, em um material que irá ingerir as pragas, tais como uma dieta artificial. O material a ser testado pode ser composto por um líquido, sólido ou chorume. O material a ser testado pode ser colocado sobre a

20 superfície e, em seguida, secado. Alternativamente, o material a ser testado pode ser misturado com uma dieta artificial liquefeita e, em seguida, dispensado na câmara de ensaio. A câmara de ensaio pode ser, por exemplo, um copo, um prato, ou uma placa de microplacas.

25

Ensaios para pragas sugadoras (pulgões, por exemplo) podem envolver a separação do material de teste para o

inseto por uma divisão, idealmente uma porção que pode ser fracionada pelas partes bucais sugadoras dos insetos sugadores, a fim de permitir a ingestão de material de ensaio. Muitas vezes o material teste é misturado com um
5 estimulante alimentar, como a sacarose, para promover a ingestão do composto teste.

Outros tipos de ensaios podem incluir microinjeção do material teste na boca, ou intestino das pragas, bem como o
10 desenvolvimento de plantas transgênicas, seguido pelo teste da capacidade das pragas para se alimentar a partir da planta transgênica. Testes vegetais podem envolver isolamento das partes das plantas normalmente consumidas, por exemplo, pequenas gaiolas anexadas a uma folha, ou
15 isolamento de plantas inteiras em gaiolas contendo insetos.

Outros métodos e abordagens para a análise de pragas são conhecidos na arte, e podem ser encontrados, por exemplo, em Robertson, J. L. & H. K. Preisler. 1992. Pesticide bioassays with arthropods. CRC, Boca Raton, FL.
20 Alternativamente, ensaios são comumente descritos nas revistas "Arthropod Management Tests" e "Journal of Economic Entomology" ou pela discussão com os membros da Sociedade Entomológica da América (ESA).

25 Exemplo 24. Transformação de células de milho com genes pesticidas da invenção

Espigas de milho foram colhidas 8-12 dias após a polinização. Embriões foram isolados das espigas, e os embriões de tamanho 0,8-1,5 mm foram utilizadas para a
30 transformação. Embriões foram plaqueados com o escutelo

para cima sobre um meio de incubação adequado, tais como meio DN62A5S (3,98 g/L de Sais N6; 1 mL/L (1000x de solução estoque) Vitaminas N6; 800 mg/L de L-asparagina; 100 mg/L de Mio-inositol, 1,4 g/L de L-prolina, 100 mg/L de Casaminoácidos, 50 g/L de sucrose; 1 mL/L (de 1 mg/ml de solução estoque) 2,4-D), e incubados por toda a noite a 25°C no escuro.

Os explantes resultantes são transferidos para uma malha quadrada (30-40 por placa), transferidos para meios osmóticos por 30-45 minutos e, em seguida, transferidos para uma placa irradiante (ver, por exemplo, Publicação PCT e No. WO/0138514 e Patente US No. 5.240.842).

Construções de DNA destinados para expressar os genes da invenção em células vegetais são acelerados em tecido vegetal utilizando um acelerador de feixe de aerossol, utilizando condições essencialmente como descritas na Publicação PCT No. WO/0138514. Após o irradiação, os embriões são incubados por 30 min em meio osmótico e, em seguida, colocados sobre meios de incubação por toda a noite, a 25°C no escuro. Para evitar danificar os explantes irradiados, eles são incubados por pelo menos 24 horas antes da transferência para o meio de recuperação. Os embriões são então espalhados em um meio para o período de recuperação, por 5 dias, a 25°C, no escuro, e em seguida, transferidos para um meio de seleção. Explantes são incubados em meios de seleção por até oito semanas, dependendo da natureza e característica de cada seleção utilizada. Após o período de seleção, o calo resultante é transferido para o meio de maturação de embrião, até a

formação de embriões somáticos maduros ser observada. Os embriões somáticos maduros resultantes são então colocados sob baixa luminosidade, e o processo de regeneração é iniciado através de métodos conhecidos na arte. Os ramos resultantes são possibilitados de enraizar sobre o meio de enraizamento, e as plantas resultantes são transferidas para os potes creche e propagadas como plantas transgênicas.

10 Materiais

Meio DN62A5S

Componentes	Por litro	Amostra
Mistura salina Chu's N6 Basal (Prod. No. C 416)	3,98g/L	Phytotechnology Labs.
Solução de vitamina Chu's N6 (Prod. No. C 149)	1 mL/L (1000x de Solução Estoque)	Phytotechnology Labs.
L-Asparagina	800 mg/L	Phytotechnology Labs.
Mio-inositol	100 mg/L	Sigma
L-Prolina	1,4 g/L	Phytotechnology Labs.
Casaminoácidos	100 mg/L	Fisher Scientific
Sucrose	50g/L	Phytotechnology Labs.
2,4-D (Prod. No. D 7299)	1 mL/L (1mg/mL de Solução Estoque)	Sigma

Ajustar o pH da solução para pH 5,8 com 1N de KOH/1N de KCl, adicionar Gelrite (Sigma) para 3g/L, e autoclavar. Após arrefecimento a 50°C, adicionar 2 ml/L a 5 mg/ml de
5 solução de nitrato de prata (Phytotechnology Labs). A receita produz cerca de 20 placas.

Todas as publicações e pedidos de patentes mencionados na especificação são indicativos do nível de habilidade das
10 pessoas qualificadas na arte à qual pertence esta invenção. Todas as publicações e pedidos de patentes são aqui incorporados por referência à mesma medida, como se cada publicação ou pedido de patente fosse especificamente e individualmente indicado a ser incorporado por referência.

15

Apesar de a invenção exposta ter sido descrita em algum detalhe a título de ilustração e de exemplo para fins de clareza de entendimento, será evidente que algumas modificações e alterações possam ser praticadas no âmbito
20 das reivindicações anexadas.

LISTAGEM DE SEQUÊNCIAS

<110> Depositante: Athenix Corporation

Nadine Carozzi
 Nicholas Duck
 Theodore Kahn
 Nalini Desai
 Shruti Jain
 Daniel John Tomso
 Rong Guo
 Julia Thissen

<120> Título da Invenção: AXMI-031, AXMI-039, AXMI-040 E AXMI-049, UMA FAMÍLIA DE NOVOS GENES DE DELTA-ENDOTOXINA E MÉTODOS PARA SEU USO

<130> Referência do documento: 45600/329694

<150> Número do Documento de Prioridade: 60/813.774

<151> Data de Depósito do Documento de Prioridade: 14-06-2006

<160> Quantidade de SEQ ID NOs.: 38

<170> Software: FastSEQ for Windows Version 4.0

<210> SEQ ID NO: 1

<211> Comprimento: 3558

<212> Tipo: DNA

<213> Organismo: *Bacillus cereus/Bacillus thuringiensis*

<220> Características:

<221> Nome/Chave: CDS

<222> Localização: (1)...(3558)

<400> Sequência: 1

atg gat tgt aat tta caa tca caa caa aat att cca tat aat gta tta	48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu	
1 5 10 15	
gca ata cca gta tct aat gtt aat tcg ttg act gat aca gtt gga gat	96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp	
20 25 30	
tta aaa aaa gca tgg gaa gaa ttt caa aaa act ggt tct ttt tca tta	144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu	
35 40 45	
aca gct tta caa caa gga ttt tct gct tca caa gga gga aca ttc aat	192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn	
50 55 60	
tat tta aca tta cta caa tca gga ata tca tta gct ggt tct ttt gtt	240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val	
65 70 75 80	

cct gga ggt act ttt gta gca cct att att aat atg gtt att ggt tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	
tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata aat	336
Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn	
100 105 110	
tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat tct	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc act	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat tat	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	
aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca aat	576
Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca cct	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	
tat ttt gtt ata gga gca aca gca cgt ttt gca gca atg caa tct tat	672
Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr	
210 215 220	
att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac gca	720
Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala	
225 230 235 240	
cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat atg	768
Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met	
245 250 255	
cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt aaa	816
Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys	
260 265 270	
gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt gat	864
Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp	
275 280 285	
acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat att	912
Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile	
290 295 300	

gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa aca	960
Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr	
305 310 315 320	
gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa gaa	1008
Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu	
325 330 335	
gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct ttt	1056
Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe	
340 345 350	
agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att tct	1104
Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser	
355 360 365	
tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc cct	1152
Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro	
370 375 380	
cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta aat	1200
Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn	
385 390 395 400	
tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa tct	1248
Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser	
405 410 415	
cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat gca	1296
Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala	
420 425 430	
gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga ata	1344
Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile	
435 440 445	
gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca gaa	1392
Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu	
450 455 460	
cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc tgt	1440
Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys	
465 470 475 480	
aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt aca	1488
Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr	
485 490 495	
caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca agt	1536
Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser	
500 505 510	
ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gaa tta gat	1584
Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp	
515 520 525	

tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa gga	1632
Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly	
530 535 540	
tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att aat	1680
Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn	
545 550 555 560	
ggg gca agt gct gta cca ctt gat tca gga aat acc tta ttt atg acg	1728
Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr	
565 570 575	
gct acg aat tta aca gct act caa tat aga att aga ata cgt tat gca	1776
Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala	
580 585 590	
aat cca aat tca aat act caa atc ggt gta cga att aca caa aat ggt	1824
Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly	
595 600 605	
tct cta att tcc agt agt aat cta aca ctt tat agt act act gat atg	1872
Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met	
610 615 620	
aat aat act tta cca cta aat gta tat gta ata gga gaa aat gga aat	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tat aca ctt caa gat tta tat aat act act aat gtt tta tca aca gga	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	
gat att aca tta caa att aca gga gga gat caa aaa ata ttt att gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	
cga ata gaa ttt gtt cct act atg cct gta cct ggt aat act aac aac	2064
Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn	
675 680 685	
aat aac ggt aat aat aac ggt aat aat aat ccc cca cac cac gtt tgt	2112
Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val Cys	
690 695 700	
gca ata gct ggt aca caa caa tct tgt tct gga ccg ccc aaa ttt gaa	2160
Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe Glu	
705 710 715 720	
caa gta agt gat tta gaa aaa att aca aca caa gta tat atg tta ttc	2208
Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu Phe	
725 730 735	
aaa tct tct ccg tat gaa gaa tta gct cta gaa gtt tcc agc tat caa	2256
Lys Ser Ser Pro Tyr Glu Glu Leu Ala Leu Glu Val Ser Ser Tyr Gln	
740 745 750	

att agt caa gta gca tta aaa gtt atg gca tta tct gat gaa cta ttt	2304
Ile Ser Gln Val Ala Leu Lys Val Met Ala Leu Ser Asp Glu Leu Phe	
755 760 765	
tgt gaa gaa aaa aac gta tta cga aaa tta gtc aat aaa gca aaa caa	2352
Cys Glu Glu Lys Asn Val Leu Arg Lys Leu Val Asn Lys Ala Lys Gln	
770 775 780	
tta tta gaa gca agt aac tta cta gta ggt gga aat ttt gaa aca act	2400
Leu Leu Glu Ala Ser Asn Leu Leu Val Gly Gly Asn Phe Glu Thr Thr	
785 790 795 800	
caa aat tgg gta ctt gga aca aat gct tat ata aat tat gat tcg ttt	2448
Gln Asn Trp Val Leu Gly Thr Asn Ala Tyr Ile Asn Tyr Asp Ser Phe	
805 810 815	
tta ttt aat gga aat tat tta tct tta caa cca gca agt gga ttt ttc	2496
Leu Phe Asn Gly Asn Tyr Leu Ser Leu Gln Pro Ala Ser Gly Phe Phe	
820 825 830	
aca tct tat gct tat caa aaa ata gat gag tca aca tta aaa cca tat	2544
Thr Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Thr Leu Lys Pro Tyr	
835 840 845	
aca cga tat aaa gtt tct ggg ttc att ggg caa agt aat caa gta gaa	2592
Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val Glu	
850 855 860	
ctt att att tct cgt tat gga aaa gaa att gat aaa ata tta aat gtt	2640
Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val	
865 870 875 880	
cca tat gca gga cct ctt cct atc act gct gat gca tca ata act tgt	2688
Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys	
885 890 895	
tgt gca cca gaa ata ggc caa tgt gat ggg gaa caa tct gat tct cat	2736
Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His	
900 905 910	
ttc ttt aac tat agc atc gat gta ggt gca ctt cac cca gaa tta aac	2784
Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn	
915 920 925	
cct ggc att gaa att ggt ctt aaa att gtg caa tca aat ggt tat ata	2832
Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile	
930 935 940	
aca att agt aat cta gaa att att gaa gaa cgt cca ctt aca gaa atg	2880
Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met	
945 950 955 960	
gaa att caa gca gtc aat cga aaa aat caa aaa tgg gaa aga gaa aaa	2928
Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys	
965 970 975	

ctt cta gaa tgt gca agt att agt gaa ctt tta caa cca att att aat Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn 980 985 990	2976
caa atc gat tca ttg ttt aaa gat gga aac tgg tat aat gat att ctt Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu 995 1000 1005	3024
cct cat gtc aca tat caa gat tta aaa aat att ata ata ccc gag tta Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu 1010 1015 1020	3072
cca aaa tta aaa cat tgg ttc ata gag aat ctc cca ggt gaa tat cat Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His 1025 1030 1035 1040	3120
gaa att gaa caa aaa atg aaa gaa gct cta aaa tat gca ttt aca caa Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln 1045 1050 1055	3168
tta gac gag aaa aat tta atc cac aat ggt cac ttt aca act aac tta Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu 1060 1065 1070	3216
ata gat tgg caa gta gaa ggt gat gct caa atg aaa gta tta gaa aat Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn 1075 1080 1085	3264
gat gct ctt gca tta caa ctt ttc aac tgg gat gct agt gct tca caa Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser Gln 1090 1095 1100	3312
tct ata aat ata tta gaa ttt gat gaa gat aag gca tat aaa ctt cgc Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg 1105 1110 1115 1120	3360
gta tat gct caa gga agc gga aca atc caa ttt gga aac tgt gaa gat Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp 1125 1130 1135	3408
gaa gct atc caa ttt aat aca aac tca ttc ata tat caa gaa aaa ata Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile 1140 1145 1150	3456
gtc tat ttc gat acc cca tca gtt aat tta cac ata caa tca gaa ggt Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly 1155 1160 1165	3504
tct gaa ttt att gta agt agt atc gat cta att gaa tta tca gac gac Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp 1170 1175 1180	3552
caa taa Gln * 1185	3558

<210> SEQ ID NO: 2

<211> Comprimento: 1185

<212> Tipo: PRT

<213> Organismo: *Bacillus cereus/Bacillus thuringiensis*

<400> Sequência: 2

```

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

```

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

Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val Glu
 850 855 860
 Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val
 865 870 875 880
 Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys
 885 890 895
 Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His
 900 905 910
 Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn
 915 920 925
 Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile
 930 935 940
 Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met
 945 950 955 960
 Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys
 965 970 975
 Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn
 980 985 990
 Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu
 995 1000 1005
 Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu
 1010 1015 1020
 Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His
 1025 1030 1035 1040
 Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln
 1045 1050 1055
 Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu
 1060 1065 1070
 Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn
 1075 1080 1085
 Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser Gln
 1090 1095 1100
 Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg
 1105 1110 1115 1120
 Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp
 1125 1130 1135
 Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile
 1140 1145 1150
 Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly
 1155 1160 1165
 Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp
 1170 1175 1180
 Gln
 1185

<210> SEQ ID NO: 3

<211> Comprimento: 3984

<212> Tipo: DNA

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo

<221> Nome/Chave: CDS

<222> Localização: (1)...(3984)

<400> Sequência: 3

atg gca aca ctt aat aat atg ttt tca gtt cct tat aat gta cta gct	48
Met Ala Thr Leu Asn Asn Met Phe Ser Val Pro Tyr Asn Val Leu Ala	
1 5 10 15	
cta ccc att atc ccc aat tct atc tta act ttt gaa gat aat cga aaa	96
Leu Pro Ile Ile Pro Asn Ser Ile Leu Thr Phe Glu Asp Asn Arg Lys	
20 25 30	
aaa ata gaa gag ggt att aaa gag ttt gaa aag act gga cgt ata aaa	144
Lys Ile Glu Glu Gly Ile Lys Glu Phe Glu Lys Thr Gly Arg Ile Lys	
35 40 45	
ccc ctt aaa gat tta ata gag ctc ata ttc aaa ggg tat agt gac gat	192
Pro Leu Lys Asp Leu Ile Glu Leu Ile Phe Lys Gly Tyr Ser Asp Asp	
50 55 60	
gaa tct tct tac gca gca tta gtt caa act atg ctt gtc gtt att ccc	240
Glu Ser Ser Tyr Ala Ala Leu Val Gln Thr Met Leu Val Val Ile Pro	
65 70 75 80	
ttg gcg ttc cct gaa tta gcc cca gtt ctt cca att att ggc gta gta	288
Leu Ala Phe Pro Glu Leu Ala Pro Val Leu Pro Ile Ile Gly Val Val	
85 90 95	
att aat ttc gtt ttt cca ggt ttg aag ggt tct gct aaa tca acc tat	336
Ile Asn Phe Val Phe Pro Gly Leu Lys Gly Ser Ala Lys Ser Thr Tyr	
100 105 110	
aca atg att aca gag atg gtt gat aaa gcc att aac caa tca ttc acg	384
Thr Met Ile Thr Glu Met Val Asp Lys Ala Ile Asn Gln Ser Phe Thr	
115 120 125	
gcc cag att aca aat ata tta aca aac aat att act ggt ata caa aat	432
Ala Gln Ile Thr Asn Ile Leu Thr Asn Asn Ile Thr Gly Ile Gln Asn	
130 135 140	
aat ata caa tct gtt tac gac gca atg agc aat gcc att gga aca aat	480
Asn Ile Gln Ser Val Tyr Asp Ala Met Ser Asn Ala Ile Gly Thr Asn	
145 150 155 160	
gac aca att cat aat ttc atc aga aat aat gat aca aca cct tgc tct	528
Asp Thr Ile His Asn Phe Ile Arg Asn Asn Asp Thr Thr Pro Cys Ser	
165 170 175	
caa aac aat cag cca gct tgt cct tgt cct cca aat aat caa tgt tta	576
Gln Asn Asn Gln Pro Ala Cys Pro Cys Pro Pro Asn Asn Gln Cys Leu	
180 185 190	
caa aaa gtt gtt gat gaa tat act aaa gca ata gct aat atc gat ctc	624
Gln Lys Val Val Asp Glu Tyr Thr Lys Ala Ile Ala Asn Ile Asp Leu	
195 200 205	
att atc cct caa ttt cat gat cca ctt act ggt gta att tcc gat tta	672
Ile Ile Pro Gln Phe His Asp Pro Leu Thr Gly Val Ile Ser Asp Leu	
210 215 220	

gct act gca aac atg tat att tta cct tta tac gct caa act gtt aat	720
Ala Thr Ala Asn Met Tyr Ile Leu Pro Leu Tyr Ala Gln Thr Val Asn	
225 230 235 240	
tta aaa tta att ttg cgt cag agt ttc att gaa ttt atg gag aaa tat	768
Leu Lys Leu Ile Leu Arg Gln Ser Phe Ile Glu Phe Met Glu Lys Tyr	
245 250 255	
aaa tat gat gaa aaa gaa act gtt ttt cag gct ttt att aat gct gat	816
Lys Tyr Asp Glu Lys Glu Thr Val Phe Gln Ala Phe Ile Asn Ala Asp	
260 265 270	
att cca gaa caa att aaa aaa ctt cgt caa gat att atc acg tat aca	864
Ile Pro Glu Gln Ile Lys Lys Leu Arg Gln Asp Ile Ile Thr Tyr Thr	
275 280 285	
aaa gat att tat atg caa ttt gaa gct cac gct ccc tac ccg act tat	912
Lys Asp Ile Tyr Met Gln Phe Glu Ala His Ala Pro Tyr Pro Thr Tyr	
290 295 300	
aat tca aaa aaa caa cta aat gac tat atc cgt tat aca aga att atc	960
Asn Ser Lys Lys Gln Leu Asn Asp Tyr Ile Arg Tyr Thr Arg Ile Ile	
305 310 315 320	
caa gta tat tgt ttg gat tta gta gca atg tgg cct acg ctt gat cga	1008
Gln Val Tyr Cys Leu Asp Leu Val Ala Met Trp Pro Thr Leu Asp Arg	
325 330 335	
gtg aat tat gca tta cct gtc caa caa aat atg aca cgc att ata ttt	1056
Val Asn Tyr Ala Leu Pro Val Gln Gln Asn Met Thr Arg Ile Ile Phe	
340 345 350	
gga gat ctt att ggc cct gta gaa act gtg cct cag gtt ccc cgc caa	1104
Gly Asp Leu Ile Gly Pro Val Glu Thr Val Pro Gln Val Pro Arg Gln	
355 360 365	
aat tca gat aac ttt cat ttt aat ttg tct gat gtt tat aga aac ccc	1152
Asn Ser Asp Asn Phe His Phe Asn Leu Ser Asp Val Tyr Arg Asn Pro	
370 375 380	
tta cct aat aat gat att ttt aat tac cgt tat ggc ggg ctt caa att	1200
Leu Pro Asn Asn Asp Ile Phe Asn Tyr Arg Tyr Gly Gly Leu Gln Ile	
385 390 395 400	
tct aaa gcg caa ttt atg aca tat tat aaa aaa ttc gga gct ttc agt	1248
Ser Lys Ala Gln Phe Met Thr Tyr Tyr Lys Lys Phe Gly Ala Phe Ser	
405 410 415	
acc cat gat gaa tat tat tat gta gat ggg cat cgc cta agt ttt aat	1296
Thr His Asp Glu Tyr Tyr Tyr Val Asp Gly His Arg Leu Ser Phe Asn	
420 425 430	
act tca gat aaa aaa aca att gaa atc aat gct cat caa aat tct cat	1344
Thr Ser Asp Lys Lys Thr Ile Glu Ile Asn Ala His Gln Asn Ser His	
435 440 445	

gat act att gaa aag gat gta ata gat caa cta ggc ggc tat gtt aca	1392
Asp Thr Ile Glu Lys Asp Val Ile Asp Gln Leu Gly Gly Tyr Val Thr	
450 455 460	
cggt atg aat atg gta agt caa gaa atc ggt agt ggc agt ata ttg ggt	1440
Arg Met Asn Met Val Ser Gln Glu Ile Gly Ser Gly Ser Ile Leu Gly	
465 470 475 480	
gga aat gcc ata gat cga act gcc ctt act acc agt aca ggt gat aaa	1488
Gly Asn Ala Ile Asp Arg Thr Ala Leu Thr Thr Ser Thr Gly Asp Lys	
485 490 495	
gac ata ctt gct gca gga ttt aac gcg ata aat aat gga aat tac aag	1536
Asp Ile Leu Ala Ala Gly Phe Asn Ala Ile Asn Asn Gly Asn Tyr Lys	
500 505 510	
aca gaa ggt cgc ggt gcc tcc cat aat cag gat gtt cca aat cat aaa	1584
Thr Glu Gly Arg Gly Ala Ser His Asn Gln Asp Val Pro Asn His Lys	
515 520 525	
gtt caa gcg att tat cct gtt caa ccg aat aaa tat gaa tat gac gat	1632
Val Gln Ala Ile Tyr Pro Val Gln Pro Asn Lys Tyr Glu Tyr Asp Asp	
530 535 540	
tat gga tct gaa gaa aag tat gga tac ctt tat act tta att cca atg	1680
Tyr Gly Ser Glu Glu Lys Tyr Gly Tyr Leu Tyr Thr Leu Ile Pro Met	
545 550 555 560	
gat gtg agt ttc att ccc caa ctc aat gca gaa gat gta att act act	1728
Asp Val Ser Phe Ile Pro Gln Leu Asn Ala Glu Asp Val Ile Thr Thr	
565 570 575	
att cct gca gaa caa gca gtc aaa att aat ggt caa tct gta ata gat	1776
Ile Pro Ala Glu Gln Ala Val Lys Ile Asn Gly Gln Ser Val Ile Asp	
580 585 590	
act gta aac acc aat tct tta ttg gaa ttt gtt aat ggt gca aat gca	1824
Thr Val Asn Thr Asn Ser Leu Leu Glu Phe Val Asn Gly Ala Asn Ala	
595 600 605	
att aaa cta tct cct ggc gaa aca gca caa tac act atg att aat ccc	1872
Ile Lys Leu Ser Pro Gly Glu Thr Ala Gln Tyr Thr Met Ile Asn Pro	
610 615 620	
gtc aat cgt tcc tac cag gtt aga gtt cgg gtt gct act gaa gga gaa	1920
Val Asn Arg Ser Tyr Gln Val Arg Val Arg Val Ala Thr Glu Gly Glu	
625 630 635 640	
acc caa tta gat att ata gca ccc gtt gat agc aat acc ctt aat ctt	1968
Thr Gln Leu Asp Ile Ile Ala Pro Val Asp Ser Asn Thr Leu Asn Leu	
645 650 655	
tca aac aca aaa aca aca gca aat gat caa aat ggg ata tta ggt aaa	2016
Ser Asn Thr Lys Thr Thr Ala Asn Asp Gln Asn Gly Ile Leu Gly Lys	
660 665 670	

caa ggt aat tat ata gta ttt cca cag cca aac att gat act gtc aca	2064
Gln Gly Asn Tyr Ile Val Phe Pro Gln Pro Asn Ile Asp Thr Val Thr	
675 680 685	
ggc acc agt tta cct cct acc gaa aac att atg aat ttt cca gta gga	2112
Gly Thr Ser Leu Pro Pro Thr Glu Asn Ile Met Asn Phe Pro Val Gly	
690 695 700	
aca ttt gat ttt acc att gta aat tct gga aac tca gat act att tta	2160
Thr Phe Asp Phe Thr Ile Val Asn Ser Gly Asn Ser Asp Thr Ile Leu	
705 710 715 720	
gat cga att gaa ttt gtt cct att gta act tct aat aaa att caa caa	2208
Asp Arg Ile Glu Phe Val Pro Ile Val Thr Ser Asn Lys Ile Gln Gln	
725 730 735	
gat ttc act att tct cct gga aca agt caa gta att tgg acc gga aat	2256
Asp Phe Thr Ile Ser Pro Gly Thr Ser Gln Val Ile Trp Thr Gly Asn	
740 745 750	
tca gct aac act ata gat att aca att att gac aac gta gat aca agc	2304
Ser Ala Asn Thr Ile Asp Ile Thr Ile Ile Asp Asn Val Asp Thr Ser	
755 760 765	
ggc tta ttc gtt caa att ttc caa aaa ggt aag caa ctg cat gga gaa	2352
Gly Leu Phe Val Gln Ile Phe Gln Lys Gly Lys Gln Leu His Gly Glu	
770 775 780	
cta act ttg att gat ccc gct cat att caa cgt aca ttc tct gaa aaa	2400
Leu Thr Leu Ile Asp Pro Ala His Ile Gln Arg Thr Phe Ser Glu Lys	
785 790 795 800	
ttt gat caa att gtt ttg tat aac cca ggt tat aat agt tct ata tct	2448
Phe Asp Gln Ile Val Leu Tyr Asn Pro Gly Tyr Asn Ser Ser Ile Ser	
805 810 815	
ggc aca gtg agt ggt tca gta agc tct att cct aag aaa ttc gaa tta	2496
Gly Thr Val Ser Gly Ser Val Ser Ser Ile Pro Lys Lys Phe Glu Leu	
820 825 830	
tct agc gac tta caa aat atc aca aac caa gtt aat aac tta ttt gca	2544
Ser Ser Asp Leu Gln Asn Ile Thr Asn Gln Val Asn Asn Leu Phe Ala	
835 840 845	
tcg agc gaa cat gat aca ctt gcc aca gat gta agt gat tat gat att	2592
Ser Ser Glu His Asp Thr Leu Ala Thr Asp Val Ser Asp Tyr Asp Ile	
850 855 860	
gaa gaa gtg att cta aaa gta gat gca tta tct gat gaa gtg ttt gga	2640
Glu Glu Val Ile Leu Lys Val Asp Ala Leu Ser Asp Glu Val Phe Gly	
865 870 875 880	
aaa gag aaa aaa gca cta cgt aaa ttg gta aat caa gcg aag cgt tta	2688
Lys Glu Lys Lys Ala Leu Arg Lys Leu Val Asn Gln Ala Lys Arg Leu	
885 890 895	

agt aaa gcg cgt aat ctt ctc cta gga gga agt ttt gat aat ttg gat	2736
Ser Lys Ala Arg Asn Leu Leu Leu Gly Gly Ser Phe Asp Asn Leu Asp	
900 905 910	
gct tgg tat agg gga cga aat gta gta act gta tct gat cat gaa ctg	2784
Ala Trp Tyr Arg Gly Arg Asn Val Val Thr Val Ser Asp His Glu Leu	
915 920 925	
ttt aaa agt gat cat ata tta tta cca cta cca aca aca ttg tat cca	2832
Phe Lys Ser Asp His Ile Leu Leu Pro Leu Pro Thr Thr Leu Tyr Pro	
930 935 940	
tct tat ctt ttc caa aaa gta gag gaa tct aaa tta aaa gca aat aca	2880
Ser Tyr Leu Phe Gln Lys Val Glu Glu Ser Lys Leu Lys Ala Asn Thr	
945 950 955 960	
cgt tat act gtc tct ggt ttc atc gcg cat gca gag gat tta gaa att	2928
Arg Tyr Thr Val Ser Gly Phe Ile Ala His Ala Glu Asp Leu Glu Ile	
965 970 975	
gtt gtt tct cgt tat ggg caa gaa ata aag aaa gtg gtg caa gtt cca	2976
Val Val Ser Arg Tyr Gly Gln Glu Ile Lys Lys Val Val Gln Val Pro	
980 985 990	
tat ggt gaa gca ttc cca ttg aca tcg agc gga cca att tgt tgt aga	3024
Tyr Gly Glu Ala Phe Pro Leu Thr Ser Ser Gly Pro Ile Cys Cys Arg	
995 1000 1005	
cca cgt tct aga gta aat ggt aaa cct gct gat cca cat ttc ttt agt	3072
Pro Arg Ser Arg Val Asn Gly Lys Pro Ala Asp Pro His Phe Phe Ser	
1010 1015 1020	
tac agc att gat gta ggt gca tta gat gtg gaa gca aat cct ggt att	3120
Tyr Ser Ile Asp Val Gly Ala Leu Asp Val Glu Ala Asn Pro Gly Ile	
1025 1030 1035 1040	
gaa tta ggg ttt cgt att gta gag cca act gga atg gca cgt gta agt	3168
Glu Leu Gly Phe Arg Ile Val Glu Pro Thr Gly Met Ala Arg Val Ser	
1045 1050 1055	
aac tta gaa att cgt gag gat cgc cca tta acg gca aat gaa ata cgt	3216
Asn Leu Glu Ile Arg Glu Asp Arg Pro Leu Thr Ala Asn Glu Ile Arg	
1060 1065 1070	
aag gtg caa cgt gct gct aga gat tgg aaa caa aag tat gat caa gag	3264
Lys Val Gln Arg Ala Ala Arg Asp Trp Lys Gln Lys Tyr Asp Gln Glu	
1075 1080 1085	
cgt gcg gaa gta aga gcc ctg att caa cct gtt tta aat caa atc aat	3312
Arg Ala Glu Val Arg Ala Leu Ile Gln Pro Val Leu Asn Gln Ile Asn	
1090 1095 1100	
gcg ttg tat gaa aat gaa gat tgg aat gga gca att cgt tct gga gtt	3360
Ala Leu Tyr Glu Asn Glu Asp Trp Asn Gly Ala Ile Arg Ser Gly Val	
1105 1110 1115 1120	

