



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

(21) **PI 0711953-4 A2**

(22) Data de Depósito: 08/06/2007
(43) Data da Publicação: 17/01/2012
(RPI 2141)



(51) *Int.Cl.:*
A01H 5/00
C12N 9/00
C12N 15/52
C12N 15/82

(54) Título: POLINUCLEOTÍDEO ISOLADO E POLIPEPTÍDEO CONFERINDO RESISTÊNCIA A GLUTAMINA SINTETASE E MÉTODOS PARA PRODUZIR PLANTAS E CÉLULAS VEGETAIS TRANSGÊNICAS APRESENTANDO PRODUÇÃO E UTILIZAÇÃO DE NITROGÊNIO MELHORADAS

(30) Prioridade Unionista: 08/06/2006 US 60/812,000

(73) Titular(es): Athenix Corporation

(72) Inventor(es): Daniel John Tomso, Laura Cooper Schouten, Nicholas B. Duck, Todd K. Hinson, Vadim Beilinson

(74) Procurador(es): Claudia Christina Schulz

(86) Pedido Internacional: PCT US2007070690 de 08/06/2007

(87) Publicação Internacional: WO 2007/146767 de 21/12/2007

(57) Resumo: POLINUCLEOTÍDEO ISOLADO E POLIPEPTÍDEO CONFERINDO RESISTÊNCIA A GLUTAMINA SINTETASE E MÉTODOS PARA PRODUZIR PLANTAS E CÉLULAS VEGETAIS TRANSGÊNICAS APRESENTANDO PRODUÇÃO E UTILIZAÇÃO DE NITROGÊNIO MELHORADAS. São providos composições e métodos para conferir resistência a herbicidas e melhorar a utilização de nitrogênio de bactérias, plantas, células vegetais, tecidos e sementes. Composições compreendendo uma seqüência codificadora para um polipeptídeo que confere resistência ou tolerância a inibidores de glutamina sintetase herbicidas são providas. As seqüências codificadoras podem ser usadas em construções de DNA ou cassetes de expressão para transformação e expressão em plantas. Composições também compreendem bactérias, plantas, células, tecidos e sementes vegetais transformadas. Em particular, polinucleotídeos isolados correspondendo a polinucleotídeos resistentes a inibidor de glutamina sintetase herbicida são providos. Adicionalmente, polipeptídeos correspondendo aos polinucleotídeos são englobados. Em particular, a presente invenção provê polinucleotídeos isolados compreendendo um variante de SEQ ID NO: 1, caracterizado pelo fato do polinucleotídeo variante codificar um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida.



PI0711953-4

**POLINUCLEOTÍDEO ISOLADO E POLIPEPTÍDEO CONFERINDO
RESISTÊNCIA A GLUTAMINA SINTETASE E MÉTODOS PARA PRODUZIR
PLANTAS E CÉLULAS VEGETAIS TRANSGÊNICAS APRESENTANDO
PRODUÇÃO E UTILIZAÇÃO DE NITROGÊNIO MELHORADAS**

5

CAMPO DA INVENÇÃO

Esta invenção está relacionada a biologia molecular vegetal, particularmente a uma nova classe de enzimas glutamina sintetases que conferem resistência a inibidores de glutamina sintetase herbicidas.

FUNDAMENTOS DA INVENÇÃO

Plantas obtêm nitrogênio de seu ambiente na forma de compostos inorgânicos, conhecidos por nitrato e amônia retirado pelas raízes, e N_2 atmosférico reduzido a amônia em nódulos de raiz fixadores de nitrogênio. A primeira etapa na assimilação de nitrogênio inorgânico na forma orgânica envolve predominantemente a incorporação de amônia com glutamato para formar glutamina, catalisada pela enzima glutamina sintetase.

Diversos herbicidas funcionam inibindo glutamina sintetase vegetal. Um exemplo típico de tal composto é o análogo de ácido glutâmico, glufosinato (ou fosfinotricina). Vários desses herbicidas inibem a glutamina sintetase presente nas plantas de cultivo, assim como em ervas daninhas, limitando, dessa forma, o uso de

tais compostos como glufosinato. Como a seletividade de herbicida é importante em qualquer herbicida comercialmente útil, seria de grande interesse ser capaz de conferir resistência em plantas selecionadas para tais herbicidas não seletivos como glufosinato, assim como a outros inibidores de glutamina sintetase herbicidas.

Enzimas que são resistentes a inibidores de glutamina sintetase herbicidas são conhecidas na técnica. Metiona sulfoximina (MSO), um análogo de glutamato, é um inibidor competitivo misturado de glutamina sintetase de folha de ervilha (Leason et al. (1982) *Phytochemistry* 21:855). Células de alfafa resistentes a fosfinotricina foram relatadas (Newmark (1983) *Nature* 305:383-384). A resistência foi devido à amplificação do gene de glutamina sintetase (Donn et al. (1984) *Journal of Molecular and Applied Genetics* 2: 621-635). O gene *Bar*, isolado de *Streptomyces hygroscopicus*, codifica a enzima fosfinotricina N-acetiltransferase (PAT). Este gene pode conferir resistência a herbicidas a base de glufosinato, já que PAT desintoxifica fosfinotricina por acetilação, o que produz um composto inativo.

Genes adicionais que são resistentes a inibidores de glutamina sintetase herbicidas são necessários onde a resistência é devido a uma mutação funcional na enzima glutamina sintetase, ao invés de uma amplificação ou inativação por acetilação da enzima.

SUMÁRIO DA INVENÇÃO

Composições e métodos para conferir resistência a inibidores de glutamina sintetase herbicidas em plantas, células, tecidos e sementes vegetais são providos. Em um avanço, os polinucleotídeos empregados nos vários métodos e composições conferem resistência a glufosinato. Composições incluem polinucleotídeos codificando polipeptídeos resistentes a inibidores de glutamina sintetase herbicidas, vetores compreendendo aqueles polinucleotídeos, e células hospedeiras compreendendo os vetores. Composições compreendendo uma seqüência codificadora para um polipeptídeo que confere resistência ou tolerância a inibidores de glutamina sintetase herbicidas são providas. Composições compreendendo uma seqüência codificadora para um polipeptídeo que resulta em utilização melhorada de nitrogênio e/ou produção intensificada em uma planta são ainda providas. As seqüências codificadoras podem ser usadas em construções de DNA ou cassetes de expressão para transformação e expressão em organismos, incluindo microorganismos e plantas. As composições também compreendem bactérias, plantas, células, tecidos e sementes vegetais transformadas.

A presente invenção provê polinucleotídeos isolados compreendendo SEQ ID NO: 1, assim como variantes da seqüência de polinucleotídeos apresentada em SEQ ID NO:1, incluindo SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, ou a

seqüência de nucleotídeos de glutamina sintetase em um hospedeiro bacteriano, como o No. de Acesso NRRL B-30930, depositado em 8 de junho de 2006 na Autoridade Depositária Intenacional da Coleção de Culturas de Pesquisas Agrícolas (Agricultural Research Culture Collection) (NRRL) International Depository Authority, 1815 N. University Street, Peoria, Illinois 61604 USA, e os polipeptídeos correspondentes àqueles polinucleotídeos. Em um avanço, Os polinucleotídeos da presente invenção compreendem pelo menos uma modificação entre os aminoácidos 125 a 175 ou entre os aminoácidos 200 a 250 correspondentes à SEQ ID NO:2, ou pelo menos uma modificação que resulte na perda de um sítio de adenilação.

15 DESCRIÇÃO DAS FIGURAS

As Figuras 1A- 1C mostram um alinhamento da seqüência de nucleotídeos dos variantes resistentes a herbicidas de glutamina sintetase, incluindo pAX3421m1 (SEQ ID NO:4), pAX3422m2 (SEQ ID NO:6), pAX3427m3 (SEQ ID NO:8), pAX3428m4 (SEQ ID NO: 10), pAX3430m6 (SEQ ID NO: 12), pAX3431m7 (SEQ ID NO: 14), pAX3432m8 (SEQ ID NO: 16), pAX3433m9 (SEQ ID NO: 18), pAX3434m10 (SEQ ID NO:20), pAX3435m11 (SEQ ID NO:22), pAX3436m12 (SEQ ID NO:24), pAX3437m13 (SEQ ID NO:26), pAX3438m14 (SEQ ID NO:28), pAX3426m15 (SEQ ID NO:30) e pAX3439m16 (SEQ ID NO:32) com a seqüência de aminoácidos *ags1* de tipo selvagem (SEQ ID NO:2).

DESCRIÇÃO DETALHADA DA INVENÇÃO

I. Composições

Composições e métodos para conferir resistência ou
5 tolerância a herbicidas, particularmente resistência ou
tolerância a inibidores de glutamina sintetase herbicidas,
em organismos são providos. Os métodos envolvem transformar
organismos com polinucleotídeos codificando um gene de
tolerância a herbicidas que codifica um polipeptídeo que é
10 resistente aos inibidores de glutamina sintetase
herbicidas. Em um avanço, os polinucleotídeos codificam um
gene de tolerância a herbicidas que codifica um
polipeptídeo que é resistente a inibição por glufosinato.
Por gene de "resistência a herbicidas" ou "tolerância a
15 herbicidas" da invenção pretende-se a seqüência de
nucleotídeos apresentada em SEQ ID NO: 3, 5, 7, 9, 11, 13,
15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43,
45, ou 47 e fragmentos e variantes desses que codificam um
polipeptídeo de resistência ou tolerância ao inibidor de
20 glutamina sintetase (inibidor-GS). Da mesma forma, um
polipeptídeo de "resistência a herbicidas" ou "tolerância a
herbicidas" da invenção é um polipeptídeo tendo a seqüência
de aminoácidos apresentada em SEQ ID NO: 4, 6, 8, 10, 12,
14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42,
25 44, 46, e fragmentos e variantes desses, que conferem
resistência ou tolerância ao inibidor-GS em uma célula
hospedeira.