15/123

tct tat cat gac tta gaa gca att gtt tta cca aca tta cca aaa tta	3408
Ser Tyr His Asp Leu Glu Ala Ile Val Leu Pro Thr Leu Pro Lys Leu	
1125 1130 1135	
aat cat tgg ttt atg tcc gat atg tta ggg gaa caa ggt tcc att tta	3456
Asn His Trp Phe Met Ser Asp Met Leu Gly Glu Gln Gly Ser Ile Leu	
1140 1145 1150	
gct caa ttc caa gaa gca tta aat cgt gcg tat acg caa ctc gaa gga	3504
Ala Gln Phe Gln Glu Ala Leu Asn Arg Ala Tyr Thr Gln Leu Glu Gly	
1155 1160 1165	
agt aca att ctg cat aat ggc cat ttc aca aca gat gca gca aat tgg	3552
Ser Thr Ile Leu His Asn Gly His Phe Thr Thr Asp Ala Ala Asn Trp	
1170 1175 1180	
acg ata gaa ggc gat gca cat cag gta gta tta gaa gat ggt aga cgt	3600
Thr Ile Glu Gly Asp Ala His Gln Val Val Leu Glu Asp Gly Arg Arg	
1185 1190 1195 1200	
gta tta cga ttg cca gat tgg tct tcg agt ttg tct caa acg att gaa	3648
Val Leu Arg Leu Pro Asp Trp Ser Ser Ser Leu Ser Gln Thr Ile Glu	
1205 1210 1215	
atc gag aat ttt gat cca gat aaa gaa tat caa tta gta ttt cat ggg	3696
Ile Glu Asn Phe Asp Pro Asp Lys Glu Tyr Gln Leu Val Phe His Gly	
1220 1225 1230	
caa gga gaa gga acg gtt acg ttg gag cat ggt gaa gaa gga gaa tat	3744
Gln Gly Glu Gly Thr Val Thr Leu Glu His Gly Glu Glu Gly Glu Tyr	
1235 1240 1245	
gtg gaa aca cac ccg cat aag ttt gcg aat ttt aca act tct caa cgt	3792
Val Glu Thr His Pro His Lys Phe Ala Asn Phe Thr Thr Ser Gln Arg	
1250 1255 1260	
caa gga att acg ttt gaa tca aat aaa gtg aca gtg act att tct tca	3840
Gln Gly Ile Thr Phe Glu Ser Asn Lys Val Thr Val Thr Ile Ser Ser	
1265 1270 1275 1280	
gaa gat gga gaa ttc tta gcg gat aat att gca att gtg gaa gtt cct	3888
Glu Asp Gly Glu Phe Leu Ala Asp Asn Ile Ala Ile Val Glu Val Pro	
1285 1290 1295	
atg ttt aac aag aat caa atg gtc aat gaa aat aga ggt gta aat ata	3936
Met Phe Asn Lys Asn Gln Met Val Asn Glu Asn Arg Gly Val Asn Ile	
1300 1305 1310	
aat agc aat aca aat atg aat agt agc aat aac agc aat aac caa taa	3984
Asn Ser Asn Thr Asn Met Asn Ser Ser Asn Asn Ser Asn Asn Gln *	
1315 1320 1325	

<210> SEQ ID NO: 4

<211> Comprimento: 1327

<212> Tipo: PRT

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo

<400> Seqüência: 4

Met	Ala	Thr	Leu	Asn	Met	Phe	Ser	Val	Pro	Tyr	Asn	Val	Leu	Ala
1				5				10					15	
Leu	Pro	Ile	Ile	Pro	Asn	Ser	Ile	Leu	Thr	Phe	Glu	Asp	Asn	Arg
			20					25				30		Lys
Lys	Ile	Glu	Glu	Gly	Ile	Lys	Glu	Phe	Glu	Lys	Thr	Gly	Arg	Ile
		35					40				45			Lys
Pro	Leu	Lys	Asp	Leu	Ile	Glu	Leu	Ile	Phe	Lys	Gly	Tyr	Ser	Asp
	50					55				60			Asp	Asp
Glu	Ser	Ser	Tyr	Ala	Ala	Leu	Val	Gln	Thr	Met	Leu	Val	Val	Ile
65					70					75				Pro
Leu	Ala	Phe	Pro	Glu	Leu	Ala	Pro	Val	Leu	Pro	Ile	Ile	Gly	Val
				85					90					95
Ile	Asn	Phe	Val	Phe	Pro	Gly	Leu	Lys	Gly	Ser	Ala	Lys	Ser	Thr
			100					105					110	Tyr
Thr	Met	Ile	Thr	Glu	Met	Val	Asp	Lys	Ala	Ile	Asn	Gln	Ser	Phe
			115				120					125		Thr
Ala	Gln	Ile	Thr	Asn	Ile	Leu	Thr	Asn	Asn	Ile	Thr	Gly	Ile	Gln
						135					140			Asn
Asn	Ile	Gln	Ser	Val	Tyr	Asp	Ala	Met	Ser	Asn	Ala	Ile	Gly	Thr
145					150					155				160
Asp	Thr	Ile	His	Asn	Phe	Ile	Arg	Asn	Asn	Asp	Thr	Thr	Pro	Cys
				165					170					175
Gln	Asn	Asn	Gln	Pro	Ala	Cys	Pro	Cys	Pro	Pro	Asn	Asn	Gln	Cys
			180					185					190	Leu
Gln	Lys	Val	Val	Asp	Glu	Tyr	Thr	Lys	Ala	Ile	Ala	Asn	Ile	Asp
		195					200					205		Leu
Ile	Ile	Pro	Gln	Phe	His	Asp	Pro	Leu	Thr	Gly	Val	Ile	Ser	Asp
	210					215					220			Leu
Ala	Thr	Ala	Asn	Met	Tyr	Ile	Leu	Pro	Leu	Tyr	Ala	Gln	Thr	Val
225					230					235				Asn
Leu	Lys	Leu	Ile	Leu	Arg	Gln	Ser	Phe	Ile	Glu	Phe	Met	Glu	Lys
				245					250					255
Lys	Tyr	Asp	Glu	Lys	Glu	Thr	Val	Phe	Gln	Ala	Phe	Ile	Asn	Ala
			260					265					270	Asp
Ile	Pro	Glu	Gln	Ile	Lys	Lys	Leu	Arg	Gln	Asp	Ile	Ile	Thr	Tyr
		275					280					285		Thr
Lys	Asp	Ile	Tyr	Met	Gln	Phe	Glu	Ala	His	Ala	Pro	Tyr	Pro	Thr
	290					295					300			Tyr
Asn	Ser	Lys	Lys	Gln	Leu	Asn	Asp	Tyr	Ile	Arg	Tyr	Thr	Arg	Ile
305					310					315				Ile
Gln	Val	Tyr	Cys	Leu	Asp	Leu	Val	Ala	Met	Trp	Pro	Thr	Leu	Asp
				325					330					Arg
Val	Asn	Tyr	Ala	Leu	Pro	Val	Gln	Gln	Asn	Met	Thr	Arg	Ile	Ile
			340					345					350	Phe
Gly	Asp	Leu	Ile	Gly	Pro	Val	G							

Ser	Lys	Ala	Gln	Phe	Met	Thr	Tyr	Tyr	Lys	Lys	Phe	Gly	Ala	Phe	Ser
				405					410					415	
Thr	His	Asp	Glu	Tyr	Tyr	Tyr	Val	Asp	Gly	His	Arg	Leu	Ser	Phe	Asn
			420					425					430		
Thr	Ser	Asp	Lys	Lys	Thr	Ile	Glu	Ile	Asn	Ala	His	Gln	Asn	Ser	His
		435					440					445			
Asp	Thr	Ile	Glu	Lys	Asp	Val	Ile	Asp	Gln	Leu	Gly	Gly	Tyr	Val	Thr
		450					455				460				
Arg	Met	Asn	Met	Val	Ser	Gln	Glu	Ile	Gly	Ser	Gly	Ser	Ile	Leu	Gly
465					470					475					480
Gly	Asn	Ala	Ile	Asp	Arg	Thr	Ala	Leu	Thr	Thr	Ser	Thr	Gly	Asp	Lys
				485					490					495	
Asp	Ile	Leu	Ala	Ala	Gly	Phe	Asn	Ala	Ile	Asn	Asn	Gly	Asn	Tyr	Lys
			500					505					510		
Thr	Glu	Gly	Arg	Gly	Ala	Ser	His	Asn	Gln	Asp	Val	Pro	Asn	His	Lys
		515					520					525			
Val	Gln	Ala	Ile	Tyr	Pro	Val	Gln	Pro	Asn	Lys	Tyr	Glu	Tyr	Asp	Asp
		530					535				540				
Tyr	Gly	Ser	Glu	Glu	Lys	Tyr	Gly	Tyr	Leu	Tyr	Thr	Leu	Ile	Pro	Met
545					550					555					560
Asp	Val	Ser	Phe	Ile	Pro	Gln	Leu	Asn	Ala	Glu	Asp	Val	Ile	Thr	Thr
				565					570					575	
Ile	Pro	Ala	Glu	Gln	Ala	Val	Lys	Ile	Asn	Gly	Gln	Ser	Val	Ile	Asp
			580					585					590		
Thr	Val	Asn	Thr	Asn	Ser	Leu	Leu	Glu	Phe	Val	Asn	Gly	Ala	Asn	Ala
		595					600					605			
Ile	Lys	Leu	Ser	Pro	Gly	Glu	Thr	Ala	Gln	Tyr	Thr	Met	Ile	Asn	Pro
		610					615				620				
Val	Asn	Arg	Ser	Tyr	Gln	Val	Arg	Val	Arg	Val	Ala	Thr	Glu	Gly	Glu
625					630					635					640
Thr	Gln	Leu	Asp	Ile	Ile	Ala	Pro	Val	Asp	Ser	Asn	Thr	Leu	Asn	Leu
			645						650					655	
Ser	Asn	Thr	Lys	Thr	Thr	Ala	Asn	Asp	Gln	Asn	Gly	Ile	Leu	Gly	Lys
			660					665					670		
Gln	Gly	Asn	Tyr	Ile	Val	Phe	Pro	Gln	Pro	Asn	Ile	Asp	Thr	Val	Thr
		675					680					685			
Gly	Thr	Ser	Leu	Pro	Pro	Thr	Glu	Asn	Ile	Met	Asn	Phe	Pro	Val	Gly
		690					695				700				
Thr	Phe	Asp	Phe	Thr	Ile	Val	Asn	Ser	Gly	Asn	Ser	Asp	Thr	Ile	Leu
705					710					715					720
Asp	Arg	Ile	Glu	Phe	Val	Pro	Ile	Val	Thr	Ser	Asn	Lys	Ile	Gln	Gln
				725					730					735	
Asp	Phe	Thr	Ile	Ser	Pro	Gly	Thr	Ser	Gln	Val	Ile	Trp	Thr	Gly	Asn
			740					745					750		
Ser	Ala	Asn	Thr												

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

Met Phe Asn Lys Asn Gln Met Val Asn Glu Asn Arg Gly Val Asn Ile
 1300 1305 1310
 Asn Ser Asn Thr Asn Met Asn Ser Ser Asn Asn Ser Asn Asn Gln
 1315 1320 1325

<210> SEQ ID NO: 5

<211> Comprimento: 3720

<212> Tipo: DNA

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo

<221> Nome/Chave: CDS

<222> Localização: (1)...(3720)

<400> Sequência: 5

atg aat gat atg aca aat tta tct aaa tta tat tca cct gta cca tac	48
Met Asn Asp Met Thr Asn Leu Ser Lys Leu Tyr Ser Pro Val Pro Tyr	
1 5 10 15	
aat gtg tta gca act cca aat gtc tta gca aca aac aaa caa cca cta	96
Asn Val Leu Ala Thr Pro Asn Val Leu Ala Thr Asn Lys Gln Pro Leu	
20 25 30	
gca gat aca gat gca tta aat aaa ttt tac aac gat ttg caa att gga	144
Ala Asp Thr Asp Ala Leu Asn Lys Phe Tyr Asn Asp Leu Gln Ile Gly	
35 40 45	
aaa gtt tca gct ttt aca att gat gca tta tgg agc ctt atg aca acc	192
Lys Val Ser Ala Phe Thr Ile Asp Ala Leu Trp Ser Leu Met Thr Thr	
50 55 60	
gga aaa tat gac tgg tgg tca att gct aaa ttt tgt tgg tct ctt gga	240
Gly Lys Tyr Asp Trp Ser Ser Ile Ala Lys Phe Cys Trp Ser Leu Gly	
65 70 75 80	
act ggt gta aca cct ctg tta ggt ata ttt tca cct att ata gat ata	288
Thr Gly Val Thr Pro Leu Leu Gly Ile Phe Ser Pro Ile Ile Asp Ile	
85 90 95	
att ttc cct gct tta ttc ggt ggt aac aaa cta agt cta ttt gaa caa	336
Ile Phe Pro Ala Leu Phe Gly Gly Asn Lys Leu Ser Leu Phe Glu Gln	
100 105 110	
cta aaa cct caa ata gaa aaa ctg att att gaa aaa tta aca gat gaa	384
Leu Lys Pro Gln Ile Glu Lys Leu Ile Ile Glu Lys Leu Thr Asp Glu	
115 120 125	
gaa aaa aat ttc tta gct caa aaa act agt gat att caa tcc tat tta	432
Glu Lys Asn Phe Leu Ala Gln Lys Thr Ser Asp Ile Gln Ser Tyr Leu	
130 135 140	
aat gat tat aaa agt gca gtt tct aaa att aat aat cca aat gtt ata	480
Asn Asp Tyr Lys Ser Ala Val Ser Lys Ile Asn Asn Pro Asn Val Ile	
145 150 155 160	

gat agc gat ttt gaa tct tta cat gca aca att aat cta aca tta tca	528
Asp Ser Asp Phe Glu Ser Leu His Ala Thr Ile Asn Leu Thr Leu Ser	
165 170 175	
aaa ata aaa gga tca tta tct tat ttc tct atc ttt aac cag cca gat	576
Lys Ile Lys Gly Ser Leu Ser Tyr Phe Ser Ile Phe Asn Gln Pro Asp	
180 185 190	
gat aga aaa cca att tat aca ata cta ggt tta cct tat tat acc ctt	624
Asp Arg Lys Pro Ile Tyr Thr Ile Leu Gly Leu Pro Tyr Tyr Thr Leu	
195 200 205	
atg gcc act atg tat tta aca tta tta cga gac gtt att tta aat aca	672
Met Ala Thr Met Tyr Leu Thr Leu Leu Arg Asp Val Ile Leu Asn Thr	
210 215 220	
aca aaa tgg aaa ata tca cca gct tct aat att tcc tat cgt caa caa	720
Thr Lys Trp Lys Ile Ser Pro Ala Ser Asn Ile Ser Tyr Arg Gln Gln	
225 230 235 240	
ttt aaa caa aac atg aat tca ttt gtt ctt aca att aaa aat aat tat	768
Phe Lys Gln Asn Met Asn Ser Phe Val Leu Thr Ile Lys Asn Asn Tyr	
245 250 255	
caa act ggt ttt aat tat ata aca aat gaa gct tac aaa gca cat cct	816
Gln Thr Gly Phe Asn Tyr Ile Thr Asn Glu Ala Tyr Lys Ala His Pro	
260 265 270	
aca aat cct tca aag act ata ctt cca ttt gaa aat aaa atg aca ttg	864
Thr Asn Pro Ser Lys Thr Ile Leu Pro Phe Glu Asn Lys Met Thr Leu	
275 280 285	
gac tgt ttt gac tat gtt gca atg tgg ccc act tta tat cca gat gat	912
Asp Cys Phe Asp Tyr Val Ala Met Trp Pro Thr Leu Tyr Pro Asp Asp	
290 295 300	
tat tat act gaa aaa aca aat ttg caa aaa act cgc tta tta ttt tcc	960
Tyr Tyr Thr Glu Lys Thr Asn Leu Gln Lys Thr Arg Leu Leu Phe Ser	
305 310 315 320	
cca ata tta ggt cgt atg cca gat tct cga agt caa tgg ctt cat agt	1008
Pro Ile Leu Gly Arg Met Pro Asp Ser Arg Ser Gln Trp Leu His Ser	
325 330 335	
aaa cct tat tcc tgg gat agt aat aaa act ttt acg ttc gat cac tat	1056
Lys Pro Tyr Ser Trp Asp Ser Asn Lys Thr Phe Thr Phe Asp His Tyr	
340 345 350	
tat atg gct gaa ctc aca cac att gac act aaa gaa ttt gat cga gta	1104
Tyr Met Ala Glu Leu Thr His Ile Asp Thr Lys Glu Phe Asp Arg Val	
355 360 365	
gac aga atc cgt cag att tat caa gaa gga tat caa aaa gaa caa caa	1152
Asp Arg Ile Arg Gln Ile Tyr Gln Glu Gly Tyr Gln Lys Glu Gln Gln	
370 375 380	

acg tat gat gat tat tac act tat ggc gga gat tct gct caa aat act	1200
Thr Tyr Asp Asp Tyr Tyr Thr Tyr Gly Gly Asp Ser Ala Gln Asn Thr	
385 390 395 400	
tct ttt aca aca gat aat cca ctt gca atc atg tat cct act cga ggt	1248
Ser Phe Thr Thr Asp Asn Pro Leu Ala Ile Met Tyr Pro Thr Arg Gly	
405 410 415	
ggg aat tat gta ggt acc gct att aaa tgg ttt gat gac aca gtg caa	1296
Gly Asn Tyr Val Gly Thr Ala Ile Lys Trp Phe Asp Asp Thr Val Gln	
420 425 430	
ggg ggt cga tct tca gga tat aca act cca tat tct gga gac cct gat	1344
Gly Gly Arg Ser Ser Gly Tyr Thr Thr Pro Tyr Ser Gly Asp Pro Asp	
435 440 445	
cct ata ata act cct gat gat cat aaa gtt aat ttt ctc tat aca gta	1392
Pro Ile Ile Thr Pro Asp Asp His Lys Val Asn Phe Leu Tyr Thr Val	
450 455 460	
aaa gat gaa tta aaa gga att gat gca tgg gtc aat tca tgg gtt ccc	1440
Lys Asp Glu Leu Lys Gly Ile Asp Ala Trp Val Asn Ser Trp Val Pro	
465 470 475 480	
att tat aca act gtt cca aat ata att gaa aat gaa atg ttc ctc act	1488
Ile Tyr Thr Thr Val Pro Asn Ile Ile Glu Asn Glu Met Phe Leu Thr	
485 490 495	
aca cta ggt ttt cct ttt gaa aaa gga ata att gat aca ggc gga gct	1536
Thr Leu Gly Phe Pro Phe Glu Lys Gly Ile Ile Asp Thr Gly Gly Ala	
500 505 510	
ggg gga gat aaa ata tat caa cta gaa cga tta aat gga tcc atg gct	1584
Gly Gly Asp Lys Ile Tyr Gln Leu Glu Arg Leu Asn Gly Ser Met Ala	
515 520 525	
atc aac tta aaa ttt aaa caa ata att aaa tta cca ttt aca aat cta	1632
Ile Asn Leu Lys Phe Lys Gln Ile Ile Lys Leu Pro Phe Thr Asn Leu	
530 535 540	
aca acc ggt aat tat cta att cgt tta cgt tat gca agt cat tcg gat	1680
Thr Thr Gly Asn Tyr Leu Ile Arg Leu Arg Tyr Ala Ser His Ser Asp	
545 550 555 560	
ata aat gca ttt act cac att cac tct gaa aat gga gct gat atc agc	1728
Ile Asn Ala Phe Thr His Ile His Ser Glu Asn Gly Ala Asp Ile Ser	
565 570 575	
agt act cct tta gga aat att act ctt cct aac acg caa aat ttc act	1776
Ser Thr Pro Leu Gly Asn Ile Thr Leu Pro Asn Thr Gln Asn Phe Thr	
580 585 590	
ttt cct act aac gat gaa tac caa ccc aat caa ccc caa tat acg acc	1824
Phe Pro Thr Asn Asp Glu Tyr Gln Pro Asn Gln Pro Gln Tyr Thr Thr	
595 600 605	

tat ata gag gga aat gct ggt aaa tat gct tta tat caa ttt aca caa Tyr Ile Glu Gly Asn Ala Gly Lys Tyr Ala Leu Tyr Gln Phe Thr Gln 610 615 620	1872
aat atc tcg ctc aca tct gga caa tat act ttc tat att caa aat aat Asn Ile Ser Leu Thr Ser Gly Gln Tyr Thr Phe Tyr Ile Gln Asn Asn 625 630 635 640	1920
agt aat act gat ctc ttt tta gat cgg att gaa ttt gtt cct atg ccg Ser Asn Thr Asp Leu Phe Leu Asp Arg Ile Glu Phe Val Pro Met Pro 645 650 655	1968
cct tct tca ata tca tta cct gat att gaa ata act aac aca gat tat Pro Ser Ser Ile Ser Leu Pro Asp Ile Glu Ile Thr Asn Thr Asp Tyr 660 665 670	2016
gaa att tgg aaa agt gat cgt cca tat ggc cat tct ata aat gga ata Glu Ile Trp Lys Ser Asp Arg Pro Tyr Gly His Ser Ile Asn Gly Ile 675 680 685	2064
ttc ata gtc agc gtg cca ttt gga aac cag aca gat act gta act ata Phe Ile Val Ser Val Pro Phe Gly Asn Gln Thr Asp Thr Val Thr Ile 690 695 700	2112
aca tat tgg aat aat ggg gaa aaa gtc cat aca gac tct caa aca ttt Thr Tyr Trp Asn Asn Gly Glu Lys Val His Thr Asp Ser Gln Thr Phe 705 710 715 720	2160
gat atg gca aga ttc caa ggt caa gat ctt aca caa tgg caa gga gct Asp Met Ala Arg Phe Gln Gly Gln Asp Leu Thr Gln Trp Gln Gly Ala 725 730 735	2208
ttt gat cgt gta act att cgt aga act cat tct gat ggt aca ttg tca Phe Asp Arg Val Thr Ile Arg Arg Thr His Ser Asp Gly Thr Leu Ser 740 745 750	2256
tta aca agt gca acc cta tac ttt gtt atc cca aaa tca tcc ttt agt Leu Thr Ser Ala Thr Leu Tyr Phe Val Ile Pro Lys Ser Ser Phe Ser 755 760 765	2304
acc cca gaa gat tta gaa aaa atc aca aac caa gtc aat cag tta ttt Thr Pro Glu Asp Leu Glu Lys Ile Thr Asn Gln Val Asn Gln Leu Phe 770 775 780	2352
act tcc tca tcc caa aca gaa tta gct aac acc gta acc gat tac gga Thr Ser Ser Ser Gln Thr Glu Leu Ala Asn Thr Val Thr Asp Tyr Gly 785 790 795 800	2400
att gat caa gtt ttg atg aaa gta gat gcg tta tca gac gat gta ttt Ile Asp Gln Val Leu Met Lys Val Asp Ala Leu Ser Asp Asp Val Phe 805 810 815	2448
ggg gta gag aaa aaa gca tta cgt aaa ctt gtc aat cag gcc aaa caa Gly Val Glu Lys Lys Ala Leu Arg Lys Leu Val Asn Gln Ala Lys Gln 820 825 830	2496

ctg agt aaa gca cgc aat gta tta gtc ggt gga aac ttt gaa gga aat	2544
Leu Ser Lys Ala Arg Asn Val Leu Val Gly Gly Asn Phe Glu Gly Asn	
835 840 845	
cat gaa tgg gta ctc ggt cgt aaa gcg gtc atg gtc gca aat gat gag	2592
His Glu Trp Val Leu Gly Arg Lys Ala Val Met Val Ala Asn Asp Glu	
850 855 860	
tta ttt aaa ggc aat cat tta tta tta cca cct cca agt cta tat cca	2640
Leu Phe Lys Gly Asn His Leu Leu Leu Pro Pro Pro Ser Leu Tyr Pro	
865 870 875 880	
tcg tat gcg tat caa aaa gtg gat gaa tcc aaa tta aaa ccg aat aca	2688
Ser Tyr Ala Tyr Gln Lys Val Asp Glu Ser Lys Leu Lys Pro Asn Thr	
885 890 895	
cga tat acc gtt tct ggt ttt gtg gcg caa agt gaa caa tta gaa gtc	2736
Arg Tyr Thr Val Ser Gly Phe Val Ala Gln Ser Glu Gln Leu Glu Val	
900 905 910	
gtt gtt tcc cgt tat ggt aaa gaa gta cat gac atg tta aat gtt cct	2784
Val Val Ser Arg Tyr Gly Lys Glu Val His Asp Met Leu Asn Val Pro	
915 920 925	
tat gaa gaa gcg tta ccg att tct tca aat gaa aag tca aat tgt tgt	2832
Tyr Glu Glu Ala Leu Pro Ile Ser Ser Asn Glu Lys Ser Asn Cys Cys	
930 935 940	
aaa ccg gct act tgc aac tat aca tct tgt gag ggg aaa gaa cca gat	2880
Lys Pro Ala Thr Cys Asn Tyr Thr Ser Cys Glu Gly Lys Glu Pro Asp	
945 950 955 960	
tct cat ttc ttc cgt tat agt atc gat gtc ggt gct tta caa cca gaa	2928
Ser His Phe Phe Arg Tyr Ser Ile Asp Val Gly Ala Leu Gln Pro Glu	
965 970 975	
gca aat cta gga att gaa ttc ggt ctt cgt att gtg aaa tca aat gga	2976
Ala Asn Leu Gly Ile Glu Phe Gly Leu Arg Ile Val Lys Ser Asn Gly	
980 985 990	
ttt gca aaa atc agt aat tta gaa atc aaa gaa gat cgt cca tta aca	3024
Phe Ala Lys Ile Ser Asn Leu Glu Ile Lys Glu Asp Arg Pro Leu Thr	
995 1000 1005	
gac caa gaa att aaa aaa gta caa cac aaa gaa caa aag tgg aaa aaa	3072
Asp Gln Glu Ile Lys Lys Val Gln His Lys Glu Gln Lys Trp Lys Lys	
1010 1015 1020	
gca ttt aac aaa gaa caa gca gag tta acg gca aca ctc caa cca aca	3120
Ala Phe Asn Lys Glu Gln Ala Glu Leu Thr Ala Thr Leu Gln Pro Thr	
1025 1030 1035 1040	
ctg aat caa atc aat gcc ttg tat caa cag gaa gat tgg aat ggt tcg	3168
Leu Asn Gln Ile Asn Ala Leu Tyr Gln Gln Glu Asp Trp Asn Gly Ser	
1045 1050 1055	

att cat cct cat gtg acg tat caa cat ctg tct gat gtt gtc tta cca	3216
Ile His Pro His Val Thr Tyr Gln His Leu Ser Asp Val Val Leu Pro	
1060 1065 1070	
acg tta cca aaa caa aca cat tgg ttt atg gaa aat cga gaa ggc gaa	3264
Thr Leu Pro Lys Gln Thr His Trp Phe Met Glu Asn Arg Glu Gly Glu	
1075 1080 1085	
cat gtt gtt ctg acg caa caa ttc caa caa gca tta gat cgt gct ttc	3312
His Val Val Leu Thr Gln Gln Phe Gln Gln Ala Leu Asp Arg Ala Phe	
1090 1095 1100	
caa caa atc gaa gaa caa aac ttg atc cac aat ggt agc ttt aca agt	3360
Gln Gln Ile Glu Glu Gln Asn Leu Ile His Asn Gly Ser Phe Thr Ser	
1105 1110 1115 1120	
gga tta aca gat tgg act gta aca gga gat gca cag att act atc tat	3408
Gly Leu Thr Asp Trp Thr Val Thr Gly Asp Ala Gln Ile Thr Ile Tyr	
1125 1130 1135	
gat gaa gat cca gta tta gaa cta gca cat tgg gat gca agc gtc tct	3456
Asp Glu Asp Pro Val Leu Glu Leu Ala His Trp Asp Ala Ser Val Ser	
1140 1145 1150	
caa acg att gag att act gat ttt gaa gaa gaa aaa gaa tac aaa ctt	3504
Gln Thr Ile Glu Ile Thr Asp Phe Glu Glu Glu Lys Glu Tyr Lys Leu	
1155 1160 1165	
cgt gta cgc gga aaa ggc aaa gga acg gtt acc gtt caa cat gaa gaa	3552
Arg Val Arg Gly Lys Gly Lys Gly Thr Val Thr Val Gln His Glu Glu	
1170 1175 1180	
gag tta gaa aca atg aca ttt aac aca act agt ttt aca acg aaa gaa	3600
Glu Leu Glu Thr Met Thr Phe Asn Thr Thr Ser Phe Thr Thr Lys Glu	
1185 1190 1195 1200	
caa acc ttc tat ttc gaa gga aat aca ata gat gta cac gtt caa tca	3648
Gln Thr Phe Tyr Phe Glu Gly Asn Thr Ile Asp Val His Val Gln Ser	
1205 1210 1215	
gag aat aat gca ttc ctt gta gac agt gtg gaa ctc att gaa gtt gta	3696
Glu Asn Asn Ala Phe Leu Val Asp Ser Val Glu Leu Ile Glu Val Val	
1220 1225 1230	
aaa gaa caa gaa gaa aaa caa taa	3720
Lys Glu Gln Glu Glu Lys Gln *	
1235	

<210> SEQ ID NO: 6

<211> Comprimento: 1239

<212> Tipo: PRT

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo

<400> Sequência: 6

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

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

27/123

```

Arg Tyr Thr Val Ser Gly Phe Val Ala Gln Ser Glu Gln Leu Glu Val
          900                      905                      910
Val Val Ser Arg Tyr Gly Lys Glu Val His Asp Met Leu Asn Val Pro
          915                      920                      925
Tyr Glu Glu Ala Leu Pro Ile Ser Ser Asn Glu Lys Ser Asn Cys Cys
          930                      935                      940
Lys Pro Ala Thr Cys Asn Tyr Thr Ser Cys Glu Gly Lys Glu Pro Asp
945          950                      955                      960
Ser His Phe Phe Arg Tyr Ser Ile Asp Val Gly Ala Leu Gln Pro Glu
          965                      970                      975
Ala Asn Leu Gly Ile Glu Phe Gly Leu Arg Ile Val Lys Ser Asn Gly
          980                      985                      990
Phe Ala Lys Ile Ser Asn Leu Glu Ile Lys Glu Asp Arg Pro Leu Thr
          995                      1000                      1005
Asp Gln Glu Ile Lys Lys Val Gln His Lys Glu Gln Lys Trp Lys Lys
1010          1015                      1020
Ala Phe Asn Lys Glu Gln Ala Glu Leu Thr Ala Thr Leu Gln Pro Thr
1025          1030                      1035                      1040
Leu Asn Gln Ile Asn Ala Leu Tyr Gln Gln Glu Asp Trp Asn Gly Ser
          1045                      1050                      1055
Ile His Pro His Val Thr Tyr Gln His Leu Ser Asp Val Val Leu Pro
          1060                      1065                      1070
Thr Leu Pro Lys Gln Thr His Trp Phe Met Glu Asn Arg Glu Gly Glu
          1075                      1080                      1085
His Val Val Leu Thr Gln Gln Phe Gln Gln Ala Leu Asp Arg Ala Phe
          1090                      1095                      1100
Gln Gln Ile Glu Glu Gln Asn Leu Ile His Asn Gly Ser Phe Thr Ser
1105          1110                      1115                      1120
Gly Leu Thr Asp Trp Thr Val Thr Gly Asp Ala Gln Ile Thr Ile Tyr
          1125                      1130                      1135
Asp Glu Asp Pro Val Leu Glu Leu Ala His Trp Asp Ala Ser Val Ser
          1140                      1145                      1150
Gln Thr Ile Glu Ile Thr Asp Phe Glu Glu Glu Lys Glu Tyr Lys Leu
          1155                      1160                      1165
Arg Val Arg Gly Lys Gly Lys Gly Thr Val Thr Val Gln His Glu Glu
          1170                      1175                      1180
Glu Leu Glu Thr Met Thr Phe Asn Thr Thr Ser Phe Thr Thr Lys Glu
1185          1190                      1195                      1200
Gln Thr Phe Tyr Phe Glu Gly Asn Thr Ile Asp Val His Val Gln Ser
          1205                      1210                      1215
Glu Asn Asn Ala Phe Leu Val Asp Ser Val Glu Leu Ile Glu Val Val
          1220                      1225                      1230
Lys Glu Gln Glu Glu Lys Gln
          1235

```

<210> SEQ ID NO: 7

<211> Comprimento: 2059

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: synaxmi-031

<221> Nome/Chave: CDS

<222> Localização: (1)...(2058)

<400> Sequência: 7

atg gat tgc aat ctc cag agc cag cag aat atc cca tac aat gtg ctt	48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu	
1 5 10 15	
gct atc cct gtt agc aat gtt aat agc ctt acg gat acg gtt ggc gat	96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp	
20 25 30	
cta aag aag gca tgg gaa gag ttc cag aag acg ggc agc ttc tcc ctt	144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu	
35 40 45	
acc gca ctc cag cag ggc ttc agc gct tcc cag ggc ggg acg ttc aat	192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn	
50 55 60	
tac ctc acc ctc ctc cag tcc ggg atc tcc ctg gct ggc agc ttc gtg	240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val	
65 70 75 80	
cct ggc ggc acc ttc gtt gct cca atc atc aat atg gtg atc ggc tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	
ctc tgg cct cat aag aat aag aat gca gat acg gaa aat ctg atc aat	336
Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn	
100 105 110	
ctg atc gat tca gaa atc cag aag cag ctg aat aag gct ctg ctg gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gct gat cgc aat gaa tgg agc agc tac ctc gag tcc atc ttc gat agc	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
agc aat aat ctg aat ggc gca atc gtt gat gct cag tgg tcc ggc acc	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gtg aat acg acc aat cgt acg ctt cgt aat cct acg gaa tca gat tac	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	
acc aat gtg gtg acg aat ttc atc gca gca gat ggg gat atc gct aat	576
Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cat atc atg aat ggg aat ttc gat gtg gca gca gct cca	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	
tac ttc gtt atc ggg gct acc gct aga ttc gca gct atg cag tcc tac	672
Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr	
210 215 220	

atc aag ttc tgc aat gca tgg atc gat aag gtt ggg ctg tca gat gct Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala 225 230 235 240	720
cag ctg acc acg cag aag gct aat ctg gat aga acg aag cag aat atg Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met 245 250 255	768
cgt aat gct atc ctt aat tac acc cag cag gtt atg aag gtt ttc aag Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys 260 265 270	816
gat tcc aag aat atg cca acg atc ggc acg aat aag ttc agc gtt gat Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp 275 280 285	864
acc tac aat gtt tac atc aag ggg atg acg ctt aat gtg ctt gat atc Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile 290 295 300	912
gtt gct atc tgg cca agc ctg tac cca gat gat tac acg tca cag acg Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr 305 310 315 320	960
gct ctg gaa cag acc cgc gtg acg ttc tcc aat atg gtg ggg cag gaa Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu 325 330 335	1008
gaa ggc acg gat ggc agc ctc aga atc tac aat acc ttc gat agc ttc Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe 340 345 350	1056
tcc tac cag cat agc cct atc cct aat aat aat gtg aat ctc atc agc Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser 355 360 365	1104
tac tac aat gat gaa ctt cag aat ctg gaa ctc ggg gtt tac acc cca Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro 370 375 380	1152
cca aag aag ggc tca ggc tac agc tac cca tac ggg ttc gtg ctg aat Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn 385 390 395 400	1200
tac gca aat agc aag tac aag tac ggc gat tcc aat gat cca gaa tcc Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser 405 410 415	1248
ctc ggc ggg ctt tcc acc ctt agc gct cca atc caa cag gtt aat gct Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala 420 425 430	1296
gct acc cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg atc Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile 435 440 445	1344

ggc gca agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc gaa	1392
Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu	
450 455 460	
cca cca ttc agc tgc acg agc acg gca aat ggg tac aag gca agc tgc	1440
Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys	
465 470 475 480	
aat cca agc gat acc aat cag aag atc aat gct ctc tac cca ttc acg	1488
Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr	
485 490 495	
cag gca aat gtg aag ggg aat acc ggg aag ctg ggc gtt ctc gca agc	1536
Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser	
500 505 510	
ctt gtt agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc gat	1584
Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp	
515 520 525	
tca gat acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag ggc	1632
Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly	
530 535 540	
tac ttc cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc aat	1680
Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn	
545 550 555 560	
ggg gct tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg acg	1728
Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr	
565 570 575	
gct acg aac ctg acg gct acc cag tac aga atc cgc atc cgt tac gct	1776
Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala	
580 585 590	
aac cca aac tca aac acg cag atc ggc gtt aga atc acg cag aac ggg	1824
Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly	
595 600 605	
agc ctc atc tcc tcc tcc aac ctc acg ctg tac tca acc acc gat atg	1872
Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met	
610 615 620	
aac aac acc ctg cca ctg aac gtt tac gtg atc ggg gaa aac ggg aac	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tac acc ctc cag gat ctc tac aac acg acg aac gtt ctg tcc acg ggc	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	
gat atc acc ctg cag atc acc ggc ggg gat cag aag ata ttc atc gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	

cgc atc gag ttc gtg cct acg atg cca gtg cca ggc aac taa 2058
 Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn *
 675 680 685

a 2059

<210> SEQ ID NO: 8
 <211> Comprimento: 685
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: SYNAXMI-031

<400> Sequência: 8
 Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu
 1 5 10 15
 Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp
 20 25 30
 Leu Lys Lys Ala Trp Glu Glu Phe Ser Gln Lys Thr Gly Ser Phe Ser Leu
 35 40 45
 Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn
 50 55 60
 Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val
 65 70 75 80
 Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp
 85 90 95
 Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn
 100 105 110
 Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp
 115 120 125
 Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser
 130 135 140
 Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr
 145 150 155 160
 Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr
 165 170 175
 Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn
 180 185 190
 Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro
 195 200 205
 Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr
 210 215 220
 Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala
 225 230 235 240
 Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met
 245 250 255
 Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys
 260 265 270
 Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp
 275 280 285
 Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile
 290 295 300
 Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr
 305 310 315 320
 Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu
 325 330 335

Glu	Gly	Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	Phe	340	345	350
Ser	Tyr	Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	Ser	355	360	365
Tyr	Tyr	Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr	Pro	370	375	380
Pro	Lys	Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu	Asn	385	390	395
Tyr	Ala	Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu	Ser	405	410	415
Leu	Gly	Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn	Ala	420	425	430
Ala	Thr	Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	Asn	Gly	Ile	435	440	445
Gly	Ala	Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Pro	Thr	Glu	450	455	460
Pro	Pro	Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	Ala	Ser	Cys	465	470	475
Asn	Pro	Ser	Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Pro	Phe	Thr	485	490	495
Gln	Ala	Asn	Val	Lys	Gly	Asn	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala	Ser	500	505	510
Leu	Val	Ser	Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu	Asp	515	520	525
Ser	Asp	Thr	Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys	Gly	530	535	540
Tyr	Phe	Pro	Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile	Asn	545	550	555
Gly	Ala	Ser	Ala	Val	Pro	Leu	Asp	Ser	Gly	Asn	Thr	Leu	Phe	Met	Thr	565	570	575
Ala	Thr	Asn	Leu	Thr	Ala	Thr	Gln	Tyr	Arg	Ile	Arg	Ile	Arg	Tyr	Ala	580	585	590
Asn	Pro	Asn	Ser	Asn	Thr	Gln	Ile	Gly	Val	Arg	Ile	Thr	Gln	Asn	Gly	595	600	605
Ser	Leu	Ile	Ser	Ser	Ser	Asn	Leu	Thr	Leu	Tyr	Ser	Thr	Thr	Asp	Met	610	615	620
Asn	Asn	Thr	Leu	Pro	Leu	Asn	Val	Tyr	Val	Ile	Gly	Glu	Asn	Gly	Asn	625	630	635
Tyr	Thr	Leu	Gln	Asp	Leu	Tyr	Asn	Thr	Thr	Asn	Val	Leu	Ser	Thr	Gly	645	650	655
Asp	Ile	Thr	Leu	Gln	Ile	Thr	Gly	Gly	Asp	Gln	Lys	Ile	Phe	Ile	Asp	660	665	670
Arg	Ile	Glu	Phe	Val	Pro	Thr	Met	Pro	Val	Pro	Gly	Asn				675	680	685

<210> SEQ ID NO: 9

<211> Comprimento: 2148

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: synaxmi-031(apo)

<221> Nome/Chave: CDS

<222> Localização: (1)...(2148)

<400> Sequência: 9

atg ggg tac tcc agc ttc gtt gct atc gct ctt ctt atg agc gtg gtt	48
Met Gly Tyr Ser Ser Phe Val Ala Ile Ala Leu Leu Met Ser Val Val	
1 5 10 15	
gtt gtt tgc aat ggg ggc aag acg tcc acc tac gtg cgt aat ctg gat	96
Val Val Cys Asn Gly Gly Lys Thr Ser Thr Tyr Val Arg Asn Leu Asp	
20 25 30	
tgc aat ctc cag agc cag cag aat atc cca tac aat gtg ctt gct atc	144
Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu Ala Ile	
35 40 45	
cct gtt agc aat gtt aat agc ctt acg gat acg gtt ggc gat cta aag	192
Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp Leu Lys	
50 55 60	
aag gca tgg gaa gag ttc cag aag acg ggc agc ttc tcc ctt acc gca	240
Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu Thr Ala	
65 70 75 80	
ctc cag cag ggc ttc agc gct tcc cag ggc ggg acg ttc aat tac ctc	288
Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn Tyr Leu	
85 90 95	
acc ctc ctc cag tcc ggg atc tcc ctg gct ggc agc ttc gtg cct ggc	336
Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val Pro Gly	
100 105 110	
ggc acc ttc gtt gct cca atc atc aat atg gtg atc ggc tgg ctc tgg	384
Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp Leu Trp	
115 120 125	
cct cat aag aat aag aat gca gat acg gaa aat ctg atc aat ctg atc	432
Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn Leu Ile	
130 135 140	
gat tca gaa atc cag aag cag ctg aat aag gct ctg ctg gat gct gat	480
Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp Ala Asp	
145 150 155 160	
cgc aat gaa tgg agc agc tac ctc gag tcc atc ttc gat agc agc aat	528
Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser Ser Asn	
165 170 175	
aat ctg aat ggc gca atc gtt gat gct cag tgg tcc ggc acc gtg aat	576
Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr Val Asn	
180 185 190	
acg acc aat cgt acg ctt cgt aat cct acg gaa tca gat tac acc aat	624
Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr Thr Asn	
195 200 205	
gtg gtg acg aat ttc atc gca gca gat ggg gat atc gct aat aat gaa	672
Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn Asn Glu	
210 215 220	

aat cat atc atg aat ggg aat ttc gat gtg gca gca gct cca tac ttc	720
Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro Tyr Phe	
225 230 235 240	
gtt atc ggg gct acc gct aga ttc gca gct atg cag tcc tac atc aag	768
Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr Ile Lys	
245 250 255	
ttc tgc aat gca tgg atc gat aag gtt ggg ctg tca gat gct cag ctg	816
Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala Gln Leu	
260 265 270	
acc acg cag aag gct aat ctg gat aga acg aag cag aat atg cgt aat	864
Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met Arg Asn	
275 280 285	
gct atc ctt aat tac acc cag cag gtt atg aag gtt ttc aag gat tcc	912
Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys Asp Ser	
290 295 300	
aag aat atg cca acg atc ggc acg aat aag ttc agc gtt gat acc tac	960
Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp Thr Tyr	
305 310 315 320	
aat gtt tac atc aag ggg atg acg ctt aat gtg ctt gat atc gtt gct	1008
Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile Val Ala	
325 330 335	
atc tgg cca agc ctg tac cca gat gat tac acg tca cag acg gct ctg	1056
Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr Ala Leu	
340 345 350	
gaa cag acc cgc gtg acg ttc tcc aat atg gtg ggg cag gaa gaa ggc	1104
Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu Glu Gly	
355 360 365	
acg gat ggc agc ctc aga atc tac aat acc ttc gat agc ttc tcc tac	1152
Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe Ser Tyr	
370 375 380	
cag cat agc cct atc cct aat aat aat gtg aat ctc atc agc tac tac	1200
Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser Tyr Tyr	
385 390 395 400	
aat gat gaa ctt cag aat ctg gaa ctc ggg gtt tac acc cca cca aag	1248
Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro Pro Lys	
405 410 415	
aag ggc tca ggc tac agc tac cca tac ggg ttc gtg ctg aat tac gca	1296
Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn Tyr Ala	
420 425 430	
aat agc aag tac aag tac ggc gat tcc aat gat cca gaa tcc ctc ggc	1344
Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser Leu Gly	
435 440 445	

ggg ctt tcc acc ctt agc gct cca atc caa cag gtt aat gct gct acc	1392
Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala Ala Thr	
450 455 460	
cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg atc ggc gca	1440
Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile Gly Ala	
465 470 475 480	
agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc gaa cca cca	1488
Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu Pro Pro	
485 490 495	
ttc agc tgc acg agc acg gca aat ggg tac aag gca agc tgc aat cca	1536
Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys Asn Pro	
500 505 510	
agc gat acc aat cag aag atc aat gct ctc tac cca ttc acg cag gca	1584
Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr Gln Ala	
515 520 525	
aat gtg aag ggg aat acc ggg aag ctg ggc gtt ctc gca agc ctt gtt	1632
Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser Leu Val	
530 535 540	
agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc gat tca gat	1680
Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp Ser Asp	
545 550 555 560	
acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag ggc tac ttc	1728
Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly Tyr Phe	
565 570 575	
cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc aat ggg gct	1776
Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn Gly Ala	
580 585 590	
tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg acg gct acg	1824
Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr Ala Thr	
595 600 605	
aac ctg acg gct acc cag tac aga atc cgc atc cgt tac gct aac cca	1872
Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala Asn Pro	
610 615 620	
aac tca aac acg cag atc ggc gtt aga atc acg cag aac ggg agc ctc	1920
Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly Ser Leu	
625 630 635 640	
atc tcc tcc tcc aac ctc acg ctg tac tca acc acc gat atg aac aac	1968
Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met Asn Asn	
645 650 655	
acc ctg cca ctg aac gtt tac gtg atc ggg gaa aac ggg aac tac acc	2016
Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn Tyr Thr	
660 665 670	

ctc cag gat ctc tac aac acg acg aac gtt ctg tcc acg ggc gat atc 2064
 Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly Asp Ile
 675 680 685

acc ctg cag atc acc ggc ggg gat cag aag ata ttc atc gat cgc atc 2112
 Thr Leu Gln ile Thr Gly Gly Asp Gln Lys ile Phe ile Asp Arg ile
 690 695 700

gag ttc gtg cct acg atg cca gtg cca ggc aac taa 2148
 Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn *
 705 710 715

<210> SEQ ID NO: 10

<211> Comprimento: 715

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: SYNAXMI-031(APO)

<400> Sequência: 10

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

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

<210> SEQ ID NO: 11
 <211> Comprimento: 2160
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: synaxmi-031(ER)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(2160)

<400> Sequência: 11

atg ggg tac tcc agc ttc gtt gct atc gct ctt ctt atg agc gtg gtt	48
Met Gly Tyr Ser Ser Phe Val Ala Ile Ala Leu Leu Met Ser Val Val	
1 5 10 15	
gtt gtt tgc aat ggg ggc aag acg tcc acc tac gtg cgt aat ctg gat	96
Val Val Cys Asn Gly Gly Lys Thr Ser Thr Tyr Val Arg Asn Leu Asp	
20 25 30	
tgc aat ctc cag agc cag cag aat atc cca tac aat gtg ctt gct atc	144
Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu Ala Ile	
35 40 45	
cct gtt agc aat gtt aat agc ctt acg gat acg gtt ggc gat cta aag	192
Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp Leu Lys	
50 55 60	
aag gca tgg gaa gag ttc cag aag acg ggc agc ttc tcc ctt acc gca	240
Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu Thr Ala	
65 70 75 80	
ctc cag cag ggc ttc agc gct tcc cag ggc ggg acg ttc aat tac ctc	288
Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn Tyr Leu	
85 90 95	
acc ctc ctc cag tcc ggg atc tcc ctg gct ggc agc ttc gtg cct ggc	336
Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val Pro Gly	
100 105 110	
ggc acc ttc gtt gct cca atc atc aat atg gtg atc ggc tgg ctc tgg	384
Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp Leu Trp	
115 120 125	
cct cat aag aat aag aat gca gat acg gaa aat ctg atc aat ctg atc	432
Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn Leu Ile	
130 135 140	
gat tca gaa atc cag aag cag ctg aat aag gct ctg ctg gat gct gat	480
Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp Ala Asp	
145 150 155 160	
cgc aat gaa tgg agc agc tac ctc gag tcc atc ttc gat agc agc aat	528
Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser Ser Asn	
165 170 175	

aat ctg aat ggc gca atc gtt gat gct cag tgg tcc ggc acc gtg aat	576
Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr Val Asn	
180 185 190	
acg acc aat cgt acg ctt cgt aat cct acg gaa tca gat tac acc aat	624
Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr Thr Asn	
195 200 205	
gtg gtg acg aat ttc atc gca gca gat ggg gat atc gct aat aat gaa	672
Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn Asn Glu	
210 215 220	
aat cat atc atg aat ggg aat ttc gat gtg gca gca gct cca tac ttc	720
Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro Tyr Phe	
225 230 235 240	
gtt atc ggg gct acc gct aga ttc gca gct atg cag tcc tac atc aag	768
Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr Ile Lys	
245 250 255	
ttc tgc aat gca tgg atc gat aag gtt ggg ctg tca gat gct cag ctg	816
Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala Gln Leu	
260 265 270	
acc acg cag aag gct aat ctg gat aga acg aag cag aat atg cgt aat	864
Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met Arg Asn	
275 280 285	
gct atc ctt aat tac acc cag cag gtt atg aag gtt ttc aag gat tcc	912
Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys Asp Ser	
290 295 300	
aag aat atg cca acg atc ggc acg aat aag ttc agc gtt gat acc tac	960
Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp Thr Tyr	
305 310 315 320	
aat gtt tac atc aag ggg atg acg ctt aat gtg ctt gat atc gtt gct	1008
Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile Val Ala	
325 330 335	
atc tgg cca agc ctg tac cca gat gat tac acg tca cag acg gct ctg	1056
Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr Ala Leu	
340 345 350	
gaa cag acc cgc gtg acg ttc tcc aat atg gtg ggg cag gaa gaa ggc	1104
Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu Glu Gly	
355 360 365	
acg gat ggc agc ctc aga atc tac aat acc ttc gat agc ttc tcc tac	1152
Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe Ser Tyr	
370 375 380	
cag cat agc cct atc cct aat aat aat gtg aat ctc atc agc tac tac	1200
Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser Tyr Tyr	
385 390 395 400	

aat gat gaa ctt cag aat ctg gaa ctc ggg gtt tac acc cca cca aag	1248
Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro Pro Lys	
405 410 415	
aag ggc tca ggc tac agc tac cca tac ggg ttc gtg ctg aat tac gca	1296
Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn Tyr Ala	
420 425 430	
aat agc aag tac aag tac ggc gat tcc aat gat cca gaa tcc ctc ggc	1344
Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser Leu Gly	
435 440 445	
ggg ctt tcc acc ctt agc gct cca atc caa cag gtt aat gct gct acc	1392
Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala Ala Thr	
450 455 460	
cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg atc ggc gca	1440
Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile Gly Ala	
465 470 475 480	
agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc gaa cca cca	1488
Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu Pro Pro	
485 490 495	
ttc agc tgc acg agc acg gca aat ggg tac aag gca agc tgc aat cca	1536
Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys Asn Pro	
500 505 510	
agc gat acc aat cag aag atc aat gct ctc tac cca ttc acg cag gca	1584
Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr Gln Ala	
515 520 525	
aat gtg aag ggg aat acc ggg aag ctg ggc gtt ctc gca agc ctt gtt	1632
Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser Leu Val	
530 535 540	
agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc gat tca gat	1680
Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp Ser Asp	
545 550 555 560	
acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag ggc tac ttc	1728
Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly Tyr Phe	
565 570 575	
cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc aat ggg gct	1776
Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn Gly Ala	
580 585 590	
tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg acg gct acg	1824
Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr Ala Thr	
595 600 605	
aac ctg acg gct acc cag tac aga atc cgc atc cgt tac gct aac cca	1872
Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala Asn Pro	
610 615 620	

aac Asn 625	tca Ser	aac Asn	acg Thr	cag Gln 645	atc Ile 630	ggc Gly	gtt Val	aga Arg	atc Ile 635	acg Thr	cag Gln	aac Asn	ggg Gly	agc Ser 640	ctc Leu	1920
atc Ile	tcc Ser	tcc Ser	tcc Ser	aac Asn 645	ctc Leu	acg Thr	ctg Leu	tac Tyr	tca Ser 650	acc Thr	acc Thr	gat Asp	atg Met	aac Asn 655	aac Asn	1968
acc Thr	ctg Leu	cca Pro	ctg Leu 660	aac Asn	gtt Val	tac Tyr	gtg Val	atc Ile 665	ggg Gly	gaa Glu	aac Asn	ggg Gly	aac Asn 670	tac Tyr	acc Thr	2016
ctc Leu	cag Gln	gat Asp 675	ctc Leu	tac Tyr	aac Asn	acg Thr	acg Thr 680	aac Asn	gtt Val	ctg Leu	tcc Ser	acg Thr 685	ggc Gly	gat Asp	atc Ile	2064
acc Thr 690	ctg Leu	cag Gln	atc Ile	acc Thr	ggc Gly	ggg Gly 695	gat Asp	cag Gln	aag Lys	ata Ile	ttc Phe 700	atc Ile	gat Asp	cgc Arg	atc Ile	2112
gag Glu 705	ttc Phe	gtg Val	cct Pro	acg Thr	atg Met 710	cca Pro	gtg Val	cca Pro	ggc Gly 715	aac Asn	aag Lys	gat Asp	gaa Glu	ctg Leu	taa *	2160

```
<210> SEQ ID NO: 12
<211> Comprimento: 719
<212> Tipo: PRT
<213> Organismo: Sequência Artificial

<220> Características:
<223> Outras Informações: SYNAXMI-031(ER)
```

<400> Seqüência: 12

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

Asn	Leu	Asn	Gly	Ala	Ile	Val	Asp	Ala	Gln	Trp	Ser	Gly	Thr	Val	Asn
			180					185					190		
Thr	Thr	Asn	Arg	Thr	Leu	Arg	Asn	Pro	Thr	Glu	Ser	Asp	Tyr	Thr	Asn
		195					200					205			
Val	Val	Thr	Asn	Phe	Ile	Ala	Ala	Asp	Gly	Asp	Ile	Ala	Asn	Asn	Glu
	210					215					220				
Asn	His	Ile	Met	Asn	Gly	Asn	Phe	Asp	Val	Ala	Ala	Ala	Pro	Tyr	Phe
225					230					235					240
Val	Ile	Gly	Ala	Thr	Ala	Arg	Phe	Ala	Ala	Met	Gln	Ser	Tyr	Ile	Lys
				245					250					255	
Phe	Cys	Asn	Ala	Trp	Ile	Asp	Lys	Val	Gly	Leu	Ser	Asp	Ala	Gln	Leu
		260						265					270		
Thr	Thr	Gln	Lys	Ala	Asn	Leu	Asp	Arg	Thr	Lys	Gln	Asn	Met	Arg	Asn
		275					280					285			
Ala	Ile	Leu	Asn	Tyr	Thr	Gln	Gln	Val	Met	Lys	Val	Phe	Lys	Asp	Ser
	290					295					300				
Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	Ser	Val	Asp	Thr	Tyr
305					310					315					320
Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp	Ile	Val	Ala
			325						330					335	
Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	Ser	Gln	Thr	Ala	Leu
			340					345					350		
Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln	Glu	Glu	Gly
		355					360					365			
Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	Phe	Ser	Tyr
	370					375					380				
Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	Ser	Tyr	Tyr
385					390					395					400
Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr	Pro	Pro	Lys
				405					410					415	
Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu	Asn	Tyr	Ala
			420					425					430		
Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu	Ser	Leu	Gly
		435					440					445			
Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn	Ala	Ala	Thr
	450					455					460				
Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	Asn	Gly	Ile	Gly	Ala
465					470					475					480
Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Pro	Thr	Glu	Pro	Pro
				485					490					495	
Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	Ala	Ser	Cys	Asn	Pro
			500					505					510		
Ser	Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Pro	Phe	Thr	Gln	Ala
		515					520					525			
Asn	Val	Lys	Gly	Asn	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala	Ser	Leu	Val
	530					535					540				
Ser	Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu	Asp	Ser	Asp
545					550					555					560
Thr	Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys	Gly	Tyr	Phe
				565					570					575	
Pro	Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile	Asn	Gly	Ala
			580					585					590		
Ser	Ala	Val	Pro	Leu	Asp	Ser	Gly	Asn	Thr	Leu	Phe	Met	Thr	Ala	Thr
		595					600					605			
Asn	Leu	Thr	Ala	Thr	Gln	Tyr	Arg	Ile	Arg	Ile	Arg	Tyr	Ala	Asn	Pro
	610					615					620				

```

Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly Ser Leu
625                               630                               635                               640
Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met Asn Asn
                               645                               650                               655
Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn Tyr Thr
                               660                               665                               670
Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly Asp Ile
                               675                               680                               685
Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp Arg Ile
690                               695                               700
Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Lys Asp Glu Leu
705                               710                               715

```

<210> SEQ ID NO: 13

<211> Comprimento: 1186

<212> Tipo: PRT

<213> Organismo: *Bacillus thuringiensis*

<400> Sequência: 13

```

Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu
1      5      10      15
Ala Ile Pro Val Ser Asn Val Asn Ala Leu Val Asp Thr Ala Gly Asp
20      25      30
Leu Lys Lys Ala Trp Glu Glu Phe Ser Ala Ser Gln Gly Gly Ala Phe Ser Leu
35      40      45
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Ala Phe Asn
50      55      60
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val
65      70      75      80
Pro Gly Gly Thr Phe Val Ala Pro Ile Val Asn Met Val Ile Gly Trp
85      90      95
Leu Trp Pro His Lys Asn Lys Thr Ala Asp Thr Glu Asn Leu Ile Lys
100     105     110
Leu Ile Asp Glu Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp
115     120     125
Gln Asp Arg Asn Asn Trp Thr Ser Phe Leu Glu Ser Ile Phe Asp Thr
130     135     140
Ser Ala Thr Val Ser Asn Ala Ile Ile Asp Ala Gln Trp Ser Gly Thr
145     150     155     160
Val Asp Thr Thr Asn Arg Gln Gln Lys Thr Pro Thr Thr Ser Asp Tyr
165     170     175
Leu Asn Val Val Gly Lys Phe Asp Ser Ala Asp Ser Ser Ile Ile Thr
180     185     190
Asn Glu Asn Gln Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro
195     200     205
Tyr Phe Val Ile Gly Ala Thr Leu Arg Leu Ser Leu Tyr Gln Ser Tyr
210     215     220
Ile Lys Phe Cys Asn Ser Trp Ile Asp Ala Val Gly Phe Ser Thr Asn
225     230     235     240
Asp Ala Asn Thr Gln Lys Ala Asn Leu Ala Arg Thr Lys Leu Thr Met
245     250     255
Arg Thr Thr Ile Asn Glu Tyr Thr Gln Arg Val Met Lys Val Phe Lys
260     265     270
Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp
275     280     285