A presente invenção provê polinucleotídeos isolados compreendendo variantes da seqüência de polinucleotídeos apresentada em SEQ ID NO:1, caracterizado pelo fato de que os variantes codificam polipeptídeos que são resistentes a inibidores de glutamina sintetase herbicidas. Em um avanço, os polinucleotídeos da presente invenção codificam polipeptídeos que compreendem pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 a 250 correspondentes a SEQ ID NO:2. Para os propósitos da presente invenção, "modificação" deve significar uma mudança na seqüência de nucleotídeos que resulte em uma mudança no polipeptídeo codificado. Uma modificação pode englobar também uma substituição de um aminoácido por outro aminoácido em uma seqüência polipeptídica. Por "correspondente a" pretende-se que as posições de aminoácidos mencionadas estejam relacionadas com as posições de aminoácidos designadas em SEQ ID NO:2, e que as substituições correspondentes a essas posições de aminoácidos podem ser encontradas em seqüências variantes quando essas seqüências variantes estão alinhadas com SEQ ID NO:2 usando métodos de alinhamento padrão.

Um plasmídeo contendo a seqüência de nucleotídeos de resistência a herbicidas da invenção foi depositado na coleção permanente da Coleção de Cultura do Serviço de Pesquisa Agrícola (*Agricultural Research Service Culture Collection*), Northern Regional Research Laboratory (NRRL), 1815 North University Street, Peoria, Illinois 61604, Estados Unidos da America, em 8 de junho de 2006, e

designado com o No. de Acesso B-30930. Este depósito será mantido sob os termos do Tratado de Budapeste sobre o Reconhecimento Internacional do Depósito de Microorganismos para os Propósitos de Procedimentos de Patentes
5 (*International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure*). Este depósito foi feito meramente como conveniência para aqueles especialistas na técnica e não é uma admissão de que um depósito é necessário sob 35 U.S.C. §112.

10

A definição da enzima "glutamina sintetase," ou "glutamina sintase" é funcional e inclui qualquer glutamina sintetase capaz de funcionar em um dado hospedeiro desejado, especialmente uma bactéria ou planta, para
15 converter ácido glutâmico para glutamina. O termo, portanto, inclui não apenas a enzima da espécie de planta em particular envolvida na transformação, mas pode incluir glutamina sintetase de outra espécie vegetal ou microrganismos, se tal glutamina sintetase é capaz de
20 funcionar na planta transformada ou células bacterianas.

O termo "inibidor de glutamina sintetase herbicida" ou "inibidor de glutamina sintase herbicida" deve incluir qualquer inibidor, competitivo ou não competitivo, que
25 diminua significativamente a atividade de glutamina sintetase de uma célula vegetal de uma determinada espécie e, como consequência disso, cause efeitos herbicidas na célula vegetal. Exemplos de inibidores de glutamina

sintetase são glufosinato, fosfinotricina, metionina sulfoximina, assim como outros análogos de ácido glutâmico.

5 O termo "glufosinato" denota o composto conhecido, em sua forma biologicamente ativa. Pode estar presente em qualquer forma enantiomérica, e pode estar sozinho ou em combinação com outros compostos inertes ou ativos que não interferem com atividade de glufosinato.

10 A. *Glutamina sintetase*

Na presente invenção, a classe de enzimas que confere resistência a herbicidas é glutamina sintetase (GS). O termo "glutamina sintetase" ou "glutamina sintase" ou "GS",
15 como aqui usado, refere-se tanto a glutamina sintetase nativa ou a um variante ou fragmento dessa.

Glutamina sintetase é uma enzima chave no metabolismo de nitrogênio; possui funções duplas em duas reações
20 bioquímicas essenciais, assimilação de amônia e biossíntese de glutamina. A glutamina produzida por GS é essencial para a síntese de proteínas, e seu nitrogênio da amida é doado para sintetizar vários metabólitos essenciais.

25 A forma comum de GS é uma enzima duodecamérica com subunidades idênticas de aproximadamente 55 kDa, codificada por *glnA*. A estrutura cristal desta enzima revelou que é composta de 12 subunidades idênticas arranjadas como dois anéis hexagonais superpostos que são mantidos juntos tanto

por interações hidrofóbicas e pontes de hidrogênio entre as subunidades. (Yamashita *et al.* (1989) *J. Biol. Chem.* 264: 17681-17690). Glutamina sintetase catalisa a formação de glutamina a partir de glutamato e amônia em uma reação dependente de ATP. Ela também catalisa a transferência de gama-glutamil de glutamina para hidroxilamina produzindo gama-glutamilhdroximato (Stadtman *et al.* (1974) *in The Enzymes* (Boyer, ed.) 3:755-807 (Academic Press, New York). A catálise de glutamina sintetase envolve a formação inicial de um intermediário gama-glutamil fosfato, seguido pelo deslocamento do grupamento fosfato ativado por amônio através da formação de um intermediário fosforilado tetraédrico. Em *E. coli*, os resíduos de Asp50 e Glu327 altamente conservados formam um bolso de ligação negativamente carregado que constitui o sítio de ligação de amônia (Liaw *et al.* (1995) *Protein Sci.* 4:2358-2365).

Numerosos inibidores potentes que imitam a geometria do intermediário tetraédrico são conhecidos para glutamina sintetase, incluindo glufosinato (ou fosfinotricina (PPT)) a L-metionina-DL-sulfoximida (MSX). A fosforilação de MSX e glufosinato é similar à fosforilação de glutamato na primeira etapa da reação enzimática normal de glutamina sintetase, e é necessária para que a inibição irreversível ocorra (Crespo *et al.* (1999) *Eur. J. Biochem.* 266:1202-1209). Em plantas, a inibição de glutamina sintetase resulta em uma preparação de amônia fitotóxica e uma falta de aminoácidos essenciais, e uma inibição de fotorrespiração e fotossíntese e, por fim, morte da planta.

Glufosinato é um composto natural isolado a partir de duas espécies de fungos Streptomicetos que inibe a atividade de glutamina sintetase. A aplicação de glufosinato resulta em níveis de glutamina reduzidos e um aumento correspondente em concentrações de amônia em tecidos vegetais, levando ao rompimento da membrana celular e interrupção de fotossíntese, resultando no ressecamento e morte da planta. Um número de análogos de glufosinato que inibem a glutamina sintetase vegetal é conhecido na técnica. Ver, por exemplo, Berlicki et al. (2005) *J. Med. Chem.* 48(20): 6340-6349 e Forlani et al. (2006) *J. Agric. Food Chem.* 54(3): 796-802; cada qual é aqui incorporada por referência em sua íntegra.

15

B. Glutamina sintetase resistente a herbicidas

Resistência a L-fosfinotricina foi relatada em células de alfafa, após uma seleção passo a passo em níveis crescentes de L-PPT, resultando na amplificação gênica (Donn et al. (1984) *J. Mol. Appl. Genet.* 2:621), e pela introdução em plantas de tabaco, batata e tomate, via transformação mediada por *Agrobacterium* do gene *Bar*, que codifica fosfinotricina acetiltransferase (PAT), uma enzima desintoxicadora (De Block et al. (1987) *EMBOJ.* 6:2513). Mutantes de enzimas glutamina sintetase vegetais que são resistentes a fosfinotricina são descritos em Patente U.S. No. 5.145.777. Esses mutantes conferem resistência pela super-expressão de glutamina sintetase.

C. Atividade de glutamina sintetase

5 Uma variedade de métodos pode ser usada para medir a
atividade de glutamina sintetase. Ver, por exemplo, Crespo
et al. (1999) *Eur. J. Biochem.* 266:1202-1209, Gawronski *et*
al. (2004) *Anal. Biochem.* 327:114-118, e Patente U.S. Nos.
5.098.838 e 5.145.777, cada qual é aqui incorporada por
10 referência em sua íntegra. A atividade pode ser medida
usando polipeptídeos glutamina sintetase purificados, ou
testando a habilidade de organismos transformados com os
polinucleotídeos da invenção para crescer na presença de
inibidores de glutamina sintetase herbicidas.

15

*D. Polinucleotídeos isolados, e variantes e fragmentos
desses*

Em alguns avanços, a presente invenção compreende
20 polinucleotídeos isolados ou recombinantes codificando
polipeptídeos que são resistentes a inibidores de glutamina
sintetase herbicidas. Um polinucleotídeo ou polipeptídeo
"isolado" ou "purificado", ou porção biologicamente ativa,
é substancialmente livre de outro material celular, ou meio
de cultura quando produzido por técnicas recombinantes, ou
25 substancialmente livre de precursores químicos ou outros
químicos quando quimicamente sintetizado. Por
"biologicamente ativo" pretende-se possuir a atividade
biológica desejada do polipeptídeo nativo, isto é,

resistência a inibidores de glutamina sintetase herbicidas. Um polinucleotídeo "isolado" pode ser livre de seqüências (por exemplo, seqüências codificando proteínas) que naturalmente flanqueiam o ácido nucléico (i.e., seqüências localizadas nas extremidades 5' e 3' do ácido nucléico) no DNA genômico do organismo do qual o polinucleotídeo é derivado. Pra propósitos da invenção, "isolado" quando usado para referir-se a polinucleotídeos exclui cromossomos isolados. Por exemplo, em vários avanços, o polinucleotídeo isolado codificando resistência a glifosato pode conter menos de cerca de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb, ou 0,1 kb de seqüência de nucleotídeos que naturalmente flanqueiam o polinucleotídeo no DNA genômico da célula da qual o polinucleotídeo é derivado.