```

Ala	Tyr	Asn	Val	Tyr	Val	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp	Met
290						295				300					
Val	Ala	Ile	Trp	Ser	Ser	Leu	Tyr	Pro	Asn	Asp	Tyr	Thr	Ser	Gln	Thr
305					310					315					320
Ala	Ile	Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln	Glu
				325					330					335	
Glu	Gly	Thr	Asp	Gly	Thr	Leu	Lys	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	Leu
			340					345					350		
Ser	Tyr	Gln	His	Ser	Leu	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	Ser
		355					360					365			
Tyr	Tyr	Thr	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Ala	Val	Tyr	Thr	Pro
		370				375					380				
Lys	Gly	Gly	Ser	Gly	Tyr	Ala	Tyr	Pro	Tyr	Gly	Phe	Ile	Leu	Asn	Tyr
385					390					395					400
Ala	Asn	Ser	Asn	Tyr	Lys	Tyr	Gly	Asp	Asn	Asp	Pro	Thr	Gly	Lys	Pro
				405					410					415	
Leu	Asn	Lys	Gln	Asp	Gly	Pro	Ile	Gln	Gln	Ile	Asn	Ala	Ala	Thr	Gln
			420					425					430		
Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Thr	Ile	Asn	Gly	Ile	Gly	Ala	Ser
		435					440					445			
Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Ala	Thr	Glu	Gln	Pro	Phe
		450				455					460				
Ser	Cys	Thr	Ser	Thr	Ala	Asn	Ser	Tyr	Lys	Ala	Ser	Cys	Asn	Pro	Ser
465					470					475					480
Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Ala	Phe	Thr	Gln	Thr	Asn
				485					490					495	
Val	Lys	Gly	Ser	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala	Ser	Leu	Val	Pro
			500					505					510		
Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu	Asp	Ser	Asp	Thr
		515					520					525			
Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys	Gly	Tyr	Phe	Pro
		530				535					540				
Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile	Asn	Gly	Ala	Ser
545					550					555					560
Ala	Val	Pro	Phe	Tyr	Ser	Gly	Asn	Thr	Leu	Phe	Met	Thr	Ala	Thr	Asn
				565					570					575	
Leu	Thr	Ala	Thr	Gln	Tyr	Lys	Ile	Arg	Ile	Arg	Tyr	Ala	Asn	Pro	Asn
			580					585					590		
Ser	Asp	Thr	Gln	Ile	Gly	Val	Leu	Ile	Thr	Gln	Asn	Gly	Ser	Gln	Ile
		595					600					605			
Ser	Asn	Ser	Asn	Leu	Thr	Leu	Tyr	Ser	Thr	Thr	Asp	Ser	Ser	Met	Ser
		610				615					620				
Ser	Asn	Leu	Pro	Gln	Asn	Val	Tyr	Val	Thr	Gly	Glu	Asn	Gly	Asn	Tyr
625					630					635					640
Thr															

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

Asp Glu
1185

<210> SEQ ID NO: 14

<211> Comprimento: 2190

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: axmi-031(truncado)

<221> Nome/Chave: CDS

<222> Localização: (1)...(2190)

<400> Sequência: 14

atg	tcg	tac	tac	cat	cac	cat	cac	cat	cac	ctc	gaa	tca	aca	agt	ttg	48
Met	Ser	Tyr	Tyr	His	His	His	His	His	His	Leu	Glu	Ser	Thr	Ser	Leu	
1				5					10					15		
tac	aaa	aaa	gca	ggc	ttc	att	gaa	gga	cgt	atg	gat	tgt	aat	tta	caa	96
Tyr	Lys	Lys	Ala	Gly	Phe	Ile	Glu	Gly	Arg	Met	Asp	Cys	Asn	Leu	Gln	
			20					25					30			
tca	caa	caa	aat	att	cca	tat	aat	gta	tta	gca	ata	cca	gta	tct	aat	144
Ser	Gln	Gln	Asn	Ile	Pro	Tyr	Asn	Val	Leu	Ala	Ile	Pro	Val	Ser	Asn	
			35				40					45				
gtt	aat	tcg	ttg	act	gat	aca	gtt	gga	gat	tta	aaa	aaa	gca	tgg	gaa	192
Val	Asn	Ser	Leu	Thr	Asp	Thr	Val	Gly	Asp	Leu	Lys	Lys	Ala	Trp	Glu	
	50					55					60					
gaa	ttt	caa	aaa	act	ggg	tct	ttt	tca	tta	aca	gct	tta	caa	caa	gga	240
Glu	Phe	Gln	Lys	Thr	Gly	Ser	Phe	Ser	Leu	Thr	Ala	Leu	Gln	Gln	Gly	
	65				70				75						80	
ttt	tct	gct	tca	caa	gga	gga	aca	ttc	aat	tat	tta	aca	tta	cta	caa	288
Phe	Ser	Ala	Ser	Gln	Gly	Gly	Thr	Phe	Asn	Tyr	Leu	Thr	Leu	Leu	Gln	
				85				90						95		
tca	gga	ata	tca	tta	gct	ggg	tct	ttt	gtt	cct	gga	ggg	act	ttt	gta	336
Ser	Gly	Ile	Ser	Leu	Ala	Gly	Ser	Phe	Val	Pro	Gly	Gly	Thr	Phe	Val	
			100					105					110			
gca	cct	att	att	aat	atg	gtt	att	ggg	tgg	tta	tgg	cca	cat	aaa	aac	384
Ala	Pro	Ile	Ile	Asn	Met	Val	Ile	Gly	Trp	Leu	Trp	Pro	His	Lys	Asn	
			115				120					125				
aaa	aat	gcg	gat	aca	gaa	aat	tta	ata	aat	tta	att	gat	tca	gaa	att	432
Lys	Asn	Ala	Asp	Thr	Glu	Asn	Leu	Ile	Asn	Leu	Ile	Asp	Ser	Glu	Ile	
			130				135				140					
caa	aaa	caa	tta	aac	aaa	gct	tta	tta	gat	gca	gat	aga	aat	gag	tgg	480
Gln	Lys	Gln	Leu	Asn	Lys	Ala	Leu	Leu	Asp	Ala	Asp	Arg	Asn	Glu	Trp	
					145		150			155					160	

agc tct tat tta gaa tct ata ttt gat tct tca aat aac cta aat ggt	528
Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser Ser Asn Asn Leu Asn Gly	
165 170 175	
gca att gta gat gca cag tgg tca ggc act gta aat act aca aat aga	576
Ala Ile Val Asp Ala Gln Trp Ser Gly Thr Val Asn Thr Thr Asn Arg	
180 185 190	
aca cta aga aat cca aca gaa tca gat tat aca aat gtt gtt aca aat	624
Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr Thr Asn Val Val Thr Asn	
195 200 205	
ttt att gca gcg gat ggt gac att gca aat aat gaa aat cac ata atg	672
Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn Asn Glu Asn His Ile Met	
210 215 220	
aat ggc aac ttt gac gta gct gca gca cct tat ttt gtt ata gga gca	720
Asn Gly Asn Phe Asp Val Ala Ala Ala Pro Tyr Phe Val Ile Gly Ala	
225 230 235 240	
aca gca cgt ttt gca gca atg caa tct tat att aaa ttt tgt aat gct	768
Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr Ile Lys Phe Cys Asn Ala	
245 250 255	
tgg att gat aaa gtt gga ttg agt gac gca cag ctt act aca caa aag	816
Trp Ile Asp Lys Val Gly Leu Ser Asp Ala Gln Leu Thr Thr Gln Lys	
260 265 270	
gct aat tta gat cgc acg aaa caa aat atg cgt aat gca att ctt aac	864
Ala Asn Leu Asp Arg Thr Lys Gln Asn Met Arg Asn Ala Ile Leu Asn	
275 280 285	
tat aca caa caa gtt atg aaa gtt ttt aaa gat tcc aaa aat atg cct	912
Tyr Thr Gln Gln Val Met Lys Val Phe Lys Asp Ser Lys Asn Met Pro	
290 295 300	
aca ata ggt act aat aaa ttt agt gtt gat acc tat aat gta tat att	960
Thr Ile Gly Thr Asn Lys Phe Ser Val Asp Thr Tyr Asn Val Tyr Ile	
305 310 315 320	
aaa gga atg aca tta aat gtt tta gat att gta gca ata tgg cct tca	1008
Lys Gly Met Thr Leu Asn Val Leu Asp Ile Val Ala Ile Trp Pro Ser	
325 330 335	
tta tat cca gat gat tat act tca caa aca gcc tta gaa caa aca cgt	1056
Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr Ala Leu Glu Gln Thr Arg	
340 345 350	
gtc act ttt tca aat atg gtt ggc caa gaa gaa ggt aca gat gga agc	1104
Val Thr Phe Ser Asn Met Val Gly Gln Glu Glu Gly Thr Asp Gly Ser	
355 360 365	
cta aga att tac aat act ttt gat tct ttt agt tat caa cat agt cca	1152
Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe Ser Tyr Gln His Ser Pro	
370 375 380	

ata cct aat aat aat gtt aat tta att tct tat tat aat gat gaa tta	1200
Ile Pro Asn Asn Asn Val Asn Leu Ile Ser Tyr Tyr Asn Asp Glu Leu	
385 390 395 400	
caa aat cta gaa tta gga gta tat acc cct cct aaa aaa gga agt gga	1248
Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro Pro Lys Lys Gly Ser Gly	
405 410 415	
tac tct tat cct tat gga ttt gtt tta aat tat gca aac agt aaa tat	1296
Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn Tyr Ala Asn Ser Lys Tyr	
420 425 430	
aaa tat ggt gat agc aat gat cca gaa tct cta gga gga tta tct aca	1344
Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser Leu Gly Gly Leu Ser Thr	
435 440 445	
cta tct gca cct ata caa caa gtt aat gca gca act caa aac agt aaa	1392
Leu Ser Ala Pro Ile Gln Gln Val Asn Ala Ala Thr Gln Asn Ser Lys	
450 455 460	
tat cta gat gga gaa atc cta aat gga ata gga gca tcc tta cct ggt	1440
Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile Gly Ala Ser Leu Pro Gly	
465 470 475 480	
tat tgt act aca gga tgt tca cca aca gaa cca cct ttt agt tgt act	1488
Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu Pro Pro Phe Ser Cys Thr	
485 490 495	
tct acc gct aat ggc tat aaa gca agc tgt aat cct tca gat aca aat	1536
Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys Asn Pro Ser Asp Thr Asn	
500 505 510	
caa aaa att aac gct tta tat cct ttt aca caa gct aat gta aag gga	1584
Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr Gln Ala Asn Val Lys Gly	
515 520 525	
aac aca gga aaa tta gga gta ctg gca agt ctt gtt tca tat gat tta	1632
Asn Thr Gly Lys Leu Gly Val Leu Ala Ser Leu Val Ser Tyr Asp Leu	
530 535 540	
aat cct aaa aat gta ttt ggt gaa tta gat tca gat aca aat aat gtt	1680
Asn Pro Lys Asn Val Phe Gly Glu Leu Asp Ser Asp Thr Asn Asn Val	
545 550 555 560	
atc tta aaa gga att cct gca gaa aaa gga tat ttt cct aat aat gcg	1728
Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly Tyr Phe Pro Asn Asn Ala	
565 570 575	
cgt cct act gtt gta aaa gaa tgg att aat ggt gca agt gct gta cca	1776
Arg Pro Thr Val Val Lys Glu Trp Ile Asn Gly Ala Ser Ala Val Pro	
580 585 590	
ctt gat tca gga aat acc tta ttt atg acg gct acg aat tta aca gct	1824
Leu Asp Ser Gly Asn Thr Leu Phe Met Thr Ala Thr Asn Leu Thr Ala	
595 600 605	

```

act caa tat aga att aga ata cgt tat gca aat cca aat tca aat act 1872
Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala Asn Pro Asn Ser Asn Thr
610 615 620

caa atc ggt gta cga att aca caa aat ggt tct cta att tcc agt agt 1920
Gln Ile Gly Val Arg Ile Thr Gln Asn Gly Ser Leu Ile Ser Ser Ser
625 630 635 640

aat cta aca ctt tat agt act act gat atg aat aat act tta cca cta 1968
Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met Asn Asn Thr Leu Pro Leu
645 650 655

aat gta tat gta ata gga gaa aat gga aat tat aca ctt caa gat tta 2016
Asn Val Tyr Val Ile Gly Glu Asn Gly Asn Tyr Thr Leu Gln Asp Leu
660 665 670

tat aat act act aat gtt tta tca aca gga gat att aca tta caa att 2064
Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly Asp Ile Thr Leu Gln Ile
675 680 685

aca gga gga gat caa aaa ata ttt att gat cga ata gaa ttt gtt cct 2112
Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp Arg Ile Glu Phe Val Pro
690 695 700

act atg cct gta cct ggt aat act aac aac aat aac ggt aat aat aac 2160
Thr Met Pro Val Pro Gly Asn Thr Asn Asn Asn Asn Gly Asn Asn Asn
705 710 715 720

ggt aat aat aat ccc cca cac cac gtc tag 2190
Gly Asn Asn Asn Pro Pro His His Val *
725

```

<210> SEQ ID NO: 15

<211> Comprimento: 729

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AXMI-031 (TRUNCADO)

<400> Sequência: 15

```

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

```

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

```

Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly Tyr Phe Pro Asn Asn Ala
      565      570      575
Arg Pro Thr Val Val Lys Glu Trp Ile Asn Gly Ala Ser Ala Val Pro
      580      585      590
Leu Asp Ser Gly Asn Thr Leu Phe Met Thr Ala Thr Asn Leu Thr Ala
      595      600      605
Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala Asn Pro Asn Ser Asn Thr
      610      615      620
Gln Ile Gly Val Arg Ile Thr Gln Asn Gly Ser Leu Ile Ser Ser Ser
      625      630      635      640
Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met Asn Asn Thr Leu Pro Leu
      645      650      655
Asn Val Tyr Val Ile Gly Glu Asn Gly Asn Tyr Thr Leu Gln Asp Leu
      660      665      670
Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly Asp Ile Thr Leu Gln Ile
      675      680      685
Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp Arg Ile Glu Phe Val Pro
      690      695      700
Thr Met Pro Val Pro Gly Asn Thr Asn Asn Asn Gly Asn Asn Asn
      705      710      715      720
Gly Asn Asn Asn Pro Pro His His Val
      725

```

<210> SEQ ID NO: 16

<211> Comprimento: 3558

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: axmi-031(ml)

<221> Nome/Chave: CDS

<222> Localização: (1)...(3558)

<400> Sequência: 16

```

atg gat tgt aat tta caa tca caa caa aat att cca tat aat gta tta      48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu
  1              5              10              15

gca ata cca gta tct aat gtt aat tcg ttg act gat aca gtt gga gat      96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp
      20              25              30

tta aaa aaa gca tgg gaa gaa ttt caa aaa act ggt tct ttt tca tta      144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu
      35              40              45

aca gct tta caa caa gga ttt tct gct tca caa gga gga aca ttc aat      192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn
      50              55              60

tat tta aca tta cta caa tca gga ata tca tta gct ggt tct ttt gtt      240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val
      65              70              75              80

```

cct gga ggt act ttt gta gca cct att att aat atg gtt att ggt tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	
tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata aat	336
Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn	
100 105 110	
tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat tct	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc act	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat tat	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	
aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca aat	576
Thr Asn Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca cct	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	
tat ttt gct ata gga gca aca gca cgt ttt gca gca atg caa tct tat	672
Tyr Phe Ala Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr	
210 215 220	
att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac gca	720
Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala	
225 230 235 240	
cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat atg	768
Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met	
245 250 255	
cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt aaa	816
Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys	
260 265 270	
gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt gat	864
Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp	
275 280 285	
acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat att	912
Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile	
290 295 300	

gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa aca	960
Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr	
305 310 315 320	
gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa gaa	1008
Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu	
325 330 335	
gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct ttt	1056
Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe	
340 345 350	
agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att tct	1104
Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser	
355 360 365	
tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc cct	1152
Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro	
370 375 380	
cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta aat	1200
Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn	
385 390 395 400	
tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa tct	1248
Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser	
405 410 415	
cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat gca	1296
Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala	
420 425 430	
gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga ata	1344
Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile	
435 440 445	
gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca gaa	1392
Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu	
450 455 460	
cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc tgt	1440
Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys	
465 470 475 480	
aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt aca	1488
Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr	
485 490 495	
caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca agt	1536
Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser	
500 505 510	
ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gag tta gat	1584
Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp	
515 520 525	

tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa gga	1632
Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly	
530 535 540	
tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att aat	1680
Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn	
545 550 555 560	
ggg gca agt gct gta cca ctt gat tca gga aat acc tta ttt atg acg	1728
Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr	
565 570 575	
gct acg aat tta aca gct act caa tat aga att aga ata cgt tat gca	1776
Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala	
580 585 590	
aat cca aat tca aat act caa atc ggt gta cga att aca caa aat ggt	1824
Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly	
595 600 605	
tct cta att tcc agt agt aat cta aca ctt tat agt act act gat atg	1872
Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met	
610 615 620	
aat aat act tta cca cta aat gta tat gta ata gga gaa aat gga aat	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tat aca ctt caa gat tta tat aat act act aat gtt tta tca aca gga	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	
gat att aca tta caa att aca gga gga gat caa aaa ata ttt att gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	
cga ata gaa ttt gtt cct act atg cct gta cct ggt aat act aac aac	2064
Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn	
675 680 685	
aat aac ggt aat aat aac ggt aat aat aat ccc cca cac cac gtt tgt	2112
Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val Cys	
690 695 700	
gca ata gct ggt aca caa caa tct tgt tct gga ccg ccc aaa ttt gaa	2160
Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe Glu	
705 710 715 720	
caa gta agt gat tta gaa aaa att aca aca caa gta tat atg tta ttc	2208
Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu Phe	
725 730 735	
aaa tct tct ccg tat gaa gaa tta gct cta gaa gtt tcc agc tat caa	2256
Lys Ser Ser Pro Tyr Glu Glu Leu Ala Leu Glu Val Ser Ser Tyr Gln	
740 745 750	

att agt caa gta gca tta aaa gtt atg gca tta tct gat gaa cta ttt Ile Ser Gln Val Ala Leu Lys Val Met Ala Leu Ser Asp Glu Leu Phe 755 760 765	2304
tgt gaa gaa aaa aac gta tta cga aaa tta gtc aat aaa gca aaa caa Cys Glu Glu Lys Asn Val Leu Arg Lys Leu Val Asn Lys Ala Lys Gln 770 775 780	2352
tta tta gaa gca agt aac tta cta gta ggt gga aat ttt gaa aca act Leu Leu Glu Ala Ser Asn Leu Leu Val Gly Gly Asn Phe Glu Thr Thr 785 790 795 800	2400
caa aat tgg gta ctt gga aca aat gct tat ata aat tat gat tcg ttt Gln Asn Trp Val Leu Gly Thr Asn Ala Tyr Ile Asn Tyr Asp Ser Phe 805 810 815	2448
tta ttt aat gga aat tat tta tct tta caa cca gca agt gga ttt ttc Leu Phe Asn Gly Asn Tyr Leu Ser Leu Gln Pro Ala Ser Gly Phe Phe 820 825 830	2496
aca tct tat gct tat caa aaa ata gat gag tca aca tta aaa cca tat Thr Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Thr Leu Lys Pro Tyr 835 840 845	2544
aca cga tat aaa gtt tct ggg ttc att ggg caa agt aat caa gta gaa Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val Glu 850 855 860	2592
ctt att att tct cgt tat gga aaa gaa att gat aaa ata tta aat gtt Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val 865 870 875 880	2640
cca tat gca gga cct ctt cct atc act gct gat gca tca ata act tgt Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys 885 890 895	2688
tgt gca cca gaa ata ggc caa tgt gat ggg gaa caa tct gat tct cat Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His 900 905 910	2736
ttc ttt aac tat agc atc gat gta ggt gca ctt cac cca gaa tta aac Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn 915 920 925	2784
cct ggc att gaa att ggt ctt aaa att gtg caa tca aat ggt tat ata Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile 930 935 940	2832
aca att agt aat cta gaa att att gaa gaa cgt cca ctt aca gaa atg Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met 945 950 955 960	2880
gaa att caa gca gtc aat cga aaa aat caa aaa tgg gaa aga gaa aaa Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys 965 970 975	2928

ctt cta gaa tgt gca agt att agt gaa ctt tta caa cca att att aat	2976
Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn	
980 985 990	
caa atc gat tca ttg ttt aaa gat gga aac tgg tat aat gat att ctt	3024
Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu	
995 1000 1005	
cct cat gtc aca tat caa gat tta aaa aat att ata ata ccc gag tta	3072
Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu	
1010 1015 1020	
cca aaa tta aaa cat tgg ttc ata gag aat ctc cca ggt gaa tat cat	3120
Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His	
1025 1030 1035 1040	
gaa att gaa caa aaa atg aaa gaa gct cta aaa tat gca ttt aca caa	3168
Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln	
1045 1050 1055	
tta gac gag aaa aat tta atc cac aat ggt cac ttt aca act aac tta	3216
Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu	
1060 1065 1070	
ata gat tgg caa gta gaa ggt gat gct caa atg aaa gta tta gaa aat	3264
Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn	
1075 1080 1085	
gat gct ctt gca tta caa ctt ttc aac tgg gat gct aat gct tca caa	3312
Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Asn Ala Ser Gln	
1090 1095 1100	
tct ata aat ata tta gaa ttt gat gaa gat aag gca tat aaa ctt cgc	3360
Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg	
1105 1110 1115 1120	
gta tat gct caa gga agc gga aca atc caa ttt gga aac tgt gaa gat	3408
Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp	
1125 1130 1135	
gaa gct atc caa ttt aat aca aac tca ttc ata tat caa gaa aaa ata	3456
Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile	
1140 1145 1150	
gtc tat ttc gat acc cca tca gtt aat tta cac ata caa tca gaa ggt	3504
Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly	
1155 1160 1165	
tct gaa ttt att gta agt agt atc gat cta att gaa tta tca gac gac	3552
Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp	
1170 1175 1180	
caa taa	3558
Gln *	
1185	

<210> SEQ ID NO: 17
 <211> Comprimento: 1185
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: AXMI-031(M1)

<400> Sequência: 17

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

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

59/123

```

Thr Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Thr Leu Lys Pro Tyr
      835                      840                      845
Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val Glu
      850                      855                      860
Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val
865                      870                      875                      880
Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys
      885                      890                      895
Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His
      900                      905                      910
Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn
      915                      920                      925
Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile
      930                      935                      940
Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met
945                      950                      955                      960
Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys
      965                      970                      975
Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn
      980                      985                      990
Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu
      995                      1000                      1005
Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu
      1010                      1015                      1020
Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His
1025                      1030                      1035                      1040
Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln
      1045                      1050                      1055
Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu
      1060                      1065                      1070
Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn
      1075                      1080                      1085
Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Asn Ala Ser Gln
      1090                      1095                      1100
Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg
1105                      1110                      1115                      1120
Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp
      1125                      1130                      1135
Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile
      1140                      1145                      1150
Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly
      1155                      1160                      1165
Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp
      1170                      1175                      1180
Gln
1185

```

<210> SEQ ID NO: 18

<211> Comprimento: 2112

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: axmi-031(ml)truncado

<221> Nome/Chave: CDS

<222> Localização: (1)...(2112)

<400> Sequência: 18

atg gat tgt aat tta caa tca caa caa aat att cca tat aat gta tta	48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu	
1 5 10 15	
gca ata cca gta tct aat gtt aat tcg ttg act gat aca gtt gga gat	96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp	
20 25 30	
tta aaa aaa gca tgg gaa gaa ttt caa aaa act ggt tct ttt tca tta	144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu	
35 40 45	
aca gct tta caa caa gga ttt tct gct tca caa gga gga aca ttc aat	192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn	
50 55 60	
tat tta aca tta cta caa tca gga ata tca tta gct ggt tct ttt gtt	240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val	
65 70 75 80	
cct gga ggt act ttt gta gca cct att att aat atg gtt att ggt tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	
tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata aat	336
Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn	
100 105 110	
tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat tct	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc act	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat tat	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	
aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca aat	576
Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca cct	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	

tat ttt gct ata gga gca aca gca cgt ttt gca gca atg caa tct tat Tyr Phe Ala Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr 210 215 220	672
att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac gca Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala 225 230 235 240	720
cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat atg Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met 245 250 255	768
cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt aaa Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys 260 265 270	816
gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt gat Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp 275 280 285	864
acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat att Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile 290 295 300	912
gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa aca Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Tyr Thr Ser Gln Thr 305 310 315 320	960
gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa gaa Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu 325 330 335	1008
gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct ttt Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe 340 345 350	1056
agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att tct Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser 355 360 365	1104
tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc cct Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro 370 375 380	1152
cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta aat Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn 385 390 395 400	1200
tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa tct Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser 405 410 415	1248
cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat gca Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala 420 425 430	1296

gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga ata	1344
Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile	
435 440 445	
gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca gaa	1392
Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu	
450 455 460	
cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc tgt	1440
Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys	
465 470 475 480	
aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt aca	1488
Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr	
485 490 495	
caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca agt	1536
Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser	
500 505 510	
ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gag tta gat	1584
Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp	
515 520 525	
tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa gga	1632
Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly	
530 535 540	
tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att aat	1680
Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn	
545 550 555 560	
ggt gca agt gct gta cca ctt gat tca gga aat acc tta ttt atg acg	1728
Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr	
565 570 575	
gct acg aat tta aca gct act caa tat aga att aga ata cgt tat gca	1776
Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala	
580 585 590	
aat cca aat tca aat act caa atc ggt gta cga att aca caa aat ggt	1824
Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly	
595 600 605	
tct cta att tcc agt agt aat cta aca ctt tat agt act act gat atg	1872
Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met	
610 615 620	
aat aat act tta cca cta aat gta tat gta ata gga gaa aat gga aat	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tat aca ctt caa gat tta tat aat act act aat gtt tta tca aca gga	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	

gat att aca tta caa att aca gga gga gat caa aaa ata ttt att gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	
cga ata gaa ttt gtt cct act atg cct gta cct ggt aat act aac aac	2064
Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn	
675 680 685	
aat aac ggt aat aat aac ggt aat aat aat ccc cca cac cac gtc tag	2112
Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val *	
690 695 700	

<210> SEQ ID NO: 19

<211> Comprimento: 703

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AXMI-031(M1)TRUNCADO

<400> Sequência: 19

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

Asp	Ser	Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	Ser	Val	Asp	275	280	285
Thr	Tyr	Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp	Ile	290	295	300
Val	Ala	Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	Ser	Gln	Thr	305	310	315
Ala	Leu	Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln	Glu	325	330	335
Glu	Gly	Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	Phe	340	345	350
Ser	Tyr	Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	Ser	355	360	365
Tyr	Tyr	Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr	Pro	370	375	380
Pro	Lys	Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu	Asn	385	390	395
Tyr	Ala	Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu	Ser	405	410	415
Leu	Gly	Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn	Ala	420	425	430
Ala	Thr	Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	Asn	Gly	Ile	435	440	445
Gly	Ala	Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Pro	Thr	Glu	450	455	460
Pro	Pro	Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	Ala	Ser	Cys	465	470	475
Asn	Pro	Ser	Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Pro	Phe	Thr	485	490	495
Gln	Ala	Asn	Val	Lys	Gly	Asn	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala	Ser	500	505	510
Leu	Val	Ser	Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu	Asp	515	520	525
Ser	Asp	Thr	Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys	Gly	530	535	540
Tyr	Phe	Pro	Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile	Asn	545	550	555
Gly	Ala	Ser	Ala	Val	Pro	Leu	Asp	Ser	Gly	Asn	Thr	Leu	Phe	Met	Thr	565	570	575
Ala	Thr	Asn	Leu	Thr	Ala	Thr	Gln	Tyr	Arg	Ile	Arg	Ile	Arg	Tyr	Ala	580	585	590
Asn	Pro	Asn	Ser	Asn	Thr	Gln	Ile	Gly	Val	Arg	Ile	Thr	Gln	Asn	Gly	595	600	605
Ser	Leu	Ile	Ser	Ser	Ser	Asn	Leu	Thr	Leu	Tyr	Ser	Thr	Thr	Asp	Met	610	615	620
Asn	Asn	Thr	Leu	Pro	Leu	Asn	Val	Tyr	Val	Ile	Gly	Glu	Asn	Gly	Asn	625	630	635
Tyr	Thr	Leu	Gln	Asp	Leu	Tyr	Asn	Thr	Thr	Asn	Val	Leu	Ser	Thr	Gly	645	650	655
Asp	Ile	Thr	Leu	Gln	Ile	Thr	Gly	Gly	Asp	Gln	Lys	Ile	Phe	Ile	Asp	660	665	670
Arg	Ile	Glu	Phe	Val	Pro	Thr	Met	Pro	Val	Pro	Gly	Asn	Thr	Asn	Asn	675	680	685
Asn	Asn	Gly	Asn	Asn	Asn	Gly	Asn	Asn	Asn	Pro	Pro	His	His	Val		690	695	700

<210> SEQ ID NO: 20
 <211> Comprimento: 2219
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: axmi-031(A-D)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(2216)