15

A presente invenção contempla ainda variantes e fragmentos dos polinucleotídeos aqui descritos. Um "fragmento" de um polinucleotídeo pode codificar uma porção biologicamente ativa de um polipeptídeo, ou pode ser um fragmento que pode ser usado como uma sonda de hibridização ou *primer* de PCR usando métodos aqui descritos em outra parte. Polinucleotídeos que são fragmentos de um polinucleotídeo compreendem pelo menos cerca de 15, 20, 50, 75, 100, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400 nucleotídeos contíguos, ou até o número de nucleotídeos presentes em polinucleotídeo de comprimento total aqui descrito. Por nucleotídeos

25

"contíguos" entende-se que são resíduos de nucleotídeos que são imediatamente adjacentes um ao outro.

Fragmentos dos polinucleotídeos da presente invenção
5 geralmente codificarão fragmentos de polipeptídeos que
retêm a atividade biológica da proteína de comprimento
total de resistência a herbicidas; i.e., resistência a
inibidores de glutamina sintetase herbicidas. Por "retém a
atividade de resistência a herbicidas" pretende-se que o
10 fragmento tenha pelo menos cerca de 30%, pelo menos cerca
de 50%, pelo menos cerca de 70%, ou pelo menos cerca de 80%
da atividade de resistência a herbicidas da proteína de
comprimento total de resistência a herbicidas aqui descrita
como SEQ ID NO:6. Métodos para medir a atividade de
15 resistência a herbicidas são bem conhecidos na técnica.
Ver, por exemplo, Patente U.S. Nos. 5.098.838 e 5.145.777,
cada qual é aqui incorporada por referência em sua íntegra.
Atividade pode também referir-se à atividade enzimática da
enzima glutamina sintetase como aqui descrito em outra
20 parte.

Um fragmento de um polinucleotídeo que codifica uma
porção biologicamente ativa de um polipeptídeo da invenção
codificará pelo menos cerca de 15, 25, 30, 50, 75, 100,
25 125, 150, 175, 200, 250, 300, 350, 400, 450 aminoácidos
contíguos, ou até o número total de aminoácidos presentes
em um polipeptídeo de comprimento total da invenção.

A invenção também engloba polinucleotídeos variantes. "Variantes" do polinucleotídeo incluem aquelas seqüências que codificam os polipeptídeos aqui descritos, mas que diferem conservativamente por causa da degeneração do código genético, assim como aqueles que são suficientemente idênticos. O termo "suficientemente idêntico" deve significar um polipeptídeo ou seqüência de polinucleotídeos que possui pelo menos cerca de 60% ou 65% de identidade de seqüência, cerca de 70% ou 75% de identidade de seqüência, cerca de 80% ou 85% de identidade de seqüência, cerca de 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99% de identidade de seqüência comparado a uma seqüência referência usando um dos programas de alinhamento usando parâmetros padrão. Um especialista na técnica reconhecerá que esses valores podem ser apropriadamente ajustados para determinar a identidade correspondente de polipeptídeos codificada por dois polinucleotídeos levando-se em consideração a degeneração de códons, similaridade de aminoácidos, posicionamento da fase de leitura, e similares.

Para determinar o percentual de identidade de duas seqüências de aminoácidos ou de dois polinucleotídeos, as seqüências são alinhadas para propósitos de comparação ótima. O percentual de identidade entre as duas seqüências é uma função do número de posições idênticas compartilhadas pelas seqüências (i.e., percentual de identidade = número de posições idênticas/número total de posições (e.g., posições se sobrepondo) x 100). Em um avanço, as duas

seqüências são do mesmo comprimento. O percentual de identidade entre duas seqüências pode ser determinado usando técnicas similares àquelas descritas abaixo, permitindo ou não espaços vazios. Quando calcular o percentual de identidade, são contadas tipicamente as correspondências exatas.

A determinação do percentual de identidade entre duas seqüências pode ser conseguida usando um algoritmo matemático. Um exemplo não limitante de um algoritmo matemático utilizado para a comparação de duas seqüências é o algoritmo de Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264, modificado como em Karlin & Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Tal algoritmo é incorporado nos programas BLASTN e BLASTX de Altschul et al. (1990) *J. Mol. Biol.* 215:403. Buscas de nucleotídeos BLAST podem ser realizadas com o programa BLASTN, pontuação = 100, tamanho de palavra = 12, para se obter polinucleotídeos homólogos a polinucleotídeos codificando resistência a herbicidas usados em métodos da invenção. Buscas BLAST de polipeptídeos podem ser efetuadas com o programa BLASTX, pontuação = 50, tamanho de palavra = 3, para se obter seqüências de aminoácidos homólogas a moléculas polipeptídicas expressas usando os métodos da invenção. Para se obter alinhamentos com espaços para fins de comparação, *Gapped BLAST* pode ser utilizado como descrito em Altschul et al. (1997) *Nucleic Acids Res.* 25:3389. Alternativamente, PSI-Blast pode ser usado para efetuar uma busca iterada que detecte relações distantes

entre moléculas. Ver Altschul et al. (1997) *supra*. Quando os programas utilizando BLAST, Gapped BLAST, e PSI-Blast, os parâmetros padrão dos respectivos programas (e.g., BLASTX e BLASTN) podem ser usados. Ver

5 www.ncbi.nlm.nih.gov. Outro exemplo não limitante de um algoritmo matemático utilizado para a comparação de seqüências é o algoritmo de ClustalW (Higgins et al. (1994) *Nucleic Acids Res.* 22:4673-4680). ClustalW compara seqüências e alinha a totalidade da seqüência de

10 aminoácidos ou de DNA e, assim, pode prover dados acerca da conservação de seqüência de toda a seqüência de aminoácidos. O algoritmo de ClustalW é usado em diversos pacotes de *software* de análise de DNA/aminoácido comercialmente disponíveis, tal como o módulo ALIGNX de

15 *Vector NTI Program Suite* (Invitrogen Corporation, Carlsbad, CA). Após o alinhamento de seqüências de aminoácidos com ClustalW, a percentual de identidade de aminoácidos pode ser avaliada. Um exemplo não limitante de um *software* útil para análise em alinhamentos de ClustalW é GENEDOC™.

20 GENEDOC™ (Karl Nicholas) permite a obtenção da similaridade e identidade de aminoácidos (ou DNA) entre múltiplos polipeptídeos. Outro exemplo não limitante de um algoritmo matemático utilizado para a comparação de seqüências é o algoritmo de Myers and Miller (1988) *CABIOS* 4: 11-17. Tal

25 algoritmo é incorporado no programa ALIGN (versão 2.0), que é parte do pacote de *software* de alinhamento de seqüências GCG. Quando utilizando o programa ALIGN para comparar seqüências de aminoácidos, uma tabela de peso de resíduos

PAM 120, uma penalidade por comprimento de espaço de 12, e uma penalidade por espaço de 4 podem ser usadas.

A menos que declarado ao contrário, GAP Versão 10, que
5 usa o algoritmo de Needleman and Wunsch (1970) *J. Mol. Biol.* 48(3):443-453 é usado para determinar a identidade ou similaridade de seqüências usando os seguintes parâmetros:
% identidade e % similaridade para uma seqüência de nucleotídeos usando peso GAP (*GAP Weight*) de 50 e peso de
10 comprimento (*Length Weight*) de 3, e a matriz de pontuação de nwsgapdna.cmp; % identidade ou % similaridade para uma seqüência de aminoácidos usando peso GAP de 8 e peso de comprimento de 2, e o programa de pontuação BLOSUM62. Programas equivalentes também podem ser usados. Por
15 "programa equivalente" pretende-se qualquer programa de comparação de seqüências que, para quaisquer duas seqüências em questão, gere um alinhamento tendo correspondências de resíduos de nucleotídeos idênticas e identidade de seqüência percentual idêntica quando
20 comparado ao alinhamento correspondente gerado pelo programa GAP Version 10.

Variantes alélicas naturalmente ocorrentes podem ser identificados com o uso de técnicas de biologia molecular
25 bem conhecidas, tal como reação de cadeia de polimerase (PCR) e técnicas de hibridização, como destacado abaixo. Polinucleotídeos variantes também incluem polinucleotídeos derivados sinteticamente que foram gerados, por exemplo, usando-se mutagênese sítio direcionada, mas que ainda

codificam o polipeptídeo tendo a atividade biológica desejada.

O especialista na técnica irá ainda ponderar que
5 mudanças podem ser introduzidas por mutação nos
polinucleotídeos da invenção, levando, dessa forma, a
mudanças na seqüência de aminoácidos dos polipeptídeos
codificados, sem alterar a atividade biológica dos
polipeptídeos. Assim, polinucleotídeos variantes isolados
10 podem ser criados introduzindo um ou mais substituições,
adições ou deleções de nucleotídeos no polinucleotídeo
correspondente aqui descrito, de forma que um ou mais
substituições, adições ou deleções de aminoácido sejam
introduzidas no polipeptídeo codificado. Mutações podem ser
15 introduzidas por técnicas padrão, tal como mutagênese sítio
direcionada e mutagênese mediada por PCR, ou técnicas de
embaralhamento de genes. Tais polinucleotídeos variantes
são também englobados pela presente invenção.

20 Polinucleotídeos variantes podem ser feitos
introduzindo mutações randomicamente ao longo de toda ou
parte da seqüência codificadora, tal como por mutagênese de
saturação, e os mutantes resultantes podem ser rastreados
para a habilidade de conferir atividade de resistência a
25 herbicidas para identificar mutantes que retêm atividade.
Após a mutagênese, o polipeptídeo codificado pode ser
expresso de forma recombinante, e a atividade do
polipeptídeo pode ser determinada usando técnicas de teste
padrão.