<400> Sequência: 20

atg gat tgt aat tta caa tca caa caa aat att cca tat aat gta tta	48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu	
1 5 10 15	
gca ata cca gta tct aat gtt aat tcg ttg act gat aca gtt gga gat	96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp	
20 25 30	
tta aaa aaa gca tgg gaa gaa ttt caa aaa act ggt tct ttt tca tta	144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu	
35 40 45	
aca gct tta caa caa gga ttt tct gct tca caa gga gga aca ttc aat	192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn	
50 55 60	
tat tta aca tta cta caa tca gga ata tca tta gct ggt tct ttt gtt	240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val	
65 70 75 80	
cct gga ggt act ttt gta gca cct att att aat atg gtt att ggt tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	
tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata aat	336
Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile Asn	
100 105 110	
tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat tct	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc act	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat tat	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	

aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca aat	576
Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca cct	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	
tat ttt gtt ata gga gca aca gca cgt ttt gca gca atg caa tct tat	672
Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr	
210 215 220	
att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac gca	720
Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala	
225 230 235 240	
cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat atg	768
Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met	
245 250 255	
cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt aaa	816
Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys	
260 265 270	
gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt gat	864
Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp	
275 280 285	
acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat att	912
Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile	
290 295 300	
gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa aca	960
Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Tyr Thr Ser Gln Thr	
305 310 315 320	
gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa gaa	1008
Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu	
325 330 335	
gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct ttt	1056
Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe	
340 345 350	
agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att tct	1104
Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser	
355 360 365	
tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc cct	1152
Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro	
370 375 380	
cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta aat	1200
Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn	
385 390 395 400	

tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa tct Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser 405 410 415	1248
cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat gca Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala 420 425 430	1296
gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga ata Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile 435 440 445	1344
gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca gaa Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu 450 455 460	1392
cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc tgt Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys 465 470 475 480	1440
aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt aca Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr 485 490 495	1488
caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca agt Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser 500 505 510	1536
ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gaa tta gat Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp 515 520 525	1584
tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa gga Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly 530 535 540	1632
tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att aat Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn 545 550 555 560	1680
ggt gca agt gct gta cca ctt gat tca gga aat acc tta ttt atg acg Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr 565 570 575	1728
gct acg aat tta aca gct act caa tat aga att aga ata cgt tat gca Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala 580 585 590	1776
aat cca aat tca aat act caa atc ggt gta cga att aca caa aat ggt Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly 595 600 605	1824
tct cta att tcc agt agt aat cta aca ctt tat agt act act gat atg Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met 610 615 620	1872

68/123

aat aat act tta cca cta aat gta tat gta ata gga gaa aat gga aat	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tat aca ctt caa gat tta tat aat act act aat gtt tta tca aca gga	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	
gat att aca tta caa att aca gga gga gat caa aaa ata ttt att gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	
cga ata gaa ttt gtt cct act atg cct gta cct ggt aat act aac aac	2064
Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn	
675 680 685	
aat aac ggt aat aat aac ggt aat aat aat ccc cca cac cac gtt tgt	2112
Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val Cys	
690 695 700	
gca ata gct ggt aca caa caa tct tgt tct gga ccg ccc aaa ttt gaa	2160
Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe Glu	
705 710 715 720	
caa gta agt gat tta gaa aaa att aca aca caa gta tat atg tta ttc	2208
Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu Phe	
725 730 735	
aaa tct ta tag	2219
Lys Ser	

<210> SEQ ID NO: 21

<211> Comprimento: 738

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AXMI-031(A-D)

<400> Sequência: 21

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

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

[illegible]

```
<210> SEQ ID NO: 22
<211> Comprimento: 2015
<212> Tipo: DNA
<213> Organismo: Sequência Artificial
```

<220> Características:
<223> Outras Informações: axmi-031(B-C)

```
<221> Nome/Chave: CDS
<222> Localização: (1)...(2012)
```

<400> Sequência: 22																	
atg	gca	ata	cca	gta	tct	aat	gtt	aat	tcg	ttg	act	gat	aca	gtt	gga	48	
Met	Ala	Ile	Pro	Val	Ser	Asn	Val	Asn	Ser	Leu	Thr	Asp	Thr	Val	Gly		
1				5					10					15			
gat	tta	aaa	aaa	gca	tggt	gaa	gaa	ttt	caa	aaa	act	ggt	tct	ttt	tca	96	
Asp	Leu	Lys	Lys	Ala	Trp	Glu	Glu	Phe	Gln	Lys	Thr	Gly	Ser	Phe	Ser		
			20					25					30				
tta	aca	gct	tta	caa	caa	gga	ttt	tct	gct	tca	caa	gga	gga	aca	ttc	144	
Leu	Thr	Ala	Leu	Gln	Gln	Gly	Phe	Ser	Ala	Ser	Gln	Gly	Gly	Thr	Phe		
		35				40						45					
aat	tat	tta	aca	tta	cta	caa	tca	gga	ata	tca	tta	gct	ggt	tct	ttt	192	
Asn	Tyr	Leu	Thr	Leu	Leu	Gln	Ser	Gly	Ile	Ser	Leu	Ala	Gly	Ser	Phe		
	50					55					60						
gtt	cct	gga	ggt	act	ttt	gta	gca	cct	att	att	aat	atg	gtt	att	ggt	240	
Val	Pro	Gly	Gly	Thr	Phe	Val	Ala	Pro	Ile	Ile	Asn	Met	Val	Ile	Gly		
65				70					75					80			

tgg tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata	288
Trp Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile	
85 90 95	
aat tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta	336
Asn Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu	
100 105 110	
gat gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat	384
Asp Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp	
115 120 125	
tct tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc	432
Ser Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly	
130 135 140	
act gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat	480
Thr Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp	
145 150 155 160	
tat aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca	528
Tyr Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala	
165 170 175	
aat aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca	576
Asn Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala	
180 185 190	
cct tat ttt gtt ata gga gca aca gca cgt ttt gca gca atg caa tct	624
Pro Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser	
195 200 205	
tat att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac	672
Tyr Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp	
210 215 220	
gca cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat	720
Ala Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn	
225 230 235 240	
atg cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt	768
Met Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe	
245 250 255	
aaa gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt	816
Lys Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val	
260 265 270	
gat acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat	864
Asp Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp	
275 280 285	
att gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa	912
Ile Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Tyr Thr Ser Gln	
290 295 300	

aca gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa	960
Thr Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln	
305 310 315 320	
gaa gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct	1008
Glu Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser	
325 330 335	
ttt agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att	1056
Phe Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile	
340 345 350	
tct tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc	1104
Ser Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr	
355 360 365	
cct cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta	1152
Pro Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu	
370 375 380	
aat tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa	1200
Asn Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu	
385 390 395 400	
tct cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat	1248
Ser Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn	
405 410 415	
gca gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga	1296
Ala Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly	
420 425 430	
ata gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca	1344
Ile Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr	
435 440 445	
gaa cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc	1392
Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser	
450 455 460	
tgt aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt	1440
Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe	
465 470 475 480	
aca caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca	1488
Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala	
485 490 495	
agt ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gaa tta	1536
Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu	
500 505 510	
gat tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa	1584
Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys	
515 520 525	

73/123

gga tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att	1632
Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile	
530 535 540	
aat ggt gca agt gct gta cca ctt gat tca gga aat acc tta ttt atg	1680
Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met	
545 550 555 560	
acg gct acg aat tta aca gct act caa tat aga att aga ata cgt tat	1728
Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr	
565 570 575	
gca aat cca aat tca aat act caa atc ggt gta cga att aca caa aat	1776
Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn	
580 585 590	
ggt tct cta att tcc agt agt aat cta aca ctt tat agt act act gat	1824
Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp	
595 600 605	
atg aat aat act tta cca cta aat gta tat gta ata gga gaa aat gga	1872
Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly	
610 615 620	
aat tat aca ctt caa gat tta tat aat act act aat gtt tta tca aca	1920
Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr	
625 630 635 640	
gga gat att aca tta caa att aca gga gga gat caa aaa ata ttt att	1968
Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile	
645 650 655	
gat cga ata gaa ttt gtt cct act atg cct gta cct ggt aat ta	2012
Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn	
660 665 670	
tag	2015

<210> SEQ ID NO: 23

<211> Comprimento: 670

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AXMI-031(B-C)

<400> Sequência: 23

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

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

75/123

```

Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile
530                    535                    540
Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met
545                    550                    555                    560
Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr
565                    570                    575
Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn
580                    585                    590
Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp
595                    600                    605
Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly
610                    615                    620
Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr
625                    630                    635                    640
Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile
645                    650                    655
Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn
660                    665                    670

```

<210> SEQ ID NO: 24
 <211> Comprimento: 2174
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: axmi-031(B-D)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(2171)

<400> Sequência: 24

```

atg gca ata cca gta tct aat gtt aat tcg ttg act gat aca gtt gga      48
Met Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly
1                    5                    10                    15

gat tta aaa aaa gca tgg gaa gaa ttt caa aaa act ggt tct ttt tca      96
Asp Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser
20                    25                    30

tta aca gct tta caa caa gga ttt tct gct tca caa gga gga aca ttc      144
Leu Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe
35                    40                    45

aat tat tta aca tta cta caa tca gga ata tca tta gct ggt tct ttt      192
Asn Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe
50                    55                    60

gtt cct gga ggt act ttt gta gca cct att att aat atg gtt att ggt      240
Val Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly
65                    70                    75                    80

tgg tta tgg cca cat aaa aac aaa aat gcg gat aca gaa aat tta ata      288
Trp Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn Leu Ile
85                    90                    95

```

aat tta att gat tca gaa att caa aaa caa tta aac aaa gct tta tta	336
Asn Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu	
100 105 110	
gat gca gat aga aat gag tgg agc tct tat tta gaa tct ata ttt gat	384
Asp Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp	
115 120 125	
tct tca aat aac cta aat ggt gca att gta gat gca cag tgg tca ggc	432
Ser Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly	
130 135 140	
act gta aat act aca aat aga aca cta aga aat cca aca gaa tca gat	480
Thr Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp	
145 150 155 160	
tat aca aat gtt gtt aca aat ttt att gca gcg gat ggt gac att gca	528
Tyr Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala	
165 170 175	
aat aat gaa aat cac ata atg aat ggc aac ttt gac gta gct gca gca	576
Asn Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala	
180 185 190	
cct tat ttt gtt ata gga gca aca gca cgt ttt gca gca atg caa tct	624
Pro Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser	
195 200 205	
tat att aaa ttt tgt aat gct tgg att gat aaa gtt gga ttg agt gac	672
Tyr Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp	
210 215 220	
gca cag ctt act aca caa aag gct aat tta gat cgc acg aaa caa aat	720
Ala Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn	
225 230 235 240	
atg cgt aat gca att ctt aac tat aca caa caa gtt atg aaa gtt ttt	768
Met Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe	
245 250 255	
aaa gat tcc aaa aat atg cct aca ata ggt act aat aaa ttt agt gtt	816
Lys Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val	
260 265 270	
gat acc tat aat gta tat att aaa gga atg aca tta aat gtt tta gat	864
Asp Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp	
275 280 285	
att gta gca ata tgg cct tca tta tat cca gat gat tat act tca caa	912
Ile Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln	
290 295 300	
aca gcc tta gaa caa aca cgt gtc act ttt tca aat atg gtt ggc caa	960
Thr Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln	
305 310 315 320	

gaa gaa ggt aca gat gga agc cta aga att tac aat act ttt gat tct Glu Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser 325 330 335	1008
ttt agt tat caa cat agt cca ata cct aat aat aat gtt aat tta att Phe Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile 340 345 350	1056
tct tat tat aat gat gaa tta caa aat cta gaa tta gga gta tat acc Ser Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr 355 360 365	1104
cct cct aaa aaa gga agt gga tac tct tat cct tat gga ttt gtt tta Pro Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu 370 375 380	1152
aat tat gca aac agt aaa tat aaa tat ggt gat agc aat gat cca gaa Asn Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu 385 390 395 400	1200
tct cta gga gga tta tct aca cta tct gca cct ata caa caa gtt aat Ser Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn 405 410 415	1248
gca gca act caa aac agt aaa tat cta gat gga gaa atc cta aat gga Ala Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly 420 425 430	1296
ata gga gca tcc tta cct ggt tat tgt act aca gga tgt tca cca aca Ile Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr 435 440 445	1344
gaa cca cct ttt agt tgt act tct acc gct aat ggc tat aaa gca agc Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser 450 455 460	1392
tgt aat cct tca gat aca aat caa aaa att aac gct tta tat cct ttt Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe 465 470 475 480	1440
aca caa gct aat gta aag gga aac aca gga aaa tta gga gta ctg gca Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala 485 490 495	1488
agt ctt gtt tca tat gat tta aat cct aaa aat gta ttt ggt gaa tta Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu 500 505 510	1536
gat tca gat aca aat aat gtt atc tta aaa gga att cct gca gaa aaa Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys 515 520 525	1584
gga tat ttt cct aat aat gcg cgt cct act gtt gta aaa gaa tgg att Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Lys Glu Trp Ile 530 535 540	1632

78/123

[illegible]

<210> SEQ ID NO: 25

<211> Comprimeto: 723

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AXMI-031 (B-D)

<400> Sequência: 25

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

Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser
 450 455 460
 Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe
 465 470 475 480
 Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala
 485 490 495
 Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu
 500 505 510
 Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys
 515 520 525
 Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile
 530 535 540
 Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met
 545 550 555 560
 Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr
 565 570 575
 Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn
 580 585 590
 Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp
 595 600 605
 Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly
 610 615 620
 Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr
 625 630 635 640
 Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile
 645 650 655
 Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn
 660 665 670
 Asn Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val
 675 680 685
 Cys Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe
 690 695 700
 Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu
 705 710 715 720
 Phe Lys Ser

<210> SEQ ID NO: 26

<211> Comprimento: 2226

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: synaxmi-031(A-D)

<221> Nome/Chave: CDS

<222> Localização: (1)...(2226)

<400> Sequência: 26

atg cag atc ctt gat tgc aat ctc cag agc cag cag aat atc cca tac 48
 Met Gln Ile Leu Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr
 1 5 10 15

aat gtg ctt gct atc cct gtt agc aat gtt aat agc ctt acg gat acg 96
 Asn Val Leu Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr
 20 25 30

gtt ggc gat cta aag aag gca tgg gaa gag ttc cag aag acg ggc agc	144
Val Gly Asp Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser	
35 40 45	
ttc tcc ctt acc gca ctc cag cag ggc ttc agc gct tcc cag ggc ggg	192
Phe Ser Leu Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly	
50 55 60	
acg ttc aat tac ctc acc ctc ctc cag tcc ggg atc tcc ctg gct ggc	240
Thr Phe Asn Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly	
65 70 75 80	
agc ttc gtg cct ggc ggc acc ttc gtt gct cca atc atc aat atg gtg	288
Ser Phe Val Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val	
85 90 95	
atc ggc tgg ctc tgg cct cat aag aat aag aat gca gat acg gaa aat	336
Ile Gly Trp Leu Trp Pro His Lys Asn Lys Asn Ala Asp Thr Glu Asn	
100 105 110	
ctg atc aat ctg atc gat tca gaa atc cag aag cag ctg aat aag gct	384
Leu Ile Asn Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala	
115 120 125	
ctg ctg gat gct gat cgc aat gaa tgg agc agc tac ctc gag tcc atc	432
Leu Leu Asp Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile	
130 135 140	
ttc gat agc agc aat aat ctg aat ggc gca atc gtt gat gct cag tgg	480
Phe Asp Ser Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp	
145 150 155 160	
tcc ggc acc gtg aat acg acc aat cgt acg ctt cgt aat cct acg gaa	528
Ser Gly Thr Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu	
165 170 175	
tca gat tac acc aat gtg gtg acg aat ttc atc gca gca gat ggg gat	576
Ser Asp Tyr Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp	
180 185 190	
atc gct aat aat gaa aat cat atc atg aat ggg aat ttc gat gtg gca	624
Ile Ala Asn Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala	
195 200 205	
gca gct cca tac ttc gtt atc ggg gct acc gct aga ttc gca gct atg	672
Ala Ala Pro Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met	
210 215 220	
cag tcc tac atc aag ttc tgc aat gca tgg atc gat aag gtt ggg ctg	720
Gln Ser Tyr Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu	
225 230 235 240	
tca gat gct cag ctg acc acg cag aag gct aat ctg gat aga acg aag	768
Ser Asp Ala Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys	
245 250 255	

cag	aat	atg	cgt	aat	gct	atc	ctt	aat	tac	acc	cag	cag	gtt	atg	aag	816
Gln	Asn	Met	Arg	Asn	Ala	Ile	Leu	Asn	Tyr	Thr	Gln	Gln	Val	Met	Lys	
			260					265					270			
gtt	ttc	aag	gat	tcc	aag	aat	atg	cca	acg	atc	ggc	acg	aat	aag	ttc	864
Val	Phe	Lys	Asp	Ser	Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	
		275					280					285				
agc	gtt	gat	acc	tac	aat	gtt	tac	atc	aag	ggg	atg	acg	ctt	aat	gtg	912
Ser	Val	Asp	Thr	Tyr	Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	
	290					295					300					
ctt	gat	atc	gtt	gct	atc	tgg	cca	agc	ctg	tac	cca	gat	gat	tac	acg	960
Leu	Asp	Ile	Val	Ala	Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	
305					310					315					320	
tca	cag	acg	gct	ctg	gaa	cag	acc	cgc	gtg	acg	ttc	tcc	aat	atg	gtg	1008
Ser	Gln	Thr	Ala	Leu	Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	
			325						330					335		
ggg	cag	gaa	gaa	ggc	acg	gat	ggc	agc	ctc	aga	atc	tac	aat	acc	ttc	1056
Gly	Gln	Glu	Glu	Gly	Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	
		340						345					350			
gat	agc	ttc	tcc	tac	cag	cat	agc	cct	atc	cct	aat	aat	aat	gtg	aat	1104
Asp	Ser	Phe	Ser	Tyr	Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	
		355					360					365				
ctc	atc	agc	tac	tac	aat	gat	gaa	ctt	cag	aat	ctg	gaa	ctc	ggg	gtt	1152
Leu	Ile	Ser	Tyr	Tyr	Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	
	370					375					380					
tac	acc	cca	cca	aag	aag	ggc	tca	ggc	tac	agc	tac	cca	tac	ggg	ttc	1200
Tyr	Thr	Pro	Pro	Lys	Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	
385					390					395					400	
gtg	ctg	aat	tac	gca	aat	agc	aag	tac	aag	tac	ggc	gat	tcc	aat	gat	1248
Val	Leu	Asn	Tyr	Ala	Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	
				405				410						415		
cca	gaa	tcc	ctc	ggc	ggg	ctt	tcc	acc	ctt	agc	gct	cca	atc	caa	cag	1296
Pro	Glu	Ser	Leu	Gly	Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	
			420					425					430			
gtt	aat	gct	gct	acc	cag	aat	agc	aag	tac	ctt	gat	ggc	gaa	atc	ctg	1344
Val	Asn	Ala	Ala	Thr	Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	
		435					440					445				
aat	ggg	atc	ggc	gca	agc	ctg	cct	ggc	tac	tgc	acg	acc	ggg	tgc	tca	1392
Asn	Gly	Ile	Gly	Ala	Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	
	450					455					460					
cct	acc	gaa	cca	cca	ttc	agc	tgc	acg	agc	acg	gca	aat	ggg	tac	aag	1440
Pro	Thr	Glu	Pro	Pro	Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	
465					470					475					480	

gca agc tgc aat cca agc gat acc aat cag aag atc aat gct ctc tac Ala Ser Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr 485 490 495	1488
cca ttc acg cag gca aat gtg aag ggg aat acc ggg aag ctg ggc gtt Pro Phe Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val 500 505 510	1536
ctc gca agc ctt gtt agc tac gat ctg aat cct aag aat gtt ttc ggg Leu Ala Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly 515 520 525	1584
gaa ctc gat tca gat acc aat aat gtt atc ctt aag ggc atc cca gca Glu Leu Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala 530 535 540	1632
gaa aag ggc tac ttc cca aat aat gca aga cca acc gtt gtg aag gaa Glu Lys Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu 545 550 555 560	1680
tgg atc aat ggg gct tcc gca gtt cca ctt gat tcc ggc aat acc ctg Trp Ile Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu 565 570 575	1728
ttc atg acg gct acg aac ctg acg gct acc cag tac aga atc cgc atc Phe Met Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile 580 585 590	1776
cgt tac gct aac cca aac tca aac acg cag atc ggc gtt aga atc acg Arg Tyr Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr 595 600 605	1824
cag aac ggg agc ctc atc tcc tcc tcc aac ctc acg ctg tac tca acc Gln Asn Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr 610 615 620	1872
acc gat atg aac aac acc ctg cca ctg aac gtt tac gtg atc ggg gaa Thr Asp Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu 625 630 635 640	1920
aac ggg aac tac acc ctc cag gat ctc tac aac acg acg aac gtt ctg Asn Gly Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu 645 650 655	1968
tcc acg ggc gat atc acc ctg cag atc acc ggc ggg gat cag aag ata Ser Thr Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile 660 665 670	2016
ttc atc gat cgc atc gag ttc gtg cct acg atg cca gtg cca ggc aac Phe Ile Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn 675 680 685	2064
acc aac aac aac aac ggc aac aac aac ggc aac aac aac ccc ccc cat Thr Asn Asn Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His 690 695 700	2112

84/123

cat gtc tgc gct ata gct ggt act cag cag tct tgc tca ggt cct cct 2160
His Val Cys Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro
705 710 715 720

aag ttc gag cag gtt tct gat ctc gag aag atc acc acc cag gtc tac 2208
Lys Phe Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr
725 730 735

atg ctg ttc aag tcc taa 2226
Met Leu Phe Lys Ser *
740

<210> SEQ ID NO: 27

<211> Comprimento: 741

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: SYNAXMI-031(A-D)

<400> Sequência: 27

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

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

Lys Phe Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr
725 730 735
Met Leu Phe Lys Ser
740

```
<210> SEQ ID NO: 28
<211> Comprimento: 2316
<212> Tipo: DNA
<213> Organismo: Sequência Artificial
```

<220> Características:
<223> Outras Informações: aposynaxmi-031(A-D)

```
<221> Nome/Chave: CDS
<222> Localização: (1)...(2316)
```

<400> Sequência: 28																
atg	cag	atc	ctt	ggg	tac	tcc	agc	ttc	ggt	gct	atc	gct	ctt	ctt	atg	48
Met	Gln	Ile	Leu	Gly	Tyr	Ser	Ser	Phe	Val	Ala	Ile	Ala	Leu	Leu	Met	
				5					10			15				
agc	gtg	ggt	ggt	ggt	tgc	aat	ggg	ggc	aag	acg	tcc	acc	tac	gtg	cgt	96
Ser	Val	Val	Val	Val	Cys	Asn	Gly	Gly	Lys	Thr	Ser	Thr	Tyr	Val	Arg	
				20					25			30				
aat	ctg	gat	tgc	aat	ctc	cag	agc	cag	cag	aat	atc	cca	tac	aat	gtg	144
Asn	Leu	Asp	Cys	Asn	Leu	Gln	Ser	Gln	Gln	Asn	Ile	Pro	Tyr	Asn	Val	
				35					40			45				
ctt	gct	atc	cct	ggt	agc	aat	ggt	aat	agc	ctt	acg	gat	acg	ggt	ggc	192
Leu	Ala	Ile	Pro	Val	Ser	Asn	Val	Asn	Ser	Leu	Thr	Asp	Thr	Val	Gly	
				50					55			60				
gat	cta	aag	aag	gca	tgg	gaa	gag	ttc	cag	aag	acg	ggc	agc	ttc	tcc	240
Asp	Leu	Lys	Lys	Ala	Trp	Glu	Glu	Phe	Gln	Lys	Thr	Gly	Ser	Phe	Ser	
65					70				75			80				
ctt	acc	gca	ctc	cag	cag	ggc	ttc	agc	gct	tcc	cag	ggc	ggg	acg	ttc	288
Leu	Thr	Ala	Leu	Gln	Gln	Gly	Phe	Ser	Ala	Ser	Gln	Gly	Gly	Thr	Phe	
				85				90			95					
aat	tac	ctc	acc	ctc	ctc	cag	tcc	ggg	atc	tcc	ctg	gct	ggc	agc	ttc	336
Asn	Tyr	Leu	Thr	Leu	Leu	Gln	Ser	Gly	Ile	Ser	Leu	Ala	Gly	Ser	Phe	
				100				105			110					
gtg	cct	ggc	ggc	acc	ttc	ggt	gct	cca	atc	atc	aat	atg	gtg	atc	ggc	384
Val	Pro	Gly	Gly	Thr	Phe	Val	Ala	Pro	Ile	Ile	Asn	Met	Val	Ile	Gly	
				115				120			125					
tgg	ctc	tgg	cct	cat	aag	aat	aag	aat	gca	gat	acg	gaa	aat	ctg	atc	432
Trp	Leu	Trp	Pro	His	Lys	Asn	Lys	Asn	Ala	Asp	Thr	Glu	Asn	Leu	Ile	
				130				135			140					
aat	ctg	atc	gat	tca	gaa	atc	cag	aag	cag	ctg	aat	aag	gct	ctg	ctg	480
Asn	Leu	Ile	Asp	Ser	Glu	Ile	Gln	Lys	Gln	Leu	Asn	Lys	Ala	Leu	Leu	
145					150				155			160				

gat gct gat cgc aat gaa tgg agc agc tac ctc gag tcc atc ttc gat	528
Asp Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp	
165 170 175	
agc agc aat aat ctg aat ggc gca atc gtt gat gct cag tgg tcc ggc	576
Ser Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly	
180 185 190	
acc gtg aat acg acc aat cgt acg ctt cgt aat cct acg gaa tca gat	624
Thr Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp	
195 200 205	
tac acc aat gtg gtg acg aat ttc atc gca gca gat ggg gat atc gct	672
Tyr Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala	
210 215 220	
aat aat gaa aat cat atc atg aat ggg aat ttc gat gtg gca gca gct	720
Asn Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala	
225 230 235 240	
cca tac ttc gtt atc ggg gct acc gct aga ttc gca gct atg cag tcc	768
Pro Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser	
245 250 255	
tac atc aag ttc tgc aat gca tgg atc gat aag gtt ggg ctg tca gat	816
Tyr Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp	
260 265 270	
gct cag ctg acc acg cag aag gct aat ctg gat aga acg aag cag aat	864
Ala Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn	
275 280 285	
atg cgt aat gct atc ctt aat tac acc cag cag gtt atg aag gtt ttc	912
Met Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe	
290 295 300	
aag gat tcc aag aat atg cca acg atc ggc acg aat aag ttc agc gtt	960
Lys Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val	
305 310 315 320	
gat acc tac aat gtt tac atc aag ggg atg acg ctt aat gtg ctt gat	1008
Asp Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp	
325 330 335	
atc gtt gct atc tgg cca agc ctg tac cca gat gat tac acg tca cag	1056
Ile Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln	
340 345 350	
acg gct ctg gaa cag acc cgc gtg acg ttc tcc aat atg gtg ggg cag	1104
Thr Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln	
355 360 365	
gaa gaa ggc acg gat ggc agc ctc aga atc tac aat acc ttc gat agc	1152
Glu Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser	
370 375 380	

ttc tcc tac cag cat agc cct atc cct aat aat aat gtg aat ctc atc	1200
Phe Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile	
385 390 395 400	
agc tac tac aat gat gaa ctt cag aat ctg gaa ctc ggg gtt tac acc	1248
Ser Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr	
405 410 415	
cca cca aag aag ggc tca ggc tac agc tac cca tac ggg ttc gtg ctg	1296
Pro Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu	
420 425 430	
aat tac gca aat agc aag tac aag tac ggc gat tcc aat gat cca gaa	1344
Asn Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu	
435 440 445	
tcc ctc ggc ggg ctt tcc acc ctt agc gct cca atc caa cag gtt aat	1392
Ser Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn	
450 455 460	
gct gct acc cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg	1440
Ala Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly	
465 470 475 480	
atc ggc gca agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc	1488
Ile Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr	
485 490 495	
gaa cca cca ttc agc tgc acg agc acg gca aat ggg tac aag gca agc	1536
Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser	
500 505 510	
tgc aat cca agc gat acc aat cag aag atc aat gct ctc tac cca ttc	1584
Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe	
515 520 525	
acg cag gca aat gtg aag ggg aat acc ggg aag ctg ggc gtt ctc gca	1632
Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala	
530 535 540	
agc ctt gtt agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc	1680
Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu	
545 550 555 560	
gat tca gat acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag	1728
Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys	
565 570 575	
ggc tac ttc cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc	1776
Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile	
580 585 590	
aat ggg gct tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg	1824
Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met	
595 600 605	

89/123

acg gct acg aac ctg acg gct acc cag tac aga atc cgc atc cgt tac	1872
Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr	
610 615 620	
gct aac cca aac tca aac acg cag atc ggc gtt aga atc acg cag aac	1920
Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn	
625 630 635 640	
ggg agc ctc atc tcc tcc tcc aac ctc acg ctg tac tca acc acc gat	1968
Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp	
645 650 655	
atg aac aac acc ctg cca ctg aac gtt tac gtg atc ggg gaa aac ggg	2016
Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly	
660 665 670	
aac tac acc ctc cag gat ctc tac aac acg acg aac gtt ctg tcc acg	2064
Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr	
675 680 685	
ggc gat atc acc ctg cag atc acc ggc ggg gat cag aag ata ttc atc	2112
Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile	
690 695 700	
gat cgc atc gag ttc gtg cct acg atg cca gtg cca ggc aac acc aac	2160
Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn	
705 710 715 720	
aac aac aac ggc aac aac aac ggc aac aac aac ccc ccc cat cat gtc	2208
Asn Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val	
725 730 735	
tgc gct ata gct ggt act cag cag tct tgc tca ggt cct cct aag ttc	2256
Cys Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe	
740 745 750	
gag cag gtt tct gat ctc gag aag atc acc acc cag gtc tac atg ctg	2304
Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu	
755 760 765	
ttc aag tcc taa	2316
Phe Lys Ser *	
770	

<210> SEQ ID NO: 29

<211> Comprimento: 771

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: APOSYNAXMI-031(A-D)

<400> Sequência: 29

Met	Gln	Ile	Leu	Gly	Tyr	Ser	Ser	Phe	Val	Ala	Ile	Ala	Leu	Leu	Met
1				5					10					15	

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

Ala Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly
 465 470 475 480
 Ile Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr
 485 490 495
 Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser
 500 505 510
 Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe
 515 520 525
 Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala
 530 535 540
 Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu
 545 550 555 560
 Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys
 565 570 575
 Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile
 580 585 590
 Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met
 595 600 605
 Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr
 610 615 620
 Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn
 625 630 635 640
 Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp
 645 650 655
 Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly
 660 665 670
 Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr
 675 680 685
 Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile
 690 695 700
 Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn
 705 710 715 720
 Asn Asn Asn Gly Asn Asn Asn Gly Asn Asn Pro Pro His His Val
 725 730 735
 Cys Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe
 740 745 750
 Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu
 755 760 765
 Phe Lys Ser
 770

<210> SEQ ID NO: 30

<211> Comprimento: 2307

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: aposyn2axmi-031(A-D)

<221> Nome/Chave: CDS

<222> Localização: (1)...(2307)

<400> Sequência: 30

atg ggc tac agc agc ttc gtc gcc atc gcg ctg ctg atg agc gtg gtg
 Met Gly Tyr Ser Ser Phe Val Ala Ile Ala Leu Leu Met Ser Val Val
 1 5 10 15

gtg	gtg	tgc	aac	ggc	ggc	aag	aca	agc	acc	tat	gtg	agg	aac	ctg	gac	96
Val	Val	Cys	Asn	Gly	Gly	Lys	Thr	Ser	Thr	Tyr	Val	Arg	Asn	Leu	Asp	
			20					25					30			
tgc	aac	ctc	cag	agc	cag	cag	aac	atc	ccc	tac	aat	gtg	ctg	gcc	atc	144
Cys	Asn	Leu	Gln	Ser	Gln	Gln	Asn	Ile	Pro	Tyr	Asn	Val	Leu	Ala	Ile	
		35					40					45				
ccc	gtc	tca	aat	gtc	aac	agc	ttg	aca	gat	act	gtt	ggg	gat	ttg	aag	192
Pro	Val	Ser	Asn	Val	Asn	Ser	Leu	Thr	Asp	Thr	Val	Gly	Asp	Leu	Lys	
	50					55					60					
aag	gca	tgg	gag	gag	ttc	cag	aag	acc	ggc	agc	ttc	agc	ttg	acg	gcg	240
Lys	Ala	Trp	Glu	Glu	Phe	Gln	Lys	Thr	Gly	Ser	Phe	Ser	Leu	Thr	Ala	
65					70				75						80	
ctg	caa	caa	ggc	ttc	tca	gca	agc	caa	gga	ggc	acc	ttc	aac	tac	ctc	288
Leu	Gln	Gln	Gly	Phe	Ser	Ala	Ser	Gln	Gly	Gly	Thr	Phe	Asn	Tyr	Leu	
			85					90						95		
acc	ttg	ctg	caa	agc	ggc	atc	agc	ttg	gcc	ggc	agc	ttc	gtg	ccc	ggc	336
Thr	Leu	Leu	Gln	Ser	Gly	Ile	Ser	Leu	Ala	Gly	Ser	Phe	Val	Pro	Gly	
			100					105					110			
ggc	acc	ttc	gtg	gcg	ccc	atc	atc	aac	atg	gtg	att	gga	tgg	ctg	tgg	384
Gly	Thr	Phe	Val	Ala	Pro	Ile	Ile	Asn	Met	Val	Ile	Gly	Trp	Leu	Trp	
		115				120						125				
ccg	cac	aag	aac	aag	aat	gct	gac	aca	gaa	aat	ttg	atc	aac	ctc	att	432
Pro	His	Lys	Asn	Lys	Asn	Ala	Asp	Thr	Glu	Asn	Leu	Ile	Asn	Leu	Ile	
	130					135					140					
gat	tca	gag	atc	cag	aag	cag	ctc	aac	aag	gcg	ctg	ctg	gat	gct	gac	480
Asp	Ser	Glu	Ile	Gln	Lys	Gln	Leu	Asn	Lys	Ala	Leu	Leu	Asp	Ala	Asp	
145					150					155					160	
aga	aat	gaa	tgg	agc	agc	tac	ctg	gag	agc	atc	ttt	gat	tca	agc	aac	528
Arg	Asn	Glu	Trp	Ser	Ser	Tyr	Leu	Glu	Ser	Ile	Phe	Asp	Ser	Ser	Asn	
				165				170						175		
aac	ctc	aac	ggc	gcc	atc	gtg	gat	gct	caa	tgg	tca	ggc	acc	gtc	aac	576
Asn	Leu	Asn	Gly	Ala	Ile	Val	Asp	Ala	Gln	Trp	Ser	Gly	Thr	Val	Asn	
			180					185					190			
acc	acc	aac	agg	acg	ctg	agg	aat	cca	aca	gaa	agt	gac	tac	acc	aat	624
Thr	Thr	Asn	Arg	Thr	Leu	Arg	Asn	Pro	Thr	Glu	Ser	Asp	Tyr	Thr	Asn	
		195					200					205				
gtg	gtg	acc	aac	ttc	att	gct	gct	gat	gga	gac	atc	gcc	aac	aat	gag	672
Val	Val	Thr	Asn	Phe	Ile	Ala	Ala	Asp	Gly	Asp	Ile	Ala	Asn	Asn	Glu	
		210				215					220					
aac	cac	atc	atg	aat	gga	aat	ttt	gat	gtt	gct	gct	gct	cca	tat	ttt	720
Asn	His	Ile	Met	Asn	Gly	Asn	Phe	Asp	Val	Ala	Ala	Ala	Pro	Tyr	Phe	
225					230					235					240	

gtg att gga gca aca gca aga ttt gct gcc atg caa tca tac atc aag	768
Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr Ile Lys	
245 250 255	
ttc tgc aat gca tgg atc gac aag gtg ggc ctc tct gat gct cag ctc	816
Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala Gln Leu	
260 265 270	
acc acc cag aag gcc aac ctg gac agg acc aag cag aac atg agg aat	864
Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met Arg Asn	
275 280 285	
gcc atc ttg aat tac acc cag cag gtg atg aag gtg ttc aag gac agc	912
Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys Asp Ser	
290 295 300	
aag aac atg cca acc atc ggc acc aac aag ttc tca gtg gac acc tac	960
Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp Thr Tyr	
305 310 315 320	
aat gtc tac atc aag ggc atg acg ctc aat gtg ctg gac atc gtc gcc	1008
Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile Val Ala	
325 330 335	
atc tgg cca agc ctc tac cct gat gac tac acc tca cag acg gcg ctg	1056
Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr Ala Leu	
340 345 350	
gag caa aca agg gtg acc ttc agc aac atg gtg ggc caa gaa gaa gga	1104
Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu Glu Gly	
355 360 365	
act gat ggc agc ttg agg atc tac aac acc ttc gac agc ttc agc tac	1152
Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe Ser Tyr	
370 375 380	
cag cat tct ccc atc ccc aac aac aat gtc aac ctc atc agc tac tac	1200
Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser Tyr Tyr	
385 390 395 400	
aat gat gag ctg caa aat ttg gag cta gga gtc tac acg ccg cca aag	1248
Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro Pro Lys	
405 410 415	
aaa gga agt gga tat tct tat cct tat ggc ttc gtg ctc aac tac gcc	1296
Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn Tyr Ala	
420 425 430	
aac agc aag tac aag tat gga gat tca aat gat cca gaa agc ctc ggc	1344
Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser Leu Gly	
435 440 445	
ggc ctc tcc acc ttg tcg gcg ccc atc cag cag gtg aac gcc gcc acc	1392
Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala Ala Thr	
450 455 460	

cag aac agc aag tac ctt gat gga gag atc ctc aat ggc att gga gct	1440
Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile Gly Ala	
465 470 475 480	
tct ctt cct ggc tac tgc acc acc ggc tgc tcg ccg acg gag ccg ccc	1488
Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu Pro Pro	
485 490 495	
ttc agc tgc acc tca aca gca aat ggc tac aag gca agc tgc aac ccc	1536
Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys Asn Pro	
500 505 510	
tcc gac acc aac cag aag atc aac gcg ctc tac ccc ttc acc caa gct	1584
Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr Gln Ala	
515 520 525	
aat gtg aag ggc aac acc ggc aag ctc ggc gtg ctg gcc agc ttg gtg	1632
Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser Leu Val	
530 535 540	
agc tac gac ctc aac ccc aag aat gtt ttt gga gag ctg gac agc gac	1680
Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp Ser Asp	
545 550 555 560	
acc aac aat gtc atc cta aag ggc atc ccg gcg gag aag ggc tac ttc	1728
Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly Tyr Phe	
565 570 575	
ccc aac aat gca agg ccg acg gtg gtg aag gag tgg atc aat gga gct	1776
Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn Gly Ala	
580 585 590	
tcg gcg gtg cct ctt gat tca ggc aac acc ttg ttc atg aca gca aca	1824
Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr Ala Thr	
595 600 605	
aat ttg acg gcg acg cag tac agg atc agg atc aga tat gcc aac ccc	1872
Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala Asn Pro	
610 615 620	
aac agc aac acc cag atc ggc gtg agg atc acc caa aat gga agc ctc	1920
Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly Ser Leu	
625 630 635 640	
atc agc agc agc aac ctc acc ctc tac tca aca act gac atg aac aac	1968
Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met Asn Asn	
645 650 655	
acc ttg ccg ctc aat gtt tat gtg att gga gaa aat ggc aac tac acc	2016
Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn Tyr Thr	
660 665 670	
ttg caa gat ctc tac aac acc acc aac gtg ctg agc acc ggc gac atc	2064
Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Gly Asp Ile	
675 680 685	

95/123

```

acc cta cag atc act gga gga gat cag aag atc ttc atc gac agg att 2112
Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp Arg Ile
690 695 700

gaa ttt gtt cca aca atg cca gtt cct ggc aac acc aac aac aac aat 2160
Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn Asn Asn
705 710 715 720

ggc aac aac aat ggc aac aac aac ccg ccg cac cat gtt tgt gcc ata 2208
Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val Cys Ala Ile
725 730 735

gct gga act caa caa agc tgc tct ggg ccg cca aaa ttt gag caa gtt 2256
Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe Glu Gln Val
740 745 750

tca gat ctg gag aag atc acc acc caa gtc tac atg ctc ttc aag agc 2304
Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu Phe Lys Ser
755 760 765

taa 2307
*
```

<210> SEQ ID NO: 31

<211> Comprimento: 768

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: APOSYN2AXMI-031(A-D)

<400> Sequência: 31

```

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

Asn	Leu	Asn	Gly	Ala	Ile	Val	Asp	Ala	Gln	Trp	Ser	Gly	Thr	Val	Asn
			180					185					190		
Thr	Thr	Asn	Arg	Thr	Leu	Arg	Asn	Pro	Thr	Glu	Ser	Asp	Tyr	Thr	Asn
		195					200					205			
Val	Val	Thr	Asn	Phe	Ile	Ala	Ala	Asp	Gly	Asp	Ile	Ala	Asn	Asn	Glu
	210					215					220				
Asn	His	Ile	Met	Asn	Gly	Asn	Phe	Asp	Val	Ala	Ala	Ala	Pro	Tyr	Phe
225					230					235					240
Val	Ile	Gly	Ala	Thr	Ala	Arg	Phe	Ala	Ala	Met	Gln	Ser	Tyr	Ile	Lys
			245					250						255	
Phe	Cys	Asn	Ala	Trp	Ile	Asp	Lys	Val	Gly	Leu	Ser	Asp	Ala	Gln	Leu
		260						265					270		
Thr	Thr	Gln	Lys	Ala	Asn	Leu	Asp	Arg	Thr	Lys	Gln	Asn	Met	Arg	Asn
		275					280					285			
Ala	Ile	Leu	Asn	Tyr	Thr	Gln	Gln	Val	Met	Lys	Val	Phe	Lys	Asp	Ser
	290					295					300				
Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	Ser	Val	Asp	Thr	Tyr
305					310					315					320
Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp	Ile	Val	Ala
			325						330					335	
Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	Ser	Gln	Thr	Ala	Leu
		340						345					350		
Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln	Glu	Glu	Gly
		355					360					365			
Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	Phe	Ser	Tyr
	370					375					380				
Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	Ser	Tyr	Tyr
385					390					395					400
Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr	Pro	Pro	Lys
			405						410					415	
Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu	Asn	Tyr	Ala
		420						425					430		
Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu	Ser	Leu	Gly
		435					440					445			
Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn	Ala	Ala	Thr
	450					455					460				
Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	Asn	Gly	Ile	Gly	Ala
465					470					475					480
Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Pro	Thr	Glu	Pro	Pro
			485						490					495	
Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	Ala	Ser	Cys	Asn	Pro
		500						505					510		
Ser	Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Pro	Phe	Thr	Gln	Ala
		515					520					525			
Asn	Val	Lys	Gly	Asn	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala	Ser	Leu	Val
	530					535					540				
Ser	Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu	Asp	Ser	Asp
545					550					555					560
Thr	Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys	Gly	Tyr	Phe
			565					570						575	
Pro	Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile	Asn	Gly	Ala
		580						585					590		
Ser	Ala	Val	Pro	Leu	Asp	Ser	Gly	Asn	Thr	Leu	Phe	Met	Thr	Ala	Thr
		595					600					605			
Asn	Leu	Thr	Ala	Thr	Gln	Tyr	Arg	Ile	Arg	Ile	Arg	Tyr	Ala	Asn	Pro
610					615						620				
Asn	Ser	Asn	Thr	Gln	Ile	Gly	Val	Arg	Ile	Thr	Gln	Asn	Gly	Ser	Leu

625					630					635					640
Ile	Ser	Ser	Ser	Asn	Leu	Thr	Leu	Tyr	Ser	Thr	Thr	Asp	Met	Asn	Asn
				645					650					655	
Thr	Leu	Pro	Leu	Asn	Val	Tyr	Val	Ile	Gly	Glu	Asn	Gly	Asn	Tyr	Thr
			660					665					670		
Leu	Gln	Asp	Leu	Tyr	Asn	Thr	Thr	Asn	Val	Leu	Ser	Thr	Gly	Asp	Ile
		675					680					685			
Thr	Leu	Gln	Ile	Thr	Gly	Gly	Asp	Gln	Lys	Ile	Phe	Ile	Asp	Arg	Ile
	690					695				700					
Glu	Phe	Val	Pro	Thr	Met	Pro	Val	Pro	Gly	Asn	Thr	Asn	Asn	Asn	Asn
705					710					715					720
Gly	Asn	Asn	Asn	Gly	Asn	Asn	Asn	Pro	Pro	His	His	Val	Cys	Ala	Ile
			725					730						735	
Ala	Gly	Thr	Gln	Gln	Ser	Cys	Ser	Gly	Pro	Pro	Lys	Phe	Glu	Gln	Val
			740					745					750		
Ser	Asp	Leu	Glu	Lys	Ile	Thr	Thr	Gln	Val	Tyr	Met	Leu	Phe	Lys	Ser
		755					760					765			

<210> SEQ ID NO: 32

<211> Comprimento: 3570

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: synaxmi-031(fl)-ER

<221> Nome/Chave: CDS

<222> Localização: (1)...(3570)

<400> Sequência: 32

atg gat tgc aat ctc cag agc cag cag aat atc cca tac aat gtg ctt	48
Met Asp Cys Asn Leu Gln Ser Gln Gln Asn Ile Pro Tyr Asn Val Leu	
1 5 10 15	
gct atc cct gtt agc aat gtt aat agc ctt acg gat acg gtt ggc gat	96
Ala Ile Pro Val Ser Asn Val Asn Ser Leu Thr Asp Thr Val Gly Asp	
20 25 30	
cta aag aag gca tgg gaa gag ttc cag aag acg ggc agc ttc tcc ctt	144
Leu Lys Lys Ala Trp Glu Glu Phe Gln Lys Thr Gly Ser Phe Ser Leu	
35 40 45	
acc gca ctc cag cag ggc ttc agc gct tcc cag ggc ggg acg ttc aat	192
Thr Ala Leu Gln Gln Gly Phe Ser Ala Ser Gln Gly Gly Thr Phe Asn	
50 55 60	
tac ctc acc ctc ctc cag tcc ggg atc tcc ctg gct ggc agc ttc gtg	240
Tyr Leu Thr Leu Leu Gln Ser Gly Ile Ser Leu Ala Gly Ser Phe Val	
65 70 75 80	
cct ggc ggc acc ttc gtt gct cca atc atc aat atg gtg atc ggc tgg	288
Pro Gly Gly Thr Phe Val Ala Pro Ile Ile Asn Met Val Ile Gly Trp	
85 90 95	

ctc tgg cct cat aag aat aag aat gca gat acg gaa aat ctg atc aat	336
Leu Trp Pro His 100 Lys Asn Lys Asn 105 Ala Asp Thr Glu Asn 110 Leu Ile Asn	
ctg atc gat tca gaa atc cag aag cag ctg aat aag gct ctg ctg gat	384
Leu Ile Asp Ser Glu Ile Gln Lys Gln Leu Asn Lys Ala Leu Leu Asp	
115 120 125	
gct gat cgc aat gaa tgg agc agc tac ctc gag tcc atc ttc gat agc	432
Ala Asp Arg Asn Glu Trp Ser Ser Tyr Leu Glu Ser Ile Phe Asp Ser	
130 135 140	
agc aat aat ctg aat ggc gca atc gtt gat gct cag tgg tcc ggc acc	480
Ser Asn Asn Leu Asn Gly Ala Ile Val Asp Ala Gln Trp Ser Gly Thr	
145 150 155 160	
gtg aat acg acc aat cgt acg ctt cgt aat cct acg gaa tca gat tac	528
Val Asn Thr Thr Asn Arg Thr Leu Arg Asn Pro Thr Glu Ser Asp Tyr	
165 170 175	
acc aat gtg gtg acg aat ttc atc gca gca gat ggg gat atc gct aat	576
Thr Asn Val Val Thr Asn Phe Ile Ala Ala Asp Gly Asp Ile Ala Asn	
180 185 190	
aat gaa aat cat atc atg aat ggg aat ttc gat gtg gca gca gct cca	624
Asn Glu Asn His Ile Met Asn Gly Asn Phe Asp Val Ala Ala Ala Pro	
195 200 205	
tac ttc gtt atc ggg gct acc gct aga ttc gca gct atg cag tcc tac	672
Tyr Phe Val Ile Gly Ala Thr Ala Arg Phe Ala Ala Met Gln Ser Tyr	
210 215 220	
atc aag ttc tgc aat gca tgg atc gat aag gtt ggg ctg tca gat gct	720
Ile Lys Phe Cys Asn Ala Trp Ile Asp Lys Val Gly Leu Ser Asp Ala	
225 230 235 240	
cag ctg acc acg cag aag gct aat ctg gat aga acg aag cag aat atg	768
Gln Leu Thr Thr Gln Lys Ala Asn Leu Asp Arg Thr Lys Gln Asn Met	
245 250 255	
cgt aat gct atc ctt aat tac acc cag cag gtt atg aag gtt ttc aag	816
Arg Asn Ala Ile Leu Asn Tyr Thr Gln Gln Val Met Lys Val Phe Lys	
260 265 270	
gat tcc aag aat atg cca acg atc ggc acg aat aag ttc agc gtt gat	864
Asp Ser Lys Asn Met Pro Thr Ile Gly Thr Asn Lys Phe Ser Val Asp	
275 280 285	
acc tac aat gtt tac atc aag ggg atg acg ctt aat gtg ctt gat atc	912
Thr Tyr Asn Val Tyr Ile Lys Gly Met Thr Leu Asn Val Leu Asp Ile	
290 295 300	
gtt gct atc tgg cca agc ctg tac cca gat gat tac acg tca cag acg	960
Val Ala Ile Trp Pro Ser Leu Tyr Pro Asp Asp Tyr Thr Ser Gln Thr	
305 310 315 320	

gct ctg gaa cag acc cgc gtg acg ttc tcc aat atg gtg ggg cag gaa	1008
Ala Leu Glu Gln Thr Arg Val Thr Phe Ser Asn Met Val Gly Gln Glu	
325 330 335	
gaa ggc acg gat ggc agc ctc aga atc tac aat acc ttc gat agc ttc	1056
Glu Gly Thr Asp Gly Ser Leu Arg Ile Tyr Asn Thr Phe Asp Ser Phe	
340 345 350	
tcc tac cag cat agc cct atc cct aat aat aat gtg aat ctc atc agc	1104
Ser Tyr Gln His Ser Pro Ile Pro Asn Asn Asn Val Asn Leu Ile Ser	
355 360 365	
tac tac aat gat gaa ctt cag aat ctg gaa ctc ggg gtt tac acc cca	1152
Tyr Tyr Asn Asp Glu Leu Gln Asn Leu Glu Leu Gly Val Tyr Thr Pro	
370 375 380	
cca aag aag ggc tca ggc tac agc tac cca tac ggg ttc gtg ctg aat	1200
Pro Lys Lys Gly Ser Gly Tyr Ser Tyr Pro Tyr Gly Phe Val Leu Asn	
385 390 395 400	
tac gca aat agc aag tac aag tac ggc gat tcc aat gat cca gaa tcc	1248
Tyr Ala Asn Ser Lys Tyr Lys Tyr Gly Asp Ser Asn Asp Pro Glu Ser	
405 410 415	
ctc ggc ggg ctt tcc acc ctt agc gct cca atc caa cag gtt aat gct	1296
Leu Gly Gly Leu Ser Thr Leu Ser Ala Pro Ile Gln Gln Val Asn Ala	
420 425 430	
gct acc cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg atc	1344
Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly Ile	
435 440 445	
ggc gca agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc gaa	1392
Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr Glu	
450 455 460	
cca cca ttc agc tgc acg agc acg gca aat ggg tac aag gca agc tgc	1440
Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser Cys	
465 470 475 480	
aat cca agc gat acc aat cag aag atc aat gct ctc tac cca ttc acg	1488
Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe Thr	
485 490 495	
cag gca aat gtg aag ggc aat acc ggg aag ctg ggc gtt ctc gca agc	1536
Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala Ser	
500 505 510	
ctt gtt agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc gat	1584
Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu Asp	
515 520 525	
tca gat acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag ggc	1632
Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys Gly	
530 535 540	

tac ttc cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc aat	1680
Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile Asn	
545 550 555 560	
ggg gct tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg acg	1728
Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met Thr	
565 570 575	
gct acg aac ctg acg gct acc cag tac aga atc cgc atc cgt tac gct	1776
Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr Ala	
580 585 590	
aac cca aac tca aac acg cag atc ggc gtt aga atc acg cag aac ggg	1824
Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn Gly	
595 600 605	
agc ctc atc tcc tcc tcc aac ctc acg ctg tac tca acc acc gat atg	1872
Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp Met	
610 615 620	
aac aac acc ctg cca ctg aac gtt tac gtg atc ggg gaa aac ggg aac	1920
Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly Asn	
625 630 635 640	
tac acc ctc cag gat ctc tac aac acg acg aac gtt ctg tcc acg ggc	1968
Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr Gly	
645 650 655	
gat atc acc ctg cag atc acc ggc ggg gat cag aag ata ttc atc gat	2016
Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile Asp	
660 665 670	
cgc atc gag ttc gtg cct acg atg cca gtg cca ggc aac acc aac aac	2064
Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn Asn	
675 680 685	
aac aac ggg aac aac aac ggc aac aac aac cct cct cat cat gtt tgc	2112
Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val Cys	
690 695 700	
gct atc gca ggc acc cag cag tcc tgc tcc ggc cca cca aag ttc gag	2160
Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe Glu	
705 710 715 720	
cag gtg tcc gat tta gaa aag ata acc acg cag gtt tac atg ctc ttc	2208
Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu Phe	
725 730 735	
aag tcc tcc cca tac gaa gaa ctt gct ctt gaa gtt agc agc tac cag	2256
Lys Ser Ser Pro Tyr Glu Glu Leu Ala Leu Glu Val Ser Ser Tyr Gln	
740 745 750	
atc agc cag gtt gca ctg aag gtt atg gct ctc agc gat gaa cta ttc	2304
Ile Ser Gln Val Ala Leu Lys Val Met Ala Leu Ser Asp Glu Leu Phe	
755 760 765	

tgc gaa gaa aag aac gtg ctg aga aag ctc gtg aac aag gca aag cag	2352
Cys Glu Glu Lys Asn Val Leu Arg Lys Leu Val Asn Lys Ala Lys Gln	
770 775 780	
ctt ctc gaa gca agc aac ctt ctc gtt ggg ggg aac ttc gag acg act	2400
Leu Leu Glu Ala Ser Asn Leu Leu Val Gly Gly Asn Phe Glu Thr Thr	
785 790 795 800	
cag aac tgg gtg ctc ggc acg aac gct tac atc aac tat gat tca ttc	2448
Gln Asn Trp Val Leu Gly Thr Asn Ala Tyr Ile Asn Tyr Asp Ser Phe	
805 810 815	
ctg ttc aac ggg aac tac ctg agc ctc cag cca gca tcc ggc ttc ttc	2496
Leu Phe Asn Gly Asn Tyr Leu Ser Leu Gln Pro Ala Ser Gly Phe Phe	
820 825 830	
acg agc tac gct tac cag aag atc gat gaa tcc acg ctt aag cct tac	2544
Thr Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Thr Leu Lys Pro Tyr	
835 840 845	
acg cgc tac aag gtg tcc ggg ttc atc ggg cag tca aac cag gtt gaa	2592
Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val Glu	
850 855 860	
ctg atc atc agc cgc tac ggg aag gaa atc gat aag atc ctt aac gtg	2640
Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val	
865 870 875 880	
cca tac gca ggg cca ctc cca atc acc gca gat gca agc atc acc tgc	2688
Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys	
885 890 895	
tgc gct cca gaa ata ggg cag tgc gat ggg gaa cag tca gat tcc cat	2736
Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His	
900 905 910	
ttc ttc aac tac agc atc gat gtt ggc gca ctc cat cca gaa ctt aac	2784
Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn	
915 920 925	
cca ggg atc gaa atc ggg ctg aag atc gtt cag agc aac ggc tac atc	2832
Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile	
930 935 940	
acc atc tcc aac ctt gaa atc atc gaa gaa aga cca tta acg gaa atg	2880
Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met	
945 950 955 960	
gaa att caa gca gtg aac cgc aag aac cag aag tgg gaa cgc gaa aag	2928
Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys	
965 970 975	
ctg ctt gaa tgc gca tcc atc tcc gaa ctc ctc cag cct atc atc aac	2976
Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn	
980 985 990	

102/123

cag ata gat agc ctc ttc aag gat ggc aac tgg tac aac gat atc cta	3024
Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu	
995 1000 1005	
cca cat gtt acc tac cag gat ctt aag aac atc atc atc cca gaa ctt	3072
Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu	
1010 1015 1020	
cct aag ctc aag cat tgg ttc atc gaa aac ctg cct ggg gaa tac cat	3120
Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His	
1025 1030 1035 1040	
gaa atc gaa cag aag atg aag gaa gca tta aag tac gca ttc act cag	3168
Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln	
1045 1050 1055	
ctc gat gaa aag aac ctg atc cat aac ggc cat ttc acc acc aac ctc	3216
Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu	
1060 1065 1070	
atc gat tgg cag gtg gaa ggc gat gct cag atg aag gtg ctt gaa aac	3264
Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn	
1075 1080 1085	
gat gca ctc gct ctc cag cta ttc aac tgg gat gct tcc gca tcc cag	3312
Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser Gln	
1090 1095 1100	
agc atc aac atc cta gag ttc gat gaa gat aag gct tac aag ctg cgt	3360
Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg	
1105 1110 1115 1120	
gtg tac gca cag ggc tcc ggg acg atc cag ttc ggc aac tgc gaa gat	3408
Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp	
1125 1130 1135	
gaa gct atc cag ttc aac acc aac tca ttc atc tac cag gaa aag ata	3456
Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile	
1140 1145 1150	
gtg tac ttc gat acg cca tcc gtt aac ctt cat atc cag agc gaa ggc	3504
Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly	
1155 1160 1165	
tcc gag ttc atc gtg agc agc atc gat ctc atc gaa ctc agc gat gat	3552
Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp	
1170 1175 1180	
cag aag gat gaa ctg taa	3570
Gln Lys Asp Glu Leu *	
1185	

<210> SEQ ID NO: 33

<211> Comprimento: 1189

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: SYNAXMI-031(FL)-ER

<400> Sequência: 33

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

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

105/123