Procedimentos de embaralhamento de genes ou PCR
sexuado (por exemplo, Smith (1994) *Nature* 370:324-325;
Patente U.S. Nos. 5.837.458; 5.830.721; 5.811.238; e
5 5.733.731. cada qual é aqui incorporada por referência)
podem ser usados para se identificar polinucleotídeos
adicionais que codificam polipeptídeos que efetuam funções
similares como aqueles aqui descritos (por exemplo,
polipeptídeos que são resistentes a inibidores de glutamina
10 sintetase herbicidas). O embaralhamento de genes envolve a
fragmentação aleatória de diversos DNAs mutantes, seguido
por seu reagrupamento por PCR em moléculas de comprimento
total. Exemplos de vários procedimentos de embaralhamento
de genes incluem, mas não são limitados a, reunião seguido
15 por tratamento com DNase, o processo de extensão alternado
(STEP), e recombinação *in vitro* por iniciação aleatória. No
método mediado por DNase, segmentos de DNA isolados a
partir de um grupo de mutantes positivos são clivados em
fragmentos randômicos com DNaseI e sujeitos a múltiplas
20 rodadas de PCR com nenhum *primer* adicionado. Os
comprimentos de fragmentos randômicos se aproximam aquele
do segmento não clivado à medida que os ciclos de PCR
prosseguem, resultando em mutações em diferentes clones se
tornando misturados e acumulando em algumas das seqüências
25 resultantes. Múltiplos ciclos de seleção e embaralhamento
levaram à intensificação funcional de diversas enzimas
(Stemmer (1994) *Nature* 370:389-391; Stemmer (1994) *Proc.*
Natl. Acad. Sci. USA 91 :10747-10751; Crameri *et al.* (1996)
Nat. Biotechnol. 14:315-319; Zhang *et al.* (1997) *Proc.*

Natl. Acad. Sci. USA 94:4504-4509; e Crameri et al. (1997) Nat. Biotechnol. 15:436-438). Estratégias de mutagênese por permutação podem ser também efetuadas. Ver, por exemplo, U.S. Provisional Application No. 60/813.095, preenchida em 5 13 de junho de 2006, aqui incorporada por referência em sua íntegra. Tais procedimentos poderiam ser efetuados, por exemplo, em polinucleotídeos codificando polipeptídeos que são resistentes a inibidores de glutamina sintetase herbicidas.

10

Em um método de hibridização, toda ou parte da seqüência de polinucleotídeos de resistência a herbicidas ou uma seqüência codificando um domínio da invenção podem ser usados para rastrear bibliotecas de cDNA ou genômico. 15 Métodos para a construção de tais bibliotecas de cDNA ou genômico são geralmente conhecidos na técnica e são descritos em Sambrook and Russell, 2001, *supra*. As chamadas sondas de hibridização podem ser fragmentos de DNA genômico, fragmentos de cDNA, fragmentos de RNA, ou outros 20 oligonucleotídeos, e podem ser rotulados com um grupamento detectável, tal como ^{32}P , ou qualquer outro marcador detectável, tal como outros radioisótopos, um composto fluorescente, uma enzima, ou um co-fator enzimático. Sondas para hibridização podem ser feitas rotulando-se 25 oligonucleotídeos sintéticos baseado na seqüência de nucleotídeos codificando resistência a herbicidas conhecida aqui descrita. *Primers* degenerados projetados com base nos resíduos de nucleotídeos ou aminoácidos conservados na seqüência de nucleotídeos ou seqüência de aminoácidos

codificada podem ser adicionalmente usados. A sonda compreende tipicamente uma região de seqüência de nucleotídeos que hibridiza sob condições estringentes a pelo menos cerca de 12, pelo menos cerca de 25, pelo menos
5 cerca de 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, ou 1400 nucleotídeos consecutivos do polinucleotídeo codificando resistência a herbicidas da invenção ou um fragmento ou variante desses. Métodos para a preparação de sondas para
10 hibridização são geralmente conhecidos na técnica e são descritos em Sambrook and Russell (2001) *supra*, e Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), ambas são aqui incorporadas por
15 referência.

A hibridização de tais seqüências pode ser conduzida sob condições estringentes. Por "condições estringentes" ou "condições de hibridização estringentes" pretende-se
20 condições sob as quais uma sonda hibridizará a sua seqüência alvo em um grau detectavelmente maior do que a outras seqüências (e.g., pelo menos 2 vezes sobre a base). Condições estringentes são dependentes de seqüência e serão diferentes em circunstâncias diferentes. Controlando a
25 estringência das condições de hibridização e/ou lavagem, seqüências alvo que são 100% complementares à sonda podem ser identificadas (sondagem homóloga). Alternativamente, condições de estringência podem ser ajustadas para permitir alguma falta de correspondência em seqüências, de forma que

níveis menores de similaridade sejam detectados (sondagem heteróloga). Geralmente, uma sonda é de menos do que cerca de 1000 nucleotídeos em comprimento, ou menos do que cerca de 500 nucleotídeos em comprimento.

5

Condições estridentes podem ser aquelas nas quais a concentração de sais é de menos do que cerca de 1,5 M de íons de Na, ou cerca de 0,01 a 1,0 M de concentração de íons de Na (ou outros sais) em pH 7,0 a 8,3 e a temperatura é de pelo menos cerca de 30 °C para sondas curtas (e.g., 10 a 50 nucleotídeos) e pelo menos cerca de 60°C para sondas longas (e.g., maiores do que 50 nucleotídeos). Condições estridentes podem ser também alcançadas com a adição de agentes desestabilizantes, tais como formamida. Condições exemplares de baixa estridência incluem hibridização com uma solução tampão de 30 a 35% de formamida, NaCl 1 M, 1% SDS (sódio dodecil sulfato) a 37°C, e uma lavagem em 1X a 2X SSC (20X SSC = NaCl 3,0 M / trissódio citrato 0,3 M) a 50 a 55°C. Condições exemplares de estridência moderada incluem hibridização em 40 a 45% formamida, NaCl 1,0 M, 1% SDS a 37°C, e uma lavagem em 0,5X a 1X SSC a 55 a 60°C. Condições exemplares de estridência elevada incluem hibridização em 50% formamida, NaCl 1 M, 1% SDS a 37°C, e uma lavagem em 0,1X SSC a 60 a 65°C por pelo menos 30 minutos. Opcionalmente, tampões de lavagem podem compreender cerca de 0,1% a cerca de 1% SDS. A duração de hibridização é de geralmente menos do que cerca de 24 horas, usualmente cerca de 4 a cerca de 12 horas.

A especificidade é tipicamente a função de lavagens pós hibridização, com os fatores críticos sendo a força iônica e temperatura da solução de lavagem final. Para híbridos DNA-DNA, o ponto de fusão (T_m) pode ser aproximado a partir da equação de Meinkoth e Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81,5^\circ\text{C} + 16,6 (\log M) + 0,41 (\%GC) - 0,61 (\% \text{ form}) - 500/L$; onde M é a molaridade de cátions monovalentes, %GC é a porcentagem de nucleotídeos guanosina e citosina no DNA, % form é a porcentagem de formamida na solução de hibridização, e L é o comprimento do híbrido em pares de bases. A T_m é a temperatura (sob definidas força iônica e pH) na qual 50% de uma seqüência alvo complementar hibridiza a uma sonda perfeitamente correspondente. A T_m é reduzida em cerca de 1 C para cada 1% de falta de correspondência; assim, condições de T_m , hibridização, e/ou lavagem podem ser ajustadas para hibridizar a seqüências da identidade desejada. Por exemplo, se seqüências com $\geq 90\%$ de identidade são procuradas, a T_m pode ser diminuída em 10°C . Por exemplo, se seqüências com $\geq 90\%$ de identidade são procuradas, a T_m pode ser diminuída em 10°C . Geralmente, condições estridentes são selecionadas para serem cerca de 5°C menores do que o ponto de fusão (T_m) para a seqüência específica e seu complemento em uma definida força iônica e pH. Contudo, condições severamente estridentes podem utilizar uma hibridização e/ou lavagem a 1, 2, 3, ou 4°C menor do que o ponto de fusão (T_m); condições moderadamente estridentes podem utilizar uma hibridização e/ou lavagem a 6, 7, 8, 9 ou 10°C menor do que o ponto de fusão (T_m);

condições de baixa estringência podem utilizar uma hibridização e/ou lavagem a 11, 12, 13, 14, 15 ou 20°C menor do que o ponto de fusão (T_m). Usando a equação, condições de hibridização e lavagem, e T_m desejada, aqueles de habilidade ordinária entenderão que variações na 5 estringência de soluções de hibridização e/ou lavagem são inerentemente descritas. Se o grau desejado de falta de correspondência resulta em uma T_m de menos do que 45°C (solução aquosa) ou 32°C (solução de formamida), a 10 concentração de SSC pode ser aumentada de forma que uma maior temperatura possa ser usada. Um guia extensivo para a hibridização de ácidos nucleicos é encontrado em Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, 15 Chapter 2 (Elsevier, New York); e Ausubel, et al., eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). Ver, Sambrook, et al., (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, 20 Plainview, New York).