```

Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn Val
865                               870                               875                               880
Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr Cys
                               885                               890                               895
Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser His
                               900                               905                               910
Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu Asn
                               915                               920                               925
Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr Ile
930                               935                               940
Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu Met
945                               950                               955                               960
Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu Lys
                               965                               970                               975
Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Leu Gln Pro Ile Ile Asn
                               980                               985                               990
Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile Leu
995                               1000                               1005
Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu Leu
1010                               1015                               1020
Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr His
1025                               1030                               1035                               1040
Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr Gln
                               1045                               1050                               1055
Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn Leu
                               1060                               1065                               1070
Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu Asn
1075                               1080                               1085
Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser Gln
1090                               1095                               1100
Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu Arg
1105                               1110                               1115                               1120
Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu Asp
                               1125                               1130                               1135
Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys Ile
1140                               1145                               1150
Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu Gly
1155                               1160                               1165
Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp Asp
1170                               1175                               1180
Gln Lys Asp Glu Leu
1185

```

<210> SEQ ID NO: 34

<211> Comprimento: 3669

<212> Tipo: DNA

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo (AXMI-049)

<221> Nome/Chave: CDS

<222> Localização: (1)...(3669)

<400> Sequência: 34

atg gat gac atg aca aat tta tcc gat gta tat tca ccc gta cct tcc	48
Met Asp Asp Met Thr Asn Leu Ser Asp Val Tyr Ser Pro Val Pro Ser	
1 5 10 15	
aat gta tta gca gct cca ctt att ctt gaa gaa aca aag aaa aaa aca	96
Asn Val Leu Ala Ala Pro Leu Ile Leu Glu Glu Thr Lys Lys Lys Thr	
20 25 30	
ccg gct gaa caa gca aaa gaa gat tta gaa aag gca ttg aga aca gga	144
Pro Ala Glu Gln Ala Lys Glu Asp Leu Glu Lys Ala Leu Arg Thr Gly	
35 40 45	
aag ttt tcc gat gca att gca caa att tta aat gat gtt ctt att aac	192
Lys Phe Ser Asp Ala Ile Ala Gln Ile Leu Asn Asp Val Leu Ile Asn	
50 55 60	
caa aag ttc agc tat caa aca gcg gtt aca gtt tcc tta tct ttg gct	240
Gln Lys Phe Ser Tyr Gln Thr Ala Val Thr Val Ser Leu Ser Leu Ala	
65 70 75 80	
agt ata gtt ctt cca gaa ata ggt ttt ttt gcc ccc ttt gtt ggt tta	288
Ser Ile Val Leu Pro Glu Ile Gly Phe Phe Ala Pro Phe Val Gly Leu	
85 90 95	
ttt ttt tct gct tta aat aaa tct caa aat ata cct acg act tca gat	336
Phe Phe Ser Ala Leu Asn Lys Ser Gln Asn Ile Pro Thr Thr Ser Asp	
100 105 110	
att ttt gaa gcg atg aaa cca gct att caa aca atg att gat cgt agt	384
Ile Phe Glu Ala Met Lys Pro Ala Ile Gln Thr Met Ile Asp Arg Ser	
115 120 125	
tta aca gat gct gaa aac aag gag atg gat gat cag gcg cac aac cta	432
Leu Thr Asp Ala Glu Asn Lys Glu Met Asp Asp Gln Ala His Asn Leu	
130 135 140	
ttt acg cga tta caa act tat caa gag caa ata gat ctc tat aaa cat	480
Phe Thr Arg Leu Gln Thr Tyr Gln Glu Gln Ile Asp Leu Tyr Lys His	
145 150 155 160	
att tta gat gca aaa caa aaa cca act ctt aac gat ata gga gat ctt	528
Ile Leu Asp Ala Lys Gln Lys Pro Thr Leu Asn Asp Ile Gly Asp Leu	
165 170 175	
cac act tcg atc gac gaa act ctc aga aca tta gat tcc aat tta gca	576
His Thr Ser Ile Asp Glu Thr Leu Arg Thr Leu Asp Ser Asn Leu Ala	
180 185 190	
ttt ttc caa aca gaa ggc tat caa gat ctc ggg tta cca tat tac aca	624
Phe Phe Gln Thr Glu Gly Tyr Gln Asp Leu Gly Leu Pro Tyr Tyr Thr	
195 200 205	
att ttt gcc aca caa tat ttg tta atc ctt tca gat aaa att aaa act	672
Ile Phe Ala Thr Gln Tyr Leu Leu Ile Leu Ser Asp Lys Ile Lys Thr	
210 215 220	

ggt atc aca tgg gga tat aat cct att aac att cca gat ttt caa aac	720
Gly Ile Thr Trp Gly Tyr Asn Pro Ile Asn Ile Pro Asp Phe Gln Asn	
225 230 235 240	
caa ttt aat aat aga ata ctt tta ttt aca aaa tac att agt ggt caa	768
Gln Phe Asn Asn Arg Ile Leu Leu Phe Thr Lys Tyr Ile Ser Gly Gln	
245 250 255	
ctt aag aaa atg tat gac aac ggt gta cat cca gca ctt tta tat caa	816
Leu Lys Lys Met Tyr Asp Asn Gly Val His Pro Ala Leu Leu Tyr Gln	
260 265 270	
aac tgt atc caa ttt gtt gct tta tgg cct act ttt tct ccc gca gat	864
Asn Cys Ile Gln Phe Val Ala Leu Trp Pro Thr Phe Ser Pro Ala Asp	
275 280 285	
tat aac ctt agt aac agc aca gat tta gaa caa aca ata agt ttt aag	912
Tyr Asn Leu Ser Asn Ser Thr Asp Leu Glu Gln Thr Ile Ser Phe Lys	
290 295 300	
agt ttc atc tcc tat aag agt tta aat gat tat aat tat gca ctt cct	960
Ser Phe Ile Ser Tyr Lys Ser Leu Asn Asp Tyr Asn Tyr Ala Leu Pro	
305 310 315 320	
gaa att aat aca tgg aca aat ttt gaa atg aca aaa tca ctt gat ttt	1008
Glu Ile Asn Thr Trp Thr Asn Phe Glu Met Thr Lys Ser Leu Asp Phe	
325 330 335	
aac aac tgt aga tcc atg gat aat tat ggt atc gcg ggt gac ctt cgt	1056
Asn Asn Cys Arg Ser Met Asp Asn Tyr Gly Ile Ala Gly Asp Leu Arg	
340 345 350	
ata aca aat cat aaa aat gaa atc cat gaa ata tct gct cct tcg gta	1104
Ile Thr Asn His Lys Asn Glu Ile His Glu Ile Ser Ala Pro Ser Val	
355 360 365	
ctc tac atg aca tca aat cag tgt aac tca aga ctt gtt ttt cca ttc	1152
Leu Tyr Met Thr Ser Asn Gln Cys Asn Ser Arg Leu Val Phe Pro Phe	
370 375 380	
gat gat cct att gtt aaa ttc gaa gga aca agt agt gtt ttc ata ggg	1200
Asp Asp Pro Ile Val Lys Phe Glu Gly Thr Ser Ser Val Phe Ile Gly	
385 390 395 400	
aca cct gga tta cct caa ttt tat aca aat ttt cag ttg act tct ggt	1248
Thr Pro Gly Leu Pro Gln Phe Tyr Thr Asn Phe Gln Leu Thr Ser Gly	
405 410 415	
aaa agt tta ata ttc cct cct cct tca gga gtt aat caa caa ggt ttc	1296
Lys Ser Leu Ile Phe Pro Pro Pro Ser Gly Val Asn Gln Gln Gly Phe	
420 425 430	
aca caa aat tat tcg gta atc cct cat cct gga ttt aaa ata gca ggt	1344
Thr Gln Asn Tyr Ser Val Ile Pro His Pro Gly Phe Lys Ile Ala Gly	
435 440 445	

att	aca	aat	atg	tca	tat	tta	cct	act	cat	cac	gat	agt	aat	gcg	cgg	1392
Ile	Thr	Asn	Met	Ser	Tyr	Leu	Pro	Thr	His	His	Asp	Ser	Asn	Ala	Arg	
	450					455					460					
aca	caa	ata	gaa	cat	atc	caa	gtt	cct	gaa	aaa	atc	ttt	cca	gaa	aat	1440
Thr	Gln	Ile	Glu	His	Ile	Gln	Val	Pro	Glu	Lys	Ile	Phe	Pro	Glu	Asn	
465					470					475					480	
att	atc	ggg	gtt	cca	gat	cca	gat	aat	aac	aat	cta	atc	cca	atc	aaa	1488
Ile	Ile	Gly	Val	Pro	Asp	Pro	Asp	Asn	Asn	Asn	Leu	Ile	Pro	Ile	Lys	
				485					490					495		
ggt	att	ccc	gca	gaa	aaa	gga	tat	ggt	gac	tca	att	gca	tat	gtg	tca	1536
Gly	Ile	Pro	Ala	Glu	Lys	Gly	Tyr	Gly	Asp	Ser	Ile	Ala	Tyr	Val	Ser	
			500					505					510			
gaa	ccg	gta	aat	ggt	gcg	agt	gca	gtt	aaa	ctt	act	tca	aat	caa	att	1584
Glu	Pro	Val	Asn	Gly	Ala	Ser	Ala	Val	Lys	Leu	Thr	Ser	Asn	Gln	Ile	
		515					520					525				
ctc	caa	atg	gaa	att	aca	aat	gta	aca	act	caa	aaa	tat	caa	gtt	cgc	1632
Leu	Gln	Met	Glu	Ile	Thr	Asn	Val	Thr	Thr	Gln	Lys	Tyr	Gln	Val	Arg	
	530					535					540					
ata	cgt	tat	gct	aca	gct	gga	gat	aca	gag	gct	aat	ata	agg	ttc	cat	1680
Ile	Arg	Tyr	Ala	Thr	Ala	Gly	Asp	Thr	Glu	Ala	Asn	Ile	Arg	Phe	His	
545					550					555					560	
att	att	gat	cca	aat	gaa	aat	aat	tta	ata	aat	ggg	cct	aat	cat	ttc	1728
Ile	Ile	Asp	Pro	Asn	Glu	Asn	Asn	Leu	Ile	Asn	Gly	Pro	Asn	His	Phe	
				565				570						575		
aca	gct	gta	tct	aat	act	caa	acg	tct	gtc	caa	ggg	gaa	aat	gga	aaa	1776
Thr	Ala	Val	Ser	Asn	Thr	Gln	Thr	Ser	Val	Gln	Gly	Glu	Asn	Gly	Lys	
			580					585					590			
tat	gta	cta	aac	aca	ctt	gta	aat	tca	ata	ata	tta	cca	tca	gga	aaa	1824
Tyr	Val	Leu	Asn	Thr	Leu	Val	Asn	Ser	Ile	Ile	Leu	Pro	Ser	Gly	Lys	
		595					600					605				
caa	aag	gtc	ttt	att	caa	aac	act	ggg	tct	caa	gat	ctc	ttt	tta	gac	1872
Gln	Lys	Val	Phe	Ile	Gln	Asn	Thr	Gly	Ser	Gln	Asp	Leu	Phe	Leu	Asp	
	610					615					620					
cgt	att	gaa	ttt	att	cca	tta	caa	cta	gaa	ctt	cct	ttc	act	tca	aaa	1920
Arg	Ile	Glu	Phe	Ile	Pro	Leu	Gln	Leu	Glu	Leu	Pro	Phe	Thr	Ser	Lys	
625					630					635					640	
tta	cct	gaa	act	act	aca	caa	cca	aat	aca	aca	aaa	aca	att	tgg	tca	1968
Leu	Pro	Glu	Thr	Thr	Thr	Gln	Pro	Asn	Thr	Thr	Lys	Thr	Ile	Trp	Ser	
				645					650					655		
ggt	caa	aaa	cct	gct	aat	aca	ctt	tct	ctt	caa	ggg	aca	gtt	tat	aat	2016
Gly	Gln	Lys	Pro	Ala	Asn	Thr	Leu	Ser	Leu	Gln	Gly	Thr	Val	Tyr	Asn	
			660					665					670			

gat gct tct atc gaa tta caa ctt tat atg aac gat aac tta gtc caa Asp Ala Ser Ile Glu Leu Gln Leu Tyr Met Asn Asp Asn Leu Val Gln 675 680 685	2064
aaa atc cct gca caa ggt cct ggt cct agt ttt gac tgt gac gac caa Lys Ile Pro Ala Gln Gly Pro Gly Pro Ser Phe Asp Cys Asp Asp Gln 690 695 700	2112
tct aaa cct att aat caa cca aat ata aaa act gaa gaa ttt aac aaa Ser Lys Pro Ile Asn Gln Pro Asn Ile Lys Thr Glu Glu Phe Asn Lys 705 710 715 720	2160
ctt gtc tta aaa gaa tta agt agt acc tat tcg tat tgt atg gga gga Leu Val Leu Lys Glu Leu Ser Ser Thr Tyr Ser Tyr Cys Met Gly Gly 725 730 735	2208
gct ttt gaa aac act tac caa att gat att aca ata gac agc aaa tcc Ala Phe Glu Asn Thr Tyr Gln Ile Asp Ile Thr Ile Asp Ser Lys Ser 740 745 750	2256
caa tct ttt act act cca gaa gat tta gaa aaa atc aca aac caa gtc Gln Ser Phe Thr Thr Pro Glu Asp Leu Glu Lys Ile Thr Asn Gln Val 755 760 765	2304
aac cag tta ttt act tcc tca tcc caa aca aaa ttg gtt caa acc gta Asn Gln Leu Phe Thr Ser Ser Ser Gln Thr Lys Leu Val Gln Thr Val 770 775 780	2352
acg gat tat gga att gat caa atg gta atg aaa gta gat gcg tta tca Thr Asp Tyr Gly Ile Asp Gln Met Val Met Lys Val Asp Ala Leu Ser 785 790 795 800	2400
gac gat gta ttt ggt gtc gag aaa aaa gca tta cgt aaa ctt gtc aat Asp Asp Val Phe Gly Val Glu Lys Lys Ala Leu Arg Lys Leu Val Asn 805 810 815	2448
cag gcc aaa caa tta agt aaa gta cga aat gta ctg gtc ggt gga aac Gln Ala Lys Gln Leu Ser Lys Val Arg Asn Val Leu Val Gly Gly Asn 820 825 830	2496
ttt gaa aaa ggt cat aaa tgg gta cta ggt cgt aaa gcg aca acg gta Phe Glu Lys Gly His Lys Trp Val Leu Gly Arg Lys Ala Thr Thr Val 835 840 845	2544
gcg gat cat gat tta ttc aaa ggg gat cat tta tta tta cca cca cca Ala Asp His Asp Leu Phe Lys Gly Asp His Leu Leu Leu Pro Pro Pro 850 855 860	2592
acc ctg tat cca tcg tat gcg tat caa aaa atc gat gaa tct aaa tta Thr Leu Tyr Pro Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Lys Leu 865 870 875 880	2640
aaa tcc aat aca cgc tat acg gtt tcc ggt ttt gtt gcg caa agt gaa Lys Ser Asn Thr Arg Tyr Thr Val Ser Gly Phe Val Ala Gln Ser Glu 885 890 895	2688

cat tta gaa gtt gtt gtt tct cgc tat ggg aaa gaa gta aat acc ctg	2736
His Leu Glu Val Val Val Ser Arg Tyr Gly Lys Glu Val Asn Thr Leu	
900 905 910	
tta cat gtc cct tat gaa gaa gca tta ccg att tct tcc gat gag cgt	2784
Leu His Val Pro Tyr Glu Glu Ala Leu Pro Ile Ser Ser Asp Glu Arg	
915 920 925	
cca aat tgc tgt aaa cca gct gct tgt cag tgt cca tct tgc aat ggt	2832
Pro Asn Cys Cys Lys Pro Ala Ala Cys Gln Cys Pro Ser Cys Asn Gly	
930 935 940	
gat gca cca gac tcc cat ttc ttt agc tat agt atc gat gtt ggt tcc	2880
Asp Ala Pro Asp Ser His Phe Phe Ser Tyr Ser Ile Asp Val Gly Ser	
945 950 955 960	
tta caa gca gat gta aat tta gga att gag ttt ggt ctt cgt att gtg	2928
Leu Gln Ala Asp Val Asn Leu Gly Ile Glu Phe Gly Leu Arg Ile Val	
965 970 975	
aaa tcc aac gga ttt gca aaa atc agt aac cta gaa atc aaa gaa gac	2976
Lys Ser Asn Gly Phe Ala Lys Ile Ser Asn Leu Glu Ile Lys Glu Asp	
980 985 990	
cgt cca tta aca gaa aaa gaa att aag aaa ata caa cgc aaa gaa caa	3024
Arg Pro Leu Thr Glu Lys Glu Ile Lys Lys Ile Gln Arg Lys Glu Gln	
995 1000 1005	
aag tgg aaa aaa gca ttt gac aaa gaa cag gca gag tta acg gca aca	3072
Lys Trp Lys Lys Ala Phe Asp Lys Glu Gln Ala Glu Leu Thr Ala Thr	
1010 1015 1020	
ctc caa cca acc ctg aac caa atc aat gcc tta tat caa aat gaa gat	3120
Leu Gln Pro Thr Leu Asn Gln Ile Asn Ala Leu Tyr Gln Asn Glu Asp	
1025 1030 1035 1040	
tgg aac ggt tcg att cac cct cat gta acg tat caa cat cta tcc gat	3168
Trp Asn Gly Ser Ile His Pro His Val Thr Tyr Gln His Leu Ser Asp	
1045 1050 1055	
gtt gtc gta cca gca tta cca aaa caa aga cat tgg ttt atg gaa gat	3216
Val Val Val Pro Ala Leu Pro Lys Gln Arg His Trp Phe Met Glu Asp	
1060 1065 1070	
cga caa ggt gaa cat tac aat gta aca caa caa ttc caa caa gca tta	3264
Arg Gln Gly Glu His Tyr Asn Val Thr Gln Gln Phe Gln Gln Ala Leu	
1075 1080 1085	
gat cgt gct ttc caa caa atc gaa gaa caa aac tta att cac aat ggt	3312
Asp Arg Ala Phe Gln Gln Ile Glu Glu Gln Asn Leu Ile His Asn Gly	
1090 1095 1100	
agc ttt gcg aat gga ttg aca gat tgg act gtc aca ggg gat gca cat	3360
Ser Phe Ala Asn Gly Leu Thr Asp Trp Thr Val Thr Gly Asp Ala His	
1105 1110 1115 1120	

111/123

gtt act atc cag gat gat gat caa gta tta gaa cta tct cat tgg gat 3408
Val Thr Ile Gln Asp Asp Asp Gln Val Leu Glu Leu Ser His Trp Asp
1125 1130 1135

gca agt gtc tct caa acg att gaa att att gat ttt gaa gaa gaa aaa 3456
Ala Ser Val Ser Gln Thr Ile Glu Ile Ile Asp Phe Glu Glu Glu Lys
1140 1145 1150

gaa tac aaa ctt cgt gta cgt gga aaa ggt aaa gga acg gta acc gtt 3504
Glu Tyr Lys Leu Arg Val Arg Gly Lys Gly Lys Gly Thr Val Thr Val
1155 1160 1165

caa cat gga gaa gaa gag tta gaa aca atg aca ttt aat gca acg agt 3552
Gln His Gly Glu Glu Glu Leu Glu Thr Met Thr Phe Asn Ala Thr Ser
1170 1175 1180

ttt aca acg caa gaa caa acc ttc tat ttc gaa gga aat aca gtg gat 3600
Phe Thr Thr Gln Glu Gln Thr Phe Tyr Phe Glu Gly Asn Thr Val Asp
1185 1190 1195 1200

ata cac gtt caa tca gag aat aat aca ttc ctg gta gac agt gta gaa 3648
Ile His Val Gln Ser Glu Asn Asn Thr Phe Leu Val Asp Ser Val Glu
1205 1210 1215

ctc att gaa att ata gaa aag 3669
Leu Ile Glu Ile Ile Glu Lys
1220

<210> SEQ ID NO: 35

<211> Comprimento: 1223

<212> Tipo: PRT

<213> Organismo: Desconhecido

<220> Características:

<223> Outras Informações: Isolado de amostra de solo (AXMI-049)

<400> Sequência: 35

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

Phe	Thr	Arg	Leu	Gln	Thr	Tyr	Gln	Glu	Gln	Ile	Asp	Leu	Tyr	Lys	His
145					150					155					160
Ile	Leu	Asp	Ala	Lys	Gln	Lys	Pro	Thr	Leu	Asn	Asp	Ile	Gly	Asp	Leu
				165					170						175
His	Thr	Ser	Ile	Asp	Glu	Thr	Leu	Arg	Thr	Leu	Asp	Ser	Asn	Leu	Ala
			180					185					190		
Phe	Phe	Gln	Thr	Glu	Gly	Tyr	Gln	Asp	Leu	Gly	Leu	Pro	Tyr	Tyr	Thr
		195					200					205			
Ile	Phe	Ala	Thr	Gln	Tyr	Leu	Leu	Ile	Leu	Ser	Asp	Lys	Ile	Lys	Thr
	210					215					220				
Gly	Ile	Thr	Trp	Gly	Tyr	Asn	Pro	Ile	Asn	Ile	Pro	Asp	Phe	Gln	Asn
225					230					235					240
Gln	Phe	Asn	Asn	Arg	Ile	Leu	Leu	Phe	Thr	Lys	Tyr	Ile	Ser	Gly	Gln
				245					250					255	
Leu	Lys	Lys	Met	Tyr	Asp	Asn	Gly	Val	His	Pro	Ala	Leu	Leu	Tyr	Gln
			260					265					270		
Asn	Cys	Ile	Gln	Phe	Val	Ala	Leu	Trp	Pro	Thr	Phe	Ser	Pro	Ala	Asp
	275						280					285			
Tyr	Asn	Leu	Ser	Asn	Ser	Thr	Asp	Leu	Glu	Gln	Thr	Ile	Ser	Phe	Lys
	290					295					300				
Ser	Phe	Ile	Ser	Tyr	Lys	Ser	Leu	Asn	Asp	Tyr	Asn	Tyr	Ala	Leu	Pro
305					310					315					320
Glu	Ile	Asn	Thr	Trp	Thr	Asn	Phe	Glu	Met	Thr	Lys	Ser	Leu	Asp	Phe
			325					330						335	
Asn	Asn	Cys	Arg	Ser	Met	Asp	Asn	Tyr	Gly	Ile	Ala	Gly	Asp	Leu	Arg
			340					345					350		
Ile	Thr	Asn	His	Lys	Asn	Glu	Ile	His	Glu	Ile	Ser	Ala	Pro	Ser	Val
		355				360						365			
Leu	Tyr	Met	Thr	Ser	Asn	Gln	Cys	Asn	Ser	Arg	Leu	Val	Phe	Pro	Phe
	370					375					380				
Asp	Asp	Pro	Ile	Val	Lys	Phe	Glu	Gly	Thr	Ser	Ser	Val	Phe	Ile	Gly
385					390					395					400
Thr	Pro	Gly	Leu	Pro	Gln	Phe	Tyr	Thr	Asn	Phe	Gln	Leu	Thr	Ser	Gly
				405					410					415	
Lys	Ser	Leu	Ile	Phe	Pro	Pro	Pro	Ser	Gly	Val	Asn	Gln	Gln	Gly	Phe
			420					425					430		
Thr	Gln	Asn	Tyr	Ser	Val	Ile	Pro	His	Pro	Gly	Phe	Lys	Ile	Ala	Gly
		435					440					445			
Ile	Thr	Asn	Met	Ser	Tyr	Leu	Pro	Thr	His	His	Asp	Ser	Asn	Ala	Arg
	450					455					460				
Thr	Gln	Ile	Glu	His	Ile	Gln	Val	Pro	Glu	Lys	Ile	Phe	Pro	Glu	Asn
465					470					475					480
Ile	Ile	Gly	Val	Pro	Asp	Pro	Asp	Asn	Asn	Asn	Leu	Ile	Pro	Ile	Lys
				485					490					495	
Gly	Ile	Pro	Ala	Glu	Lys	Gly	Tyr	Gly	Asp	Ser	Ile	Ala	Tyr	Val	Ser
			500					505					510		
Glu	Pro	Val	Asn	Gly	Ala	Ser	Ala	Val	Lys	Leu	Thr	Ser	Asn	Gln	Ile
		515					520					525			
Leu	Gln	Met	Glu	Ile	Thr	Asn	Val	Thr	Thr	Gln	Lys	Tyr	Gln	Val	Arg
	530					535					540				
Ile	Arg	Tyr	Ala	Thr	Ala	Gly	Asp	Thr	Glu	Ala	Asn	Ile	Arg	Phe	His
545					550					555					560
Ile	Ile	Asp	Pro	Asn	Glu	Asn	Asn	Leu	Ile	Asn	Gly	Pro	Asn	His	Phe
				565					570					575	
Thr	Ala	Val	Ser	Asn	Thr	Gln	Thr	Ser	Val	Gln	Gly	Glu	Asn	Gly	Lys
			580					585					590		

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

114/123

```

Trp Asn Gly Ser Ile His Pro His Val Thr Tyr Gln His Leu Ser Asp
      1045      1050      1055
Val Val Val Pro Ala Leu Pro Lys Gln Arg His Trp Phe Met Glu Asp
      1060      1065      1070
Arg Gln Gly Glu His Tyr Asn Val Thr Gln Gln Phe Gln Gln Ala Leu
      1075      1080      1085
Asp Arg Ala Phe Gln Gln Ile Glu Glu Gln Asn Leu Ile His Asn Gly
      1090      1095      1100
Ser Phe Ala Asn Gly Leu Thr Asp Trp Thr Val Thr Gly Asp Ala His
      1105      1110      1115      1120
Val Thr Ile Gln Asp Asp Gln Val Leu Glu Leu Ser His Trp Asp
      1125      1130      1135
Ala Ser Val Ser Gln Thr Ile Glu Ile Ile Asp Phe Glu Glu Glu Lys
      1140      1145      1150
Glu Tyr Lys Leu Arg Val Arg Gly Lys Gly Lys Gly Thr Val Thr Val
      1155      1160      1165
Gln His Gly Glu Glu Glu Leu Glu Thr Met Thr Phe Asn Ala Thr Ser
      1170      1175      1180
Phe Thr Thr Gln Glu Gln Thr Phe Tyr Phe Glu Gly Asn Thr Val Asp
      1185      1190      1195      1200
Ile His Val Gln Ser Glu Asn Asn Thr Phe Leu Val Asp Ser Val Glu
      1205      1210      1215
Leu Ile Glu Ile Ile Glu Lys
      1220

```

<210> SEQ ID NO: 36
 <211> Comprimento: 4
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: Peptídeo sinal

<400> Sequência: 36
 Lys Asp Glu Leu
 1

<210> SEQ ID NO: 37
 <211> Comprimento: 3657
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: aposynaxmi-031(fl)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(3657)

<400> Sequência: 37
 atg cag atc ctt ggg tac tcc agc ttc gtt gct atc gct ctt ctt atg 48
 Met Gln Ile Leu Gly Tyr Ser Ser Phe Val Ala Ile Ala Leu Leu Met
 1 5 10 15

agc	gtg	gtt	gtt	gtt	tgc	aat	ggg	ggc	aag	acg	tcc	acc	tac	gtg	cgt	96
Ser	Val	Val	Val	Val	Cys	Asn	Gly	Gly	Lys	Thr	Ser	Thr	Tyr	Val	Arg	
			20					25					30			
aat	ctg	gat	tgc	aat	ctc	cag	agc	cag	cag	aat	atc	cca	tac	aat	gtg	144
Asn	Leu	Asp	Cys	Asn	Leu	Gln	Ser	Gln	Gln	Asn	Ile	Pro	Tyr	Asn	Val	
		35					40					45				
ctt	gct	atc	cct	gtt	agc	aat	gtt	aat	agc	ctt	acg	gat	acg	gtt	ggc	192
Leu	Ala	Ile	Pro	Val	Ser	Asn	Val	Asn	Ser	Leu	Thr	Asp	Thr	Val	Gly	
	50					55					60					
gat	cta	aag	aag	gca	tgg	gaa	gag	ttc	cag	aag	acg	ggc	agc	ttc	tcc	240
Asp	Leu	Lys	Lys	Ala	Trp	Glu	Glu	Phe	Gln	Lys	Thr	Gly	Ser	Phe	Ser	
65					70				75						80	
ctt	acc	gca	ctc	cag	cag	ggc	ttc	agc	gct	tcc	cag	ggc	ggg	acg	ttc	288
Leu	Thr	Ala	Leu	Gln	Gln	Gly	Phe	Ser	Ala	Ser	Gln	Gly	Gly	Thr	Phe	
				85				90						95		
aat	tac	ctc	acc	ctc	ctc	cag	tcc	ggg	atc	tcc	ctg	gct	ggc	agc	ttc	336
Asn	Tyr	Leu	Thr	Leu	Leu	Gln	Ser	Gly	Ile	Ser	Leu	Ala	Gly	Ser	Phe	
			100					105					110			
gtg	cct	ggc	ggc	acc	ttc	gtt	gct	cca	atc	atc	aat	atg	gtg	atc	ggc	384
Val	Pro	Gly	Gly	Thr	Phe	Val	Ala	Pro	Ile	Ile	Asn	Met	Val	Ile	Gly	
		115					120					125				
tgg	ctc	tgg	cct	cat	aag	aat	aag	aat	gca	gat	acg	gaa	aat	ctg	atc	432
Trp	Leu	Trp	Pro	His	Lys	Asn	Lys	Asn	Ala	Asp	Thr	Glu	Asn	Leu	Ile	
	130					135					140					
aat	ctg	atc	gat	tca	gaa	atc	cag	aag	cag	ctg	aat	aag	gct	ctg	ctg	480
Asn	Leu	Ile	Asp	Ser	Glu	Ile	Gln	Lys	Gln	Leu	Asn	Lys	Ala	Leu	Leu	
145					150					155					160	
gat	gct	gat	cgc	aat	gaa	tgg	agc	agc	tac	ctc	gag	tcc	atc	ttc	gat	528
Asp	Ala	Asp	Arg	Asn	Glu	Trp	Ser	Ser	Tyr	Leu	Glu	Ser	Ile	Phe	Asp	
				165					170					175		
agc	agc	aat	aat	ctg	aat	ggc	gca	atc	gtt	gat	gct	cag	tgg	tcc	ggc	576
Ser	Ser	Asn	Asn	Leu	Asn	Gly	Ala	Ile	Val	Asp	Ala	Gln	Trp	Ser	Gly	
			180					185					190			
acc	gtg	aat	acg	acc	aat	cgt	acg	ctt	cgt	aat	cct	acg	gaa	tca	gat	624
Thr	Val	Asn	Thr	Thr	Asn	Arg	Thr	Leu	Arg	Asn	Pro	Thr	Glu	Ser	Asp	
		195					200					205				
tac	acc	aat	gtg	gtg	acg	aat	ttc	atc	gca	gca	gat	ggg	gat	atc	gct	672
Tyr	Thr	Asn	Val	Val	Thr	Asn	Phe	Ile	Ala	Ala	Asp	Gly	Asp	Ile	Ala	
	210					215					220					
aat	aat	gaa	aat	cat	atc	atg	aat	ggg	aat	ttc	gat	gtg	gca	gca	gct	720
Asn	Asn	Glu	Asn	His	Ile	Met	Asn	Gly	Asn	Phe	Asp	Val	Ala	Ala	Ala	
225					230					235					240	

cca	tac	ttc	gtt	atc	ggg	gct	acc	gct	aga	ttc	gca	gct	atg	cag	tcc	768
Pro	Tyr	Phe	Val	Ile	Gly	Ala	Thr	Ala	Arg	Phe	Ala	Ala	Met	Gln	Ser	
				245					250					255		
tac	atc	aag	ttc	tgc	aat	gca	tgg	atc	gat	aag	gtt	ggg	ctg	tca	gat	816
Tyr	Ile	Lys	Phe	Cys	Asn	Ala	Trp	Ile	Asp	Lys	Val	Gly	Leu	Ser	Asp	
			260					265					270			
gct	cag	ctg	acc	acg	cag	aag	gct	aat	ctg	gat	aga	acg	aag	cag	aat	864
Ala	Gln	Leu	Thr	Thr	Gln	Lys	Ala	Asn	Leu	Asp	Arg	Thr	Lys	Gln	Asn	
		275					280					285				
atg	cgt	aat	gct	atc	ctt	aat	tac	acc	cag	cag	gtt	atg	aag	gtt	ttc	912
Met	Arg	Asn	Ala	Ile	Leu	Asn	Tyr	Thr	Gln	Gln	Val	Met	Lys	Val	Phe	
	290					295					300					
aag	gat	tcc	aag	aat	atg	cca	acg	atc	ggc	acg	aat	aag	ttc	agc	gtt	960
Lys	Asp	Ser	Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	Ser	Val	
305					310					315					320	
gat	acc	tac	aat	gtt	tac	atc	aag	ggg	atg	acg	ctt	aat	gtg	ctt	gat	1008
Asp	Thr	Tyr	Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp	
				325					330					335		
atc	gtt	gct	atc	tgg	cca	agc	ctg	tac	cca	gat	gat	tac	acg	tca	cag	1056
Ile	Val	Ala	Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	Ser	Gln	
			340					345					350			
acg	gct	ctg	gaa	cag	acc	cgc	gtg	acg	ttc	tcc	aat	atg	gtg	ggg	cag	1104
Thr	Ala	Leu	Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln	
		355					360					365				
gaa	gaa	ggc	acg	gat	ggc	agc	ctc	aga	atc	tac	aat	acc	ttc	gat	agc	1152
Glu	Glu	Gly	Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser	
	370					375					380					
ttc	tcc	tac	cag	cat	agc	cct	atc	cct	aat	aat	aat	gtg	aat	ctc	atc	1200
Phe	Ser	Tyr	Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile	
385					390					395					400	
agc	tac	tac	aat	gat	gaa	ctt	cag	aat	ctg	gaa	ctc	ggg	gtt	tac	acc	1248
Ser	Tyr	Tyr	Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr	
				405					410					415		
cca	cca	aag	aag	ggc	tca	ggc	tac	agc	tac	cca	tac	ggg	ttc	gtg	ctg	1296
Pro	Pro	Lys	Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu	
			420					425					430			
aat	tac	gca	aat	agc	aag	tac	aag	tac	ggc	gat	tcc	aat	gat	cca	gaa	1344
Asn	Tyr	Ala	Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu	
		435					440					445				
tcc	ctc	ggc	ggg	ctt	tcc	acc	ctt	agc	gct	cca	atc	caa	cag	gtt	aat	1392
Ser	Leu	Gly	Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn	
	450					455					460					

gct gct acc cag aat agc aag tac ctt gat ggc gaa atc ctg aat ggg Ala Ala Thr Gln Asn Ser Lys Tyr Leu Asp Gly Glu Ile Leu Asn Gly 465 470 475 480	1440
atc ggc gca agc ctg cct ggc tac tgc acg acc ggg tgc tca cct acc Ile Gly Ala Ser Leu Pro Gly Tyr Cys Thr Thr Gly Cys Ser Pro Thr 485 490 495	1488
gaa cca cca ttc agc tgc acg agc acg gca aat ggg tac aag gca agc Glu Pro Pro Phe Ser Cys Thr Ser Thr Ala Asn Gly Tyr Lys Ala Ser 500 505 510	1536
tgc aat cca agc gat acc aat cag aag atc aat gct ctc tac cca ttc Cys Asn Pro Ser Asp Thr Asn Gln Lys Ile Asn Ala Leu Tyr Pro Phe 515 520 525	1584
acg cag gca aat gtg aag ggg aat acc ggg aag ctg ggc gtt ctc gca Thr Gln Ala Asn Val Lys Gly Asn Thr Gly Lys Leu Gly Val Leu Ala 530 535 540	1632
agc ctt gtt agc tac gat ctg aat cct aag aat gtt ttc ggg gaa ctc Ser Leu Val Ser Tyr Asp Leu Asn Pro Lys Asn Val Phe Gly Glu Leu 545 550 555 560	1680
gat tca gat acc aat aat gtt atc ctt aag ggc atc cca gca gaa aag Asp Ser Asp Thr Asn Asn Val Ile Leu Lys Gly Ile Pro Ala Glu Lys 565 570 575	1728
ggc tac ttc cca aat aat gca aga cca acc gtt gtg aag gaa tgg atc Gly Tyr Phe Pro Asn Asn Ala Arg Pro Thr Val Val Lys Glu Trp Ile 580 585 590	1776
aat ggg gct tcc gca gtt cca ctt gat tcc ggc aat acc ctg ttc atg Asn Gly Ala Ser Ala Val Pro Leu Asp Ser Gly Asn Thr Leu Phe Met 595 600 605	1824
acg gct acg aac ctg acg gct acc cag tac aga atc cgc atc cgt tac Thr Ala Thr Asn Leu Thr Ala Thr Gln Tyr Arg Ile Arg Ile Arg Tyr 610 615 620	1872
gct aac cca aac tca aac acg cag atc ggc gtt aga atc acg cag aac Ala Asn Pro Asn Ser Asn Thr Gln Ile Gly Val Arg Ile Thr Gln Asn 625 630 635 640	1920
ggg agc ctc atc tcc tcc tcc aac ctc acg ctg tac tca acc acc gat Gly Ser Leu Ile Ser Ser Ser Asn Leu Thr Leu Tyr Ser Thr Thr Asp 645 650 655	1968
atg aac aac acc ctg cca ctg aac gtt tac gtg atc ggg gaa aac ggg Met Asn Asn Thr Leu Pro Leu Asn Val Tyr Val Ile Gly Glu Asn Gly 660 665 670	2016
aac tac acc ctc cag gat ctc tac aac acg acg aac gtt ctg tcc acg Asn Tyr Thr Leu Gln Asp Leu Tyr Asn Thr Thr Asn Val Leu Ser Thr 675 680 685	2064

ggc gat atc acc ctg cag atc acc ggc ggg gat cag aag ata ttc atc Gly Asp Ile Thr Leu Gln Ile Thr Gly Gly Asp Gln Lys Ile Phe Ile 690 695 700	2112
gat cgc atc gag ttc gtg cct acg atg cca gtg cca ggc aac acc aac Asp Arg Ile Glu Phe Val Pro Thr Met Pro Val Pro Gly Asn Thr Asn 705 710 715 720	2160
aac aac aac ggg aac aac aac ggc aac aac aac cct cct cat cat gtt Asn Asn Asn Gly Asn Asn Asn Gly Asn Asn Asn Pro Pro His His Val 725 730 735	2208
tgc gct atc gca ggc acc cag cag tcc tgc tcc ggc cca cca aag ttc Cys Ala Ile Ala Gly Thr Gln Gln Ser Cys Ser Gly Pro Pro Lys Phe 740 745 750	2256
gag cag gtg tcc gat tta gaa aag ata acc acg cag gtt tac atg ctc Glu Gln Val Ser Asp Leu Glu Lys Ile Thr Thr Gln Val Tyr Met Leu 755 760 765	2304
ttc aag tcc tcc cca tac gaa gaa ctt gct ctt gaa gtt agc agc tac Phe Lys Ser Ser Pro Tyr Glu Glu Leu Ala Leu Glu Val Ser Ser Tyr 770 775 780	2352
cag atc agc cag gtt gca ctg aag gtt atg gct ctc agc gat gaa cta Gln Ile Ser Gln Val Ala Leu Lys Val Met Ala Leu Ser Asp Glu Leu 785 790 795 800	2400
ttc tgc gaa gaa aag aac gtg ctg aga aag ctc gtg aac aag gca aag Phe Cys Glu Glu Lys Asn Val Leu Arg Lys Leu Val Asn Lys Ala Lys 805 810 815	2448
cag ctt ctc gaa gca agc aac ctt ctc gtt ggg ggg aac ttc gag acg Gln Leu Leu Glu Ala Ser Asn Leu Leu Val Gly Gly Asn Phe Glu Thr 820 825 830	2496
act cag aac tgg gtg ctc ggc acg aac gct tac atc aac tat gat tca Thr Gln Asn Trp Val Leu Gly Thr Asn Ala Tyr Ile Asn Tyr Asp Ser 835 840 845	2544
ttc ctg ttc aac ggg aac tac ctg agc ctc cag cca gca tcc ggc ttc Phe Leu Phe Asn Gly Asn Tyr Leu Ser Leu Gln Pro Ala Ser Gly Phe 850 855 860	2592
ttc acg agc tac gct tac cag aag atc gat gaa tcc acg ctt aag cct Phe Thr Ser Tyr Ala Tyr Gln Lys Ile Asp Glu Ser Thr Leu Lys Pro 865 870 875 880	2640
tac acg cgc tac aag gtg tcc ggg ttc atc ggg cag tca aac cag gtt Tyr Thr Arg Tyr Lys Val Ser Gly Phe Ile Gly Gln Ser Asn Gln Val 885 890 895	2688
gaa ctg atc atc agc cgc tac ggg aag gaa atc gat aag atc ctt aac Glu Leu Ile Ile Ser Arg Tyr Gly Lys Glu Ile Asp Lys Ile Leu Asn 900 905 910	2736

gtg cca tac gca ggg cca ctc cca atc acc gca gat gca agc atc acc	2784
Val Pro Tyr Ala Gly Pro Leu Pro Ile Thr Ala Asp Ala Ser Ile Thr	
915 920 925	
tgk tgc gct cca gaa ata ggg cag tgc gat ggg gaa cag tca gat tcc	2832
Cys Cys Ala Pro Glu Ile Gly Gln Cys Asp Gly Glu Gln Ser Asp Ser	
930 935 940	
cat ttc ttc aac tac agc atc gat gtt ggc gca ctc cat cca gaa ctt	2880
His Phe Phe Asn Tyr Ser Ile Asp Val Gly Ala Leu His Pro Glu Leu	
945 950 955 960	
aac cca ggg atc gaa atc ggg ctg aag atc gtt cag agc aac ggc tac	2928
Asn Pro Gly Ile Glu Ile Gly Leu Lys Ile Val Gln Ser Asn Gly Tyr	
965 970 975	
atc acc atc tcc aac ctt gaa atc atc gaa gaa aga cca tta acg gaa	2976
Ile Thr Ile Ser Asn Leu Glu Ile Ile Glu Glu Arg Pro Leu Thr Glu	
980 985 990	
atg gaa att caa gca gtg aac cgc aag aac cag aag tgg gaa cgc gaa	3024
Met Glu Ile Gln Ala Val Asn Arg Lys Asn Gln Lys Trp Glu Arg Glu	
995 1000 1005	
aag ctg ctt gaa tgc gca tcc atc tcc gaa ctc ctc cag cct atc atc	3072
Lys Leu Leu Glu Cys Ala Ser Ile Ser Glu Leu Gln Pro Ile Ile	
1010 1015 1020	
aac cag ata gat agc ctc ttc aag gat ggc aac tgg tac aac gat atc	3120
Asn Gln Ile Asp Ser Leu Phe Lys Asp Gly Asn Trp Tyr Asn Asp Ile	
1025 1030 1035 1040	
cta cca cat gtt acc tac cag gat ctt aag aac atc atc atc cca gaa	3168
Leu Pro His Val Thr Tyr Gln Asp Leu Lys Asn Ile Ile Ile Pro Glu	
1045 1050 1055	
ctt cct aag ctc aag cat tgg ttc atc gaa aac ctg cct ggg gaa tac	3216
Leu Pro Lys Leu Lys His Trp Phe Ile Glu Asn Leu Pro Gly Glu Tyr	
1060 1065 1070	
cat gaa atc gaa cag aag atg aag gaa gca tta aag tac gca ttc act	3264
His Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr	
1075 1080 1085	
cag ctc gat gaa aag aac ctg atc cat aac ggc cat ttc acc acc aac	3312
Gln Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn	
1090 1095 1100	
ctc atc gat tgg cag gtg gaa ggc gat gct cag atg aag gtg ctt gaa	3360
Leu Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu	
1105 1110 1115 1120	
aac gat gca ctc gct ctc cag cta ttc aac tgg gat gct tcc gca tcc	3408
Asn Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser	
1125 1130 1135	

120/123

cag agc atc aac atc cta gag ttc gat gaa gat aag gct tac aag ctg	3456
Gln Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu	
1140 1145 1150	
cgt gtg tac gca cag ggc tcc ggg acg atc cag ttc ggc aac tgc gaa	3504
Arg Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu	
1155 1160 1165	
gat gaa gct atc cag ttc aac acc aac tca ttc atc tac cag gaa aag	3552
Asp Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys	
1170 1175 1180	
ata gtg tac ttc gat acg cca tcc gtt aac ctt cat atc cag agc gaa	3600
Ile Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu	
1185 1190 1195 1200	
ggc tcc gag ttc atc gtg agc agc atc gat ctc atc gaa ctc agc gat	3648
Gly Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp	
1205 1210 1215	
gat cag tag	3657
Asp Gln *	

<210> SEQ ID NO: 38

<211> Comprimento: 1218

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: APOSYNAXMI-031(FL)

<400> Sequência: 38

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

Ser	Ser	Asn	Asn	Leu	Asn	Gly	Ala	Ile	Val	Asp	Ala	Gln	Trp	Ser	Gly		
			180					185					190				
Thr	Val	Asn	Thr	Thr	Asn	Arg	Thr	Leu	Arg	Asn	Pro	Thr	Glu	Ser	Asp		
		195					200					205					
Tyr	Thr	Asn	Val	Val	Thr	Asn	Phe	Ile	Ala	Ala	Asp	Gly	Asp	Ile	Ala		
	210					215					220						
Asn	Asn	Glu	Asn	His	Ile	Met	Asn	Gly	Asn	Phe	Asp	Val	Ala	Ala	Ala		
225				230						235					240		
Pro	Tyr	Phe	Val	Ile	Gly	Ala	Thr	Ala	Arg	Phe	Ala	Ala	Met	Gln	Ser		
			245						250					255			
Tyr	Ile	Lys	Phe	Cys	Asn	Ala	Trp	Ile	Asp	Lys	Val	Gly	Leu	Ser	Asp		
	260						265						270				
Ala	Gln	Leu	Thr	Thr	Gln	Lys	Ala	Asn	Leu	Asp	Arg	Thr	Lys	Gln	Asn		
	275						280					285					
Met	Arg	Asn	Ala	Ile	Leu	Asn	Tyr	Thr	Gln	Gln	Val	Met	Lys	Val	Phe		
	290					295					300						
Lys	Asp	Ser	Lys	Asn	Met	Pro	Thr	Ile	Gly	Thr	Asn	Lys	Phe	Ser	Val		
305				310						315					320		
Asp	Thr	Tyr	Asn	Val	Tyr	Ile	Lys	Gly	Met	Thr	Leu	Asn	Val	Leu	Asp		
			325					330						335			
Ile	Val	Ala	Ile	Trp	Pro	Ser	Leu	Tyr	Pro	Asp	Asp	Tyr	Thr	Ser	Gln		
		340					345						350				
Thr	Ala	Leu	Glu	Gln	Thr	Arg	Val	Thr	Phe	Ser	Asn	Met	Val	Gly	Gln		
	355						360					365					
Glu	Glu	Gly	Thr	Asp	Gly	Ser	Leu	Arg	Ile	Tyr	Asn	Thr	Phe	Asp	Ser		
	370					375					380						
Phe	Ser	Tyr	Gln	His	Ser	Pro	Ile	Pro	Asn	Asn	Asn	Val	Asn	Leu	Ile		
385				390						395					400		
Ser	Tyr	Tyr	Asn	Asp	Glu	Leu	Gln	Asn	Leu	Glu	Leu	Gly	Val	Tyr	Thr		
			405					410						415			
Pro	Pro	Lys	Lys	Gly	Ser	Gly	Tyr	Ser	Tyr	Pro	Tyr	Gly	Phe	Val	Leu		
		420					425						430				
Asn	Tyr	Ala	Asn	Ser	Lys	Tyr	Lys	Tyr	Gly	Asp	Ser	Asn	Asp	Pro	Glu		
	435						440					445					
Ser	Leu	Gly	Gly	Leu	Ser	Thr	Leu	Ser	Ala	Pro	Ile	Gln	Gln	Val	Asn		
	450					455					460						
Ala	Ala	Thr	Gln	Asn	Ser	Lys	Tyr	Leu	Asp	Gly	Glu	Ile	Leu	Asn	Gly		
465				470						475					480		
Ile	Gly	Ala	Ser	Leu	Pro	Gly	Tyr	Cys	Thr	Thr	Gly	Cys	Ser	Pro	Thr		
			485					490						495			
Glu	Pro	Pro	Phe	Ser	Cys	Thr	Ser	Thr	Ala	Asn	Gly	Tyr	Lys	Ala	Ser		
		500					505						510				
Cys	Asn	Pro	Ser	Asp	Thr	Asn	Gln	Lys	Ile	Asn	Ala	Leu	Tyr	Pro	Phe		
	515					520						525					
Thr	Gln	Ala	Asn	Val	Lys	Gly	Asn	Thr	Gly	Lys	Leu	Gly	Val	Leu	Ala		
	530					535					540						
Ser	Leu	Val	Ser	Tyr	Asp	Leu	Asn	Pro	Lys	Asn	Val	Phe	Gly	Glu	Leu		
545				550						555					560		
Asp	Ser	Asp	Thr	Asn	Asn	Val	Ile	Leu	Lys	Gly	Ile	Pro	Ala	Glu	Lys		
			565					570						575			
Gly	Tyr	Phe	Pro	Asn	Asn	Ala	Arg	Pro	Thr	Val	Val	Lys	Glu	Trp	Ile		
		580					585						590				
Asn	Gly	Ala	Ser	Ala	Val	Pro	Leu	Asp	Ser	Gly	Asn	Thr	Leu	Phe	Met		
	595					600						605					
Thr	Ala	Thr	Asn	Leu	Thr	Ala	Thr	Gln	Tyr	Arg	Ile	Arg	Ile	Arg	Tyr		
	610					615					620						

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

123/123

His Glu Ile Glu Gln Lys Met Lys Glu Ala Leu Lys Tyr Ala Phe Thr
1075 1080 1085
Gln Leu Asp Glu Lys Asn Leu Ile His Asn Gly His Phe Thr Thr Asn
1090 1095 1100
Leu Ile Asp Trp Gln Val Glu Gly Asp Ala Gln Met Lys Val Leu Glu
1105 1110 1115 1120
Asn Asp Ala Leu Ala Leu Gln Leu Phe Asn Trp Asp Ala Ser Ala Ser
1125 1130 1135
Gln Ser Ile Asn Ile Leu Glu Phe Asp Glu Asp Lys Ala Tyr Lys Leu
1140 1145 1150
Arg Val Tyr Ala Gln Gly Ser Gly Thr Ile Gln Phe Gly Asn Cys Glu
1155 1160 1165
Asp Glu Ala Ile Gln Phe Asn Thr Asn Ser Phe Ile Tyr Gln Glu Lys
1170 1175 1180
Ile Val Tyr Phe Asp Thr Pro Ser Val Asn Leu His Ile Gln Ser Glu
1185 1190 1195 1200
Gly Ser Glu Phe Ile Val Ser Ser Ile Asp Leu Ile Glu Leu Ser Asp
1205 1210 1215
Asp Gln

REIVINDICAÇÕES

1. Molécula de ácido nucléico isolada caracterizada por compreender uma seqüência de nucleotídeos selecionada do grupo que consiste em:

a) seqüência de nucleotídeos das SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37, ou um complemento destas;

b) seqüência de nucleotídeos tendo pelo menos 90% de identidade de seqüência com a seqüência de nucleotídeos de SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37, ou um complemento destas;

c) seqüência de nucleotídeos que codifica um polipeptídeo compreendendo a seqüência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

d) seqüência de nucleotídeos que codifica um polipeptídeo compreendendo uma seqüência de aminoácidos tendo pelo menos 90% de identidade de seqüência com a seqüência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38; e,

e) seqüência de nucleotídeos da endotoxina delta do inserto de DNA do plasmídeo depositado com Nos. de Acesso B-30935, B-30936, B-30937 ou B-50046, ou um complemento destas.

2. Molécula de ácido nucléico isolada de acordo com a reivindicação 1, caracterizada pelo fato de que a seqüência

de nucleotídeos citada é uma sequência sintética que foi projetada para expressão em uma planta.

3. Vetor caracterizado por compreender a molécula de
5 ácido nucléico de acordo com a reivindicação 1.

4. Vetor de acordo com a reivindicação 3, caracterizado
por compreender adicionalmente uma molécula de ácido
nucléico codificando um polipeptídeo heterólogo.
10

5. Célula hospedeira caracterizada por conter o vetor de
acordo com a reivindicação 3.

6. Célula hospedeira de acordo com a reivindicação 5,
15 caracterizada por ser uma célula hospedeira bacteriana.

7. Célula hospedeira de acordo com a reivindicação 5,
caracterizada por ser uma célula vegetal.

20 8. Planta transgênica caracterizada por compreender a
célula hospedeira de acordo com a reivindicação 7.

9. Planta transgênica de acordo com a reivindicação 8,
caracterizada pelo fato de que a planta citada é
25 selecionada do grupo que consiste em milho, sorgo, trigo,
repolho, girassol, tomate, crucíferas, pimentas, batata,
algodão, arroz, soja, beterraba, cana-de-açúcar, tabaco,
cevada e colza.

10. Semente transgênica caracterizada por compreender a molécula de ácido nucléico de acordo com a reivindicação 1.

11. Polipeptídeo isolado com atividade pesticida
5 caracterizado por ser selecionado do grupo que consiste em:

a) polipeptídeo compreendendo a seqüência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

10 b) polipeptídeo compreendendo uma seqüência de aminoácidos tendo pelo menos 90% de identidade de seqüência com a seqüência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

15 c) polipeptídeo que é codificado por uma seqüência de nucleotídeos de SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

d) polipeptídeo que é codificado por uma seqüência de nucleotídeos que é pelo menos 90% idêntica à seqüência de nucleotídeos de SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18,
20 20, 22, 24, 26, 28, 30, 32, 34 ou 37; e,

e) polipeptídeo codificado pela seqüência de nucleotídeos da endotoxina delta do inserto de DNA do plasmídeo depositado com Nos. de Acesso B-30935, B-30936, B-30937 ou B-50046.

25

12. Polipeptídeo de acordo com a reivindicação 11, caracterizado por compreender adicionalmente seqüências de aminoácido heterólogas.

13. Anticorpo caracterizado por se ligar seletivamente ao polipeptídeo de acordo com a reivindicação 11.

14. Composição caracterizada por compreender o
5 polipeptídeo de acordo com a reivindicação 11.

15. Composição de acordo com a reivindicação 14,
caracterizada por ser selecionada do grupo que consiste em
um pó, poeira, pellet, grânulo, spray, emulsão, colóide e
10 solução.

16. Composição de acordo com a reivindicação 14,
caracterizada por ser preparada por dessecação,
liofilização, homogeneização, extração, filtração,
15 centrifugação, sedimentação ou concentração de uma cultura
de células de *Bacillus thuringiensis*.

17. Composição de acordo com a reivindicação 14,
caracterizada por compreender de cerca de 1% a cerca de 99%
20 em peso do polipeptídeo citado.

18. Método para controle de população de pragas de
lepidópteros ou coleópteros caracterizado por compreender
pôr em contato a população citada com uma quantidade que
25 tenha efeito de pesticida do polipeptídeo de acordo com a
reivindicação 11.

19. Método para exterminar uma praga de lepidópteros ou
coleópteros caracterizado por compreender pôr em contato a
30 praga citada com, ou alimentar a praga citada com, uma

quantidade que tenha efeito de pesticida do polipeptídeo de acordo com a reivindicação 11.

20. Método para produzir um polipeptídeo com atividade
5 pesticida caracterizado por compreender cultivar a célula hospedeira de acordo com a reivindicação 4 sob condições nas quais a molécula de ácido nucléico codificando o polipeptídeo é expressa.

10 21. Planta caracterizada por ter incorporado de maneira estável no seu genoma uma construção de DNA compreendendo uma sequência de nucleotídeos que codifica uma proteína tendo atividade pesticida, em que a sequência de nucleotídeo citada é selecionada do grupo que consiste em:

15

a) sequência de nucleotídeos das SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

b) sequência de nucleotídeos tendo pelo menos 90% de identidade de sequência com a sequência de nucleotídeos de
20 SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

c) sequência de nucleotídeos que codifica um polipeptídeo compreendendo a sequência de aminoácidos de
SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27,
25 29, 31, 33, 35 ou 38;

d) sequência de nucleotídeos que codifica um polipeptídeo tendo pelo menos 90% de identidade de sequência de aminoácidos com a sequência de aminoácidos de
SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27,
30 29, 31, 33, 35 ou 38; e,

e) sequência de nucleotídeos da endotoxina delta do inserto de DNA do plasmídeo depositado com Nos. de Acesso B-30935, B-30936, B-30937, B-50046, ou um complemento destas;

5 em que a sequência de nucleotídeos citada está operacionalmente ligada a um promotor que dirige a expressão de uma sequência codificante em uma célula vegetal.

10 22. Planta de acordo com a reivindicação 21, caracterizada por ser uma célula vegetal.

23. Método para proteger uma planta de uma praga caracterizado por introduzir na planta citada ou célula
15 desta pelo menos um vetor de expressão compreendendo uma sequência de nucleotídeos que codifique um polipeptídeo pesticida, em que a sequência de nucleotídeos é selecionada do grupo que consiste em:

20 a) sequência de nucleotídeos das SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

b) sequência de nucleotídeos tendo pelo menos 90% de identidade de sequência com a sequência de nucleotídeos de
25 SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

c) sequência de nucleotídeos que codifica um polipeptídeo compreendendo a sequência de aminoácidos de
SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

d) sequência de nucleotídeos que codifica um polipeptídeo tendo pelo menos 90% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38; e,

e) sequência de nucleotídeos da endotoxina delta do inserto de DNA do plasmídeo depositado com Nos. de Acesso B-30935, B-30936, B-30937 ou B-50046, ou um complemento destas.

24. Método de acordo com a reivindicação 23, caracterizado pelo fato de que a planta citada produz um polipeptídeo pesticida tendo atividade pesticida contra uma praga de lepidópteros ou coleópteros.

25. Método para proteger uma planta de uma praga de nematóides caracterizado por compreender:

a) introduzir na planta citada ou célula desta pelo menos um vetor de expressão compreendendo uma sequência de nucleotídeos que codifique um polipeptídeo pesticida, em que o polipeptídeo pesticida citado possui um tamanho molecular maior que 22 kDa; e,

b) cultivar a planta citada em um campo contendo nematóides de cisto.

26. Método de acordo com a reivindicação 25, caracterizado pelo fato de que o nematóide de cisto citado é *Heterodera sp.*

27. Método de acordo com a reivindicação 26, caracterizado pelo fato de que o nematóide de cisto citado é *Heterodera schactii*.

5 28. Método de acordo com a reivindicação 25, caracterizado pelo fato de que o polipeptídeo pesticida citado é maior que cerca de 40kDa.

10 29. Método de acordo com a reivindicação 28, caracterizado pelo fato de que o polipeptídeo pesticida citado é maior que cerca de 70kDa.

15 30. Método de acordo com a reivindicação 29, caracterizado pelo fato de que o polipeptídeo citado é selecionado do grupo que consiste em:

a) polipeptídeo compreendendo a sequência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

20 b) polipeptídeo compreendendo uma sequência de aminoácidos tendo pelo menos 90% de identidade de sequência com a sequência de aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38;

25 c) polipeptídeo que é codificado por uma sequência de nucleotídeos de SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37;

30 d) polipeptídeo que é codificado por uma sequência de nucleotídeos que é pelo menos 90% idêntica à sequência de nucleotídeos de SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37; e,

e) polipeptídeo codificado pela sequência de nucleotídeos da endotoxina delta do inserto de DNA do plasmídeo depositado com Nos. de Acesso B-30935, B-30936, B-30937 ou B-50046, ou um complemento destas.

5

31. Método de acordo com a reivindicação 25, caracterizado pelo fato de que a planta citada possui uma produção aumentada comparada a uma planta que não expressa a sequência de nucleotídeos codificando uma proteína

10 pesticida.

RESUMO

MOLÉCULA DE ÁCIDO NUCLÉICO ISOLADA, VETOR, CÉLULA HOSPEDEIRA, PLANTA E SEMENTE TRANSGÊNICAS, POLIPEPTÍDEO ISOLADO COM ATIVIDADE PESTICIDA E MÉTODO PARA SUA PRODUÇÃO, ANTICORPO, COMPOSIÇÃO E MÉTODOS PARA CONTROLE DE POPULAÇÃO DE PRAGAS, PARA EXTERMINAR UMA PRAGA DE LEPIDÓPTEROS OU COLEÓPTEROS, PARA PROTEGER UMA PLANTA DE UMA PRAGA E DE UMA PRAGA DE NEMATÓIDES

São fornecidos composições e métodos para conferir atividade pesticida a bactérias, plantas, células vegetais, tecidos e sementes. São fornecidas composições compreendendo uma sequência codificante de um polipeptídeo de delta-endotoxina. As sequências codificantes podem ser usadas em construções de DNA ou cassetes de expressão para transformação e expressão em plantas e bactérias. Composições também compreendem bactérias, plantas, células vegetais, tecidos e sementes transformadas. Em particular, são fornecidas moléculas de ácido nucléico de delta-endotoxinas isoladas. Adicionalmente, são englobadas sequências de aminoácidos correspondentes aos polinucleotídeos e anticorpos que se ligam especificamente a estas sequências de aminoácidos. Em particular, a presente invenção proporciona moléculas de ácido nucléico compreendendo as sequências de nucleotídeos codificantes das sequências de aminoácidos mostradas em SEQ ID NO:2, 4, 6, 8, 10, 12, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35 ou 38, ou a sequência de nucleotídeos apresentada em SEQ ID NO:1, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ou 37, bem como variantes e fragmentos destas.