E. Proteínas Isoladas e Variantes e Fragmentos destas

Polipeptídeos de resistência a herbicidas são também 25 englobados na presente invenção. Um polipeptídeo de resistência a herbicidas que é substancialmente livre de material celular inclui preparações de polipeptídeos tendo menos do que cerca de 30%, 20%, 10%, ou 5% (por peso seco) de polipeptídeo que não de resistência a herbicidas (também

referido aqui como uma "proteína contaminante"). Na presente invenção, "proteína de resistência a herbicidas" deve significar um polipeptídeo que é resistente a inibidores de glutamina sintetase herbicidas. Em alguns avanços, a proteína de resistência a herbicida confere resistência para glufosinato. Fragmentos, porções biologicamente ativas, e variantes destas também são providos, e podem ser usados para praticar os métodos da presente invenção.

10

"Fragmentos" de "porções biologicamente ativas" incluem fragmentos de polipeptídeos compreendendo uma porção de uma seqüência de aminoácidos codificando uma proteína de resistência a herbicidas e que retenha atividade de resistência a herbicida. Uma porção biologicamente de uma proteína de resistência a herbicidas pode ser um polipeptídeo que é, por exemplo, de 10, 25, 50, 100 ou mais aminoácidos em tamanho. Tais porções biologicamente ativas podem ser preparadas por técnicas de recombinação e avaliadas para atividade de resistência a herbicida.

20

Por "variantes" entende-se que são proteínas ou polipeptídeos tendo uma seqüência de aminoácido que é pelo menos de cerca de 60%, 65%, cerca de 70%, 75%, cerca de 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99% idêntica a SEQ ID NO: 2, caracterizado pelo fato de que um ou mais aminoácidos correspondendo às posições 125 até 175 e/ou posições 200-250 da SEQ ID NO:2 tenham sido

25

modificados de tal forma que o polipeptídeo é resistente ao inibidor de glutamina sintetase herbicida, ou caracterizado pelo fato de que um ou mais aminoácidos tenham sido modificados de tal forma que exista uma perda de um ou mais sítios de adenilação no polipeptídeo resultante. Esta proteína pode ser alterada de diversas formas incluindo substituições, deleções, encurtamentos, e inserções de aminoácidos de um ou mais aminoácidos da SEQ ID NO:4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, ou 46, incluindo até 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 ou mais substituições de aminoácidos, deleções ou inserções. Um especialista na técnica irá reconhecer que esses valores podem ser apropriadamente ajustados para determinar a identidade correspondente do polipeptídeo codificado por dois polinucleotídeos levando-se em conta a degeneração de códon, a similaridade de aminoácidos, o posicionamento da fase de leitura, e similares.

Por exemplo, substituições conservadas de aminoácidos podem ser feitas em um ou mais resíduos de aminoácidos não essenciais. Um resíduo de aminoácido "não essencial" é um resíduo que pode ser modificado a partir da seqüência selvagem de um polipeptídeo sem alterar a atividade biológica, enquanto que um resíduo de aminoácido

"essencial" é necessário para a atividade biológica. Uma "substituição conservada de aminoácido" é uma na qual o resíduo de aminoácido é substituído por um resíduo de aminoácido tendo uma cadeia lateral similar. Famílias de resíduos de aminoácidos tendo cadeias laterais similares foram definidas na técnica. Estas famílias incluem aminoácidos com cadeias laterais básicas (por exemplo, lisina, arginina, histidina), cadeias laterais ácidas (por exemplo, ácido aspártico, ácido glutâmico), cadeias laterais sem carga (por exemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadeias laterais apolares (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). Substituições de aminoácidos podem ser feitas em regiões não conservadas que retenham a função. Em geral, tais substituições não seriam feitas para resíduos de aminoácidos conservados, ou para resíduos de aminoácidos residindo em um motivo conservado, onde tais resíduos são essenciais para a atividade do polipeptídeo. Entretanto, um especialista na técnica entenderia que variantes funcionais podem ter pequenas alterações conservativas ou não conservativas ou modificações nos resíduos conservados.

Variantes também incluem polipeptídeos codificados por um polinucleotídeo que hibridiza com o polinucleotídeo que codifica um polipeptídeo que é resistente ao inibidor

glutamina sintetase herbicida, ou uma complemento desse, sob condições estridentes. Variantes incluem polipeptídeos que diferem na seqüência de aminoácido devido a mutagênese. Proteínas variantes englobadas pela presente invenção são

5 biologicamente ativas, isto é, continuam a possuir a atividade biológica desejada da proteína nativa, que é reter a atividade de resistência a herbicidas. Métodos para medir a atividade de resistência a herbicidas são bem conhecidos na técnica. Ver, por exemplo, Patente U.S. Nos.

10 5.098.838 e 5.145.777, cada uma das quais é aqui incorporada por referência na sua íntegra.

Genes bacterianos muito freqüentemente possuem múltiplos códons de iniciação para metionina nas

15 proximidades do início do quadro de iniciação de leitura. Freqüentemente, a iniciação da tradução em um ou mais destes códons de iniciação levará à geração de uma proteína funcional. Esses códons de iniciação podem incluir códons ATG. Entretanto, bactérias como *Bacillus* sp. também

20 reconhecem o códon GTG como um códon de iniciação, e proteínas que iniciam a tradução em códons GTG contém uma metionina como primeiro aminoácido. Adicionalmente, não é freqüentemente determinado *a priori* quais desses códons são

25 usados naturalmente na bactéria. Então, é compreendido que o uso de um desses códons alternativos de metionina pode levar à geração de variantes que conferem resistência a herbicidas. Essas proteínas de resistência a herbicidas são englobadas na presente invenção e podem ser usadas nos métodos da presente invenção.

Anticorpos para os polipeptídeos da presente invenção, ou para variantes ou fragmentos desses, também são englobados. Métodos para a produção de anticorpos são bem conhecidos na técnica (ver, por exemplo, Harlow and Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Patente U.S. No. 4.196.265).

10 *F. Construções de polinucleotídeos*

Os polinucleotídeos empregados nos métodos e composições da invenção podem ser modificados para se obter ou aumentar a expressão nas células vegetais. Os polinucleotídeos codificando os polipeptídeos da invenção podem ser providos em cassetes de expressão para expressão na planta de interesse. Um "cassete de expressão em planta" inclui uma construção de DNA que é capaz de resultar na expressão de um polinucleotídeo em uma célula vegetal. O cassete pode incluir na direção de transcrição 5'-3', uma região de iniciação transcricional (ou seja, promotor) operacionalmente ligado a um ou mais polinucleotídeos de interesse, e uma região de terminação de tradução e de transcrição (ou seja, região de terminação) funcional em plantas. O cassete pode conter adicionalmente pelo menos um polinucleotídeo adicional para ser introduzido no organismo, como um gene marcador selecionável. Alternativamente, o(s) polinucleotídeo(s) adicional(is) pode ser provido em múltiplos cassetes de expressão. Tal cassete de expressão é provido com uma pluralidade de

sítios de restrição para inserção do polinucleotídeo a estar sob regulação transcricional das regiões reguladoras.

"Heterólogo" geralmente refere-se ao polinucleotídeo ou polipeptídeo que não é endógeno da célula ou não é endógeno ao local no genoma nativo no qual está presente, e tenha sido adicionado à célula por infecção, transfecção, microinjeção, eletroporação, microprojeção, ou similares. Por "operacionalmente ligado" pretende-se uma ligação funcional entre dois polinucleotídeos. Por exemplo, quando um promotor é operacionalmente ligado a uma seqüência de DNA, a seqüência promotora inicia e medeia a transcrição da seqüência de DNA. É reconhecido que os polinucleotídeos ligados operacionalmente podem ou não ser contíguos e, onde usados para referenciar a junção de duas regiões codificadoras de polipeptídeos, os polipeptídeos são expressos na mesma fase de leitura.

O promotor pode ser qualquer seqüência de polinucleotídeos que demonstrar atividade na célula vegetal escolhida, parte da planta, ou planta. O promotor pode ser nativo ou análogo, ou exógeno ou heterólogo, à planta hospedeira e/ou à seqüência de DNA da invenção. Quando um promotor é "nativo" ou "análogo" à planta hospedeira, pretende-se que este promotor seja encontrado na planta nativa na qual o promotor é introduzido. Quando o promotor é "exógeno" ou "heterólogo" à seqüência de DNA da invenção, pretende-se que o promotor não seja nativo ou naturalmente ocorrente para a seqüência de DNA a qual está

operacionalmente ligado na invenção. O promotor pode ser induzível ou constitutivo. Este pode ser naturalmente-ocorrente, pode ser composto de porções de vários promotores naturalmente-ocorrentes, ou pode ser parcialmente ou totalmente sintético. Uma orientação para o desenho de promotores é fornecida por estudos de estrutura de promotor, tais como os de Harley and Reynolds (1987) *Nucleic Acids Res.* 15: 2343-2361. Ainda, a localização do promotor em relação ao início da transcrição pode ser otimizada. Ver, por exemplo, Roberts et al. (1979) *Proc. Natl Acad. Sci. USA*, 76: 760-764. Vários promotores adequados para o uso em plantas são bem conhecidos na técnica.

Por exemplo, promotores constitutivos adequados para o uso em plantas incluem: os promotores para viroses de plantas, tais como o promotor de caulimovírus da faixa clorótica de amendoim (PClSV) (Patente U.S. No. 5.850.019); o promotor 35S de vírus mosaico de couve-flor (CaMV) (Odell et al. (1985) *Nature* 313: 810-812); promotores dos genes da metiltransferase do vírus de *Chlorella* (Patente U.S. No. 5.563.328) e o promotor transcrito em tamanho integral de vírus mosaico da escrofulária (FMV) (Patente U.S. No. 5.378.619); os promotores de genes tais qual actina de arroz (McElroy et al. (1990) *Plant Cell* 2:163-171); ubiquitina (Christensen et al. (1989) *Plant Mol. Biol.* 12: 619-632 e Christensen et al. (1992) *Plant Mol. Biol.* 18: 675-689); pEMU (Last et al. (1991) *Theor. Appl. Genet.* 81: 581-588); MAS (Velten et al. (19M) *EMBO J.* 3:2723-2730);

histona H3 de milho (Lepelet et al. (1992) *Mol. Gen. Genet.* 231: 276-285 e Atanassova et al. (1992) *Plant J.* 2(3): 291-300); ALS3 de *Brassica napus* (PCT application WO 97/41228); e promotores de vários genes de *Agrobacterium* (ver Patente 5 U.S. Nos. 4.771.002; 5.102.796; 5.182.200; e 5.428.147).

Promotores induzíveis adequados para uso em plantas incluem: o promotor de sistema ACE1, o qual responde a cobre (Mett et al. (1993) *PNAS* 90:4567-4571); o promotor do 10 gene In2 de milho, o qual responde a protetores de herbicida benzenosulfonamida (Hershey et al. (1991) *Mol. Gen. Genetics* 227:229-237 e Gatz et al. (1994) *Mol. Gen. Genetics* 243:32-38); e o promotor do repressor Tet de TnIO (Gatz et al. (1991) *Mol. Gen. Genet.* 227:229-237). Outro 15 promoter induzível para uso em plantas é um que responda a um agente indutor ao qual plantas normalmente não respondem. Um exemplar de promotor induzível deste tipo é o promotor induzível proveniente de um gene de hormônio esteróide, o qual a atividade transcricional é induzida 20 pelo hormônio glucocorticosteróide (Sчена et al. (1991) *Proc. Natl. Acad. Sci. USA* 88: 10421) ou a aplicação recente de um ativador transcricional quimérico, XVE, para uso em um sistema de expressão de planta induzível a base de um receptor de estrogênio e ativado por estradiol (Zuo 25 et al. (2000) *Plant J.* 24: 265-273). Outros promotores induzíveis para uso em plantas são descritos em EP 332104, PCT WO 93/21334 e PCT WO 97/06269 os quais são aqui incorporados por referência na íntegra. Promotores compostos de porções de outros promotores e parcialmente ou

totalmente sintéticos podem também ser utilizados. Ver, por exemplo, Ni et al. (1995) *Plant J.* 7: 661-676 e PCT WO 95/14098 descrevendo tais promotores para uso em plantas.

5 O promotor pode incluir, ou ser modificado para incluir, um ou mais elementos intensificadores. Em alguns avanços, o promotor pode incluir uma pluralidade de elementos intensificadores. Promotores contendo elementos intensificadores provêm para altos níveis de transcrição em
10 comparação com promotores que não os incluem. Elementos intensificadores adequados para o uso em plantas incluem o elemento intensificador PC1SV (Patente U.S. No. 5.850.019), o elemento intensificador CaMV 35S (Patente U.S. Nos. 5.106.739 e 5.164.316) e o elemento intensificador FMV
15 (Maiti et al. (1997) *Transgenic Res.* 6:143-156). Ver também PCT WO 96/23898.

Freqüentemente, tais construções podem conter regiões não transcritas a 5' e 3'. Tais construções podem conter
20 uma "seqüência sinal" ou "seqüência líder" para facilitar o transporte co-traducional ou pós-traducional do peptídeo de interesse para certas estruturas intracelulares como o cloroplasto (ou outro plastídeo), retículo endoplasmático, ou complexo de Golgi, ou para ser secretado. Por exemplo, a
25 construção pode ser elaborada para conter um peptídeo sinal para facilitar a transferência do peptídeo para o retículo endoplasmático. Por "seqüência sinal" pretende-se uma seqüência que é sabida ou suspeita de estar na origem do transporte co-tradução ou pós-tradução do peptídeo através

da membrana celular. Em eucariotos, isso envolve tipicamente secreção para o complexo de Golgi, com alguma glicosilação resultante. Por "seqüência líder" pretende-se qualquer seqüência que, quando traduzida, resulte em uma

5 seqüência de aminoácido suficiente para engatilhar o transporte co-traducional da cadeia peptídica para uma organela sub-celular. Sendo assim, isto inclui seqüências líder direcionando o transporte e/ou glicosilação por passagem por dentro do retículo endoplasmático, passagem

10 para vacúolos, plastídeos incluindo cloroplastos, mitocôndria, e similares. Também pode ser preferível construir o cassete de expressão em planta para conter um íntron, de forma que o processamento do íntron deste mRNA seja necessário para a expressão.

15 Por "região 3' não traduzida" pretende-se um polinucleotídeo localizado abaixo da seqüência codificadora. Seqüências sinal de poliadenilação e outras seqüências codificando sinais regulatórios capazes de afetar a adição de tratos de ácido poliadenilados à

20 extremidade 3' do mRNA precursor são regiões 3' não traduzidas. Por "regiões 5' não traduzidas" pretende-se um polinucleotídeo localizado acima da seqüência codificadora.

Outros elementos não traduzidos acima ou abaixo

25 incluem os intensificadores. Intensificadores são polinucleotídeos que agem para aumentar a expressão de uma região promotora. Intensificadores são bem conhecidos na técnica e incluem, mas não são limitados a, uma região intensificadora SV40 e o elemento intensificador 35 S.

A região de terminação pode ser nativa em relação à região de iniciação de transcrição, pode ser nativa com seqüência da presente invenção, ou pode ser derivada de
5 outra fonte. Regiões de terminação convenientes são disponíveis a partir do plasmídeo Ti de *A. tumefaciens*, tais como as regiões de terminação de octopina sintase e de nopalina sintase. Ver também Guerineau et al. (1991) *Mol. Gen. Genet.* 262: 141-144; Proudfoot (1991) *Cell* 64: 671 -
10 674; Sanfacon et al. (1991) *Genes Dev.* 5: 141-149; Mogen et al. (1990) *Plant Cell* 2: 1261-1272; Munroe et al. (1990) *Gene* 91: 151-158; Ballas et al. (1989) *Nucleic Acids Res.* 17: 7891-7903; e Joshi et al. (1987) *Nucleic Acid Res.* 15: 9627-9639.

15

Em um aspecto da invenção, seqüências sintéticas de DNA são desenhadas para um dado polipeptídeo, tais como os polipeptídeos da invenção. A expressão da fase aberta de leitura da seqüência sintética de DNA em uma célula resulta
20 na produção de um polipeptídeo da invenção. Seqüências de DNA sintéticas podem ser úteis para simplesmente remover sítios de restrição indesejáveis, para facilitar estratégias de clonagem de DNA, para alterar ou remover qualquer potencial *bias* de códon, para alterar ou aumentar
25 o conteúdo de GC, para remover ou alterar quadros de leitura alternados, e/ou para alterar ou remover sítios de reconhecimento de *splice* de íntron/éxon, sítios de poliadenilação, seqüências Shine-Delgarno, elementos promotores indesejáveis e similares que possam estar

presentes na seqüência de DNA nativa. Também é possível que as seqüências de DNA sintéticas possam ser utilizadas para introduzir outros melhoramentos para uma seqüência de DNA, tais como a introdução de uma seqüência de íntron, criação de uma seqüência de DNA que é expressa como proteína de fusão para seqüências com direcionamento para organelas, como peptídeos de trânsito para cloroplasto, peptídeos direcionados para apoplasto/vacúolos, ou seqüências peptídicas que resultem na retenção do peptídeo resultante no retículo endoplasmático. Genes sintéticos também podem ser sintetizados utilizando-se códons preferenciais da célula hospedeira para aumento da expressão, ou podem ser sintetizados utilizando-se códons em uma freqüência de uso preferida pelo hospedeiro. Ver, por exemplo, Campbell and Gowri (1990) *Plant Physiol.* 92: 1-11; Patente U.S. Nos. 6.320.100; 6.075.185; 5.380.831; e 5.436.391, U.S. Published Application Nos. 20040005600 e 20010003849, e Murray et al. (1989) *Nucleic Acids Res.* 17: 477-498, aqui incorporadas por referência.

20

Em um avanço, os polinucleotídeos de interesse são direcionados para o cloroplasto para expressão. Dessa maneira, aonde o polinucleotídeo de interesse não for diretamente inserido no cloroplasto, o cassete de expressão irá conter adicionalmente um polinucleotídeo codificando um peptídeo de trânsito para direcionar o nucleotídeo de interesse para os cloroplastos. Tais peptídeos de trânsito são conhecidos na técnica. Ver, por exemplo, Von Heijne et al. (1991) *Plant Mol. Biol. Rep.* 9: 104-126; Clark et al.

(1989) *J. Biol. Chem.* 264: 17544-17550; 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.

5

Os polinucleotídeos de interesse a serem direcionados para o cloroplasto podem ser otimizados para expressão no cloroplasto para dar conta de diferenças de utilização de códon entre o núcleo vegetal e esta organela. Desta
10 maneira, os polinucleotídeos de interesse podem ser sintetizados usando-se códon de preferência do cloroplasto. Ver, por exemplo, Patente U.S. No. 5.380.831, aqui incorporada como referencia.

15 Este cassete de expressão vegetal pode ser inserido em um vetor de transformação vegetal. Por "vetor de transformação" pretende-se uma molécula de DNA que permite que a célula seja transformada. Tais moléculas podem consistir de um ou mais cassetes de expressão, e podem ser
20 organizadas dentro de uma ou mais moléculas de DNA vetores. Por exemplo, vetores binários são vetores de transformação de plantas que são utilizados em dois vetores de DNA não contíguos para codificar todos os requisitos para funções cis- e trans-atuantes para transformação de células
25 vegetais (Hellens and Mullineaux (2000) *Trends in Plant Science* 5: 446-451). "Vetor" refere-se a uma construção de polinucleotídeo projetada para transferência entre diferentes células hospedeiras. "Vetor de expressão" refere-se a um vetor que tenha a habilidade de incorporar,

integrar e expressar seqüências de DNA heterólogas ou fragmentos em uma célula exógena.

Vetor de transformação vegetal compreende um ou mais
5 vetores de DNA para alcançar a transformação de plantas. Por exemplo, é uma prática comum na técnica utilizar vetores de transformação de plantas que compreendam mais de um segmento de DNA contíguo. Estes vetores são freqüentemente referidos na técnica como vetores binários.

10 Vetores binários, assim como vetores com plasmídeos ajudantes são mais freqüentemente usados para transformação mediada por *Agrobacterium*, onde o tamanho e a complexidade dos segmentos de DNA devem alcançar uma eficiência de transformação bem grande, e esta é vantajosa para separar

15 as funções em moléculas separadas de DNA. Vetores binários tipicamente contêm um vetor plasmidial que contém as seqüências cis-atuantes requeridas para a transferência do T-DNA (tais como fronteira esquerda e fronteira direita), um marcador selecionável que é construído para ser capaz de

20 expressar na célula vegetal, e um "polinucleotídeo de interesse" (um polinucleotídeo construído para ser capaz de expressão na célula vegetal para a qual se deseja a geração de transgênicos). Também presentes nesse vetor plasmidial estão seqüências necessárias para replicação bacteriana. As

25 seqüências cis-atuantes são arranjadas de forma que se permita a transferência eficiente para dentro da célula vegetal e a expressão nesta. Por exemplo, a seqüência de um marcador selecionável e a seqüência de interesse são localizadas entre as fronteiras esquerda e direita.

Freqüentemente, um segundo vetor plasmidial contém os fatores trans-atuantes que medeiam a transferência do T-DNA de *Agrobacterium* para as células vegetais. Esse plasmídeo normalmente contém as funções de virulência (genes Vir) que permitem a infecção das células vegetais por *Agrobacterium*, e a transferência de DNA por clivagem nas seqüências das bordas e transferência mediada por vir, como é conhecido na técnica (Hellens and Mullineaux (2000) *Trends in Plant Science*, 5: 446-451). Diversos tipos de linhagens de *Agrobacterium* (por exemplo, LBA4404, GV3101, EHA101, EHA105, etc.) podem ser usadas para a transformação de plantas. O segundo vetor plasmidial não é necessário para a introdução dos polinucleotídeos nas plantas por outros métodos, tais como microprojeção, microinjeção, eletroporação, glicol polietileno, etc.

G. Expressão de genes de tolerância a herbicidas e tolerância a insetos

As plantas tolerantes de inibidores de glutamina sintetase aqui descritas podem exibir, ainda, resistência ou tolerância para um ou mais herbicidas (além de inibidores GS) e/ou uma ou mais pestes tais como insetos, nematóides ou fungos. Em alguns avanços, uma ou mais das plantas descritas aqui exibem tolerância ou resistência para um ou mais herbicidas além de inibidores de GS. Um número de genes que conferem resistência a herbicida, tanto transgênicos e não transgênicos, estão disponíveis. Genes conferindo resistência a um herbicida que inibe o ponto de crescimento ou meristema, tais como uma imidazolinona ou

uma sulfoniluréia podem ser adequados. Genes exemplares nesta categoria codificam para enzimas mutantes ALS e AHAS são descritos, por exemplo, na Patente U.S. Nos. 5.767.366 e 5.928.937. Patente U.S. Nos. 4.761.373 e 5.013.659 são 5 direcionados para plantas resistentes a vários herbicidas de imidazolinona ou de sulfonamida.

Genes de resistência a glifosato, tais como genes de EPSP sintase resistente a glifosato, são particularmente 10 úteis nos métodos e composições divulgados aqui. Ver, por exemplo, Pedido de Patente U.S. Nos. 11/500.718, 11/185.342, 11/185.560, 11/315.678, 11/312.866, 11/400.598, 11/605.824, e 11/651.752, Patente U.S. No. 4.940.835 e Patente U.S. No. 4.769.061, cada uma das quais é aqui 15 incorporada por referência em sua íntegra. Patente U.S. No. 5.554.798 descreve plantas de milho transgênicas resistentes a glifosato, nas quais a resistência é conferida por um gene de 5-enolpiruvil-3-fosfochiquimato sintase (EPSP) alterado.

20

Genes de resistência para compostos fosfono tais como glufosinato de amônio ou fosfinotricina, e ácidos piridinoxi ou fenoxi propiônico e ciclohexonas também são adequados. Ver Pedido de patente Européia No. 0 242 246. 25 Ver também, Patente U.S. Nos. 5.879.903, 5.276.268 e 5.561.236. Outros herbicidas adequados incluem estes que inibem a fotossíntese, tais como uma triazina e uma benzonitrila (nitrilase) (ver Patente U.S. No. 4.810.648) assim como herbicidas tais como ácido 2,2-

dicloropropiônico, setoxidim, haloxifope, herbicidas de imidazolinona, herbicidas de sulfoniluréia, herbicidas de triazolopirimidina, herbicidas de s-triazina e bromoxinil.

5 As proteínas inseticidas úteis para a invenção podem ser expressas em uma ou mais plantas divulgadas aqui. Genes úteis para resistência a insetos ou pestes incluem, por exemplo, genes para endotoxina codificando toxinas identificadas em organismos *Bacillus*. Genes codificando
10 toxinas de *Bacillus thuringiensis* (Bt) de diversas subespécies têm sido clonados e clones recombinantes foram verificados como sendo tóxicos para larvas de insetos lepidópteros, díptera e coleóptera. Ver, por exemplo, Pedido de Patente U.S. Nos. 10/782.020, 10/782.141,
15 10/782.570, 10/783.417, 10/781.979, 10/782.096, 10/926.819, e 11/343.533, cada qual é aqui incorporada por referência em sua íntegra. Vários outros genes para delta-endotoxina tais como CryIAa, CryIAb, CryIAc, CryIB, CryIC, CryID, CryIEa, CryIFa, Cry3A, Cry9A, Cry9C e Cry9B; assim como
20 genes codificando proteínas inseticidas vegetativas tais como Vip1, Vip2 e Vip3), também são úteis nos métodos e composições aqui descritos. Uma lista completa de toxinas Bt pode ser encontrada na rede mundial em www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/.

25

H. Plantas e partes de plantas

Por "planta" pretende-se plantas inteiras, órgãos de planta (por exemplo, folhas, caules, raízes, etc.),

sementes, células vegetais, propágulos, embriões e descendentes dos mesmos. Células vegetais podem ser diferenciadas ou indiferenciadas (por exemplo, calo, de células de cultura em suspensão, protoplastos, células de 5 folha, células de raiz, células de floema, pólen). A presente invenção pode ser usada para introdução de polinucleotídeos em qualquer espécie de planta, incluindo, mas não se limitando a, monocotiledôneas e dicotiledôneas. Exemplos de plantas de interesse incluem, mas não são 10 limitadas a, milho, sorgo, trigo, girassol, tomate, crucíferas, pimentas, batata, algodão, arroz, soja, beterraba, cana-de-açúcar, tabaco, cevada, e colzas, *Brassica* sp., alfafa, centeio, milheto, cártamo, amendoim, batata doce, mandioca, café, coco, abacaxi, frutas 15 cítricas, cacau, chá, banana, abacate, figo, goiaba, manga, azeitona, mamão, castanhas, macadâmia, amêndoa, aveia, vegetais, plantas ornamentais, e coníferas.

Verduras incluem, mas não estão limitadas a, tomates, 20 alface, feijão verde, feijões, ervilhas, e membros do gênero *Curcumis* tais como pepino, cantalupo, e meloeiro. Plantas ornamentais incluem, mas não são limitadas a, azaléia, hortênsia, hibisco, rosas, tulipas, narcisos, petúnias, cravo, bico-de-papagaio, e crisântemo. Plantas de 25 cultivo também são de interesse, incluindo, por exemplo, milho, sorgo, trigo, girassol, tomate, crucíferas, pimentas, batatas, algodão, arroz, soja, beterraba, cana-de-açúcar, tabaco, cevada, colza, etc.

Esta invenção é adequada para qualquer membro de plantas da família das monocotiledôneas incluindo, mas não se limitando a, milho, arroz, cevada, aveia, trigo, sorgo, centeio, cana-de-açúcar, abacaxi, inhames, cebola, banana, 5 coco, e tâmaras.

II. Métodos

A. Métodos para crescimento de propriedades 10 agronomicamente importantes em plantas

Métodos para aumentar propriedades agronomicamente importantes de plantas também são providos. Esses métodos compreendem introduzir dentro de uma planta ou de uma 15 célula vegetal uma seqüência de nucleotídeos codificando uma enzima glutamina sintetase derivada de bactéria. Por "enzima glutamina sintetase derivada de bactéria" pretende-se uma enzima glutamina sintetase isolada a partir de uma bactéria, ou um variante biologicamente ativo ou fragmento 20 dessa. Em um avanço, a seqüência de nucleotídeos compreende uma variante da SEQ ID NO: 1, onde o polinucleotídeo variante é ao menos 80% idêntico a SEQ ID NO: 1. Em outro avanço, a seqüência de nucleotídeos compreende um polinucleotídeo tendo ao menos uma modificação entre os 25 aminoácidos 125 a 175, ao menos uma modificação entre os aminoácidos 200 a 250 correspondendo a SEQ ID NO: 2, ou ao menos uma modificação que resulte na perda de um ou mais sítios de adenilação. Em outro avanço, o polinucleotídeo é selecionado a partir do grupo constituído da SEQ ID NOS: 3,

5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, ou 47. A expressão dessas enzimas em plantas resulta no aumento das capacidades da planta de assimilação e/ou utilização de nitrogênio, assim como
5 melhoram características agronômicas tais como rendimento da planta.

Conforme definido aqui, o "rendimento" da planta refere-se à qualidade e/ou quantidade de biomassa produzida
10 pela planta. Por "biomassa" pretende-se qualquer produto vegetal medido. Um aumento na produção de biomassa é qualquer melhoramento no rendimento do produto de planta medido. O método compreende introduzir dentro de uma planta de interesse um polinucleotídeo codificando uma enzima
15 glutamina sintetase derivada de bactéria. Em um avanço, a enzima glutamina sintetase é resistente aos inibidores de glutamina sintetase herbicidas, apesar da resistência não ser necessária para se alcançar o aumento das propriedades agronômicas. Em outro avanço, a enzima glutamina sintetase
20 aumentou a atividade enzimática relativa a uma enzima glutamina sintetase controle como definido *infra*.

Embora não vinculado a qualquer teoria ou mecanismo em particular, a expressão de uma enzima glutamina sintetase
25 derivada de bactéria em uma planta pode levar a uma atividade aumentada (resultando em um aumento do rendimento e/ou utilização de nitrogênio) comparada com a glutamina sintetase derivada de planta (incluindo a glutamina sintetase endógena na qual a sintetase derivada de bactéria

é expressa de maneira heteróloga) devido a mecanismos regulatórios diferentes para a GS bacteriana em comparação com a GS de plantas. A atividade enzimática da GS bacteriana é regulada de uma maneira que é diferente das 5 enzimas GS de plantas (Moorhead and Smith (2003) *Plant Physiol* 133: 492-498, aqui incorporada por referência em sua íntegra). Em um sistema bacteriano, o *status* do nitrogênio na célula é sentido pela proteína PII. Sob condições de altas concentrações de nitrogênio, PII inicia 10 uma cascata de sinalização que causa a adenilação de subunidades individuais de enzimas GS bacterianas. Adenilação de GS bacterianas causa um decréscimo na atividade enzimática. Então, a atividade enzimática de uma enzima GS bacteriana pode ser modulada pela extensão da 15 adenilação do duodecâmero da GS. Em contraste, desde que plantas não possuem um análogo da cascata de sinalização PII, é improvável que as células vegetais causem a adenilação de uma enzima GS bacteriana.

20 O desenvolvimento de variedades de plantas que usam nitrogênio mais eficientemente irá reduzir a necessidade de excessivas entradas de nitrogênio, economizar o custo da produção para os fazendeiros, beneficiar os fazendeiros nos países em desenvolvimento que não têm acesso às entradas de 25 fertilizantes, e reduzir a poluição associada com a aplicação excessiva de fertilizantes de nitrogênio. Adicionalmente, prover plantas com rendimento aumentado como resultado da atividade de glutamina sintetase melhorada possui diversas aplicações comerciais. Por

exemplo, aumentar a biomassa da folha de planta pode elevar o rendimento de vegetais folhosos para consumo humano ou animal. Ademais, o aumento da biomassa foliar pode ser usado para aumentar a produção de farmacêuticos derivados de planta ou produtos industriais.

De acordo com a presente invenção, as plantas expressando glutamina sintetase derivada de bactéria podem exibir conteúdos nitrogenados aumentados, composições de aminoácidos ou proteínas alterados, características de crescimento vigoroso, rendimento vegetativo aumentado ou melhores rendimentos e qualidades de sementes. Estas plantas podem ser identificadas examinando-se qualquer dos parâmetros a seguir: 1) a taxa de crescimento, medida em termos da taxa de crescimento por peso fresco ou seco; 2) rendimento vegetativo da planta madura, em termos de peso fresco ou seco; 3) rendimento da semente ou fruta; 4) o peso da semente ou fruta; 5) o conteúdo total de nitrogênio da planta; 6) o conteúdo total de nitrogênio da fruta ou semente; 7) o conteúdo de aminoácido livre da planta; 8) o conteúdo de aminoácido livre da fruta ou semente; 9) o conteúdo de proteína total da planta; e 10) o conteúdo de proteína total da fruta ou semente. Os procedimentos e métodos para examinar estes parâmetros são bem conhecidos para aqueles especialistas na técnica. Estes métodos podem envolver testes enzimáticos e imuno-ensaios para medir os níveis de enzima/proteína; testes para medir a composição de aminoácidos, agrupamento de aminoácidos livres ou conteúdo total de nitrogênio de vários tecidos de planta;

medição da taxa de crescimento em termos de peso fresco e ganho ao longo do tempo; ou medição de rendimento da planta em termos de peso seco total e/ou peso total da semente.

5 A medição pode ser *in vitro* em uma célula expressando a enzima glutamina sintetase ou no material da planta coletado a partir de uma planta expressando a enzima, ou pode ser *in vivo* em uma planta expressando a enzima. O rastreamento pode ser realizado sob condições de

10 deficiência de nitrogênio ou sob condições não limitantes de nitrogênio. Condições de nitrogênio são descritas em relação ao nutriente nitrogenado disponível. Condições deficientes de nitrogênio incluem aquelas que fazem com que o crescimento de uma planta controle cesse ou que seja

15 diminuído de tal forma como a reduzir significativamente o tamanho ou qualidade da planta controle. Condições não limitantes de nitrogênio incluem aquelas tendo quantidades suficientes de nutrientes nitrogenados para sustentar o crescimento saudável da planta. Condições de nitrogênio que

20 constituem não limitantes ou deficientes são conhecidas na técnica para a maioria, se não todas, as variedades de plantas de interesse. Orientações adicionais podem ser encontradas em, por exemplo, Hewitt (1966) *Sand and Water Culture Methods Used in the Study of Plant Nutrition*, 2nd

25 ed., Farnham Royal (Bucks), Commonwealth Agricultural Bureaux; e, Hewitt (1975) *Plant Mineral Nutrition*, London, English University Press.

Para os propósitos da presente invenção, uma melhoria em qualquer das características acima é em relação ao controle do crescimento da planta ou da célula vegetal sob condições similares. Uma planta ou célula vegetal "controle" é aquela que expressa uma enzima glutamina sintetase que não é uma glutamina sintetase derivada de bactéria. Uma melhoria em qualquer um desses parâmetros pode compreender qualquer aumento incluindo, mas não se limitando a, pelo menos um aumento de 1%, pelo menos um aumento de 3%, pelo menos um aumento de 5%, pelo menos um aumento de 10%, pelo menos um aumento de 20%, pelo menos de 30%, pelo menos de 50%, pelo menos de 70%, pelo menos de 100% ou um aumento maior em um ou mais desses parâmetros.

Em vários avanços, a enzima GS derivada de bactéria tem atividade enzimática melhorada quando comparada a outras enzimas glutamina sintetases derivadas de bactérias ou plantas. Uma enzima glutamina sintetase com atividade melhorada é uma que tem atividade acima da atividade da enzima AGS1 divulgada aqui como SEQ ID NO: 2. Em alguns avanços, a enzima GS derivada de bactéria tem atividade melhorada devido a mutações funcionais na enzima, ao invés da super expressão da enzima em um sistema. A atividade pode ser medida por qualquer método conhecido na técnica. Ao menos que especificado de outra maneira, a enzima glutamina sintetase derivada de bactéria com a atividade melhorada é aquela que tenha atividade melhorada quando comparada à atividade da mesma ou substancialmente a mesma

concentração de SEQ ID NO: 2 quando expressa em um sistema bacteriano (por exemplo, em *E. coli*).

B. Transformação de Planta

5 Métodos da invenção envolvem introduzir um ou mais polinucleotídeos em uma planta. Por "introduzir" pretende-se apresentar à planta o polinucleotídeo de tal maneira que este polinucleotídeo ganhe acesso ao interior de uma célula de planta. Os métodos da invenção não requerem que um
10 método particular para introdução de um polinucleotídeo na planta seja usado, apenas que o polinucleotídeo ganhe acesso ao inteiro de ao menos uma célula da planta.

A introdução de um polinucleotídeo em células vegetais
15 é realizada por uma das diversas técnicas conhecidas na técnica, incluindo, mas não se limitando a eletroporação ou transformação química (Ver, por exemplo, Ausubel, ed. (1994) *Current Protocols in Molecular Biology* (John Wiley and Sons, Inc., Indianapolis, IN). Marcadores conferindo
20 resistência a substâncias tóxicas são úteis na identificação de células transformadas (que tenham apanhado e expressado a seqüência de polinucleotídeos teste) de células não transformadas (estas não contendo ou não expressando a seqüência de polinucleotídeos teste). Em um
25 aspecto da invenção, as seqüências de polinucleotídeos aqui descritas são úteis como marcadores para avaliar a introdução do DNA em células vegetais. Métodos para se detectar a presença de um transgene em uma planta, órgão de planta (por exemplo, folhas, caules, raízes, etc.),

semente, célula vegetal, propágulo, embrião ou descendentes destes são bem conhecidos na técnica. "Plantas transgênicas" ou "plantas transformadas" ou plantas, células, tecidos ou sementes "transformadas de forma estável" referem-se a plantas que tenham incorporado ou integrado polinucleotídeos exógenos na célula vegetal. Por "transformação estável" pretende-se que a construção de polinucleotídeos introduzida em uma planta se integre no genoma dessa planta e seja capaz de ser herdada pela progênie desta.

Em geral, os métodos de transformação de plantas envolvem transferir DNA heterólogo para dentro de uma célula vegetal alvo (por exemplo, embriões imaturos ou maduros, culturas de suspensão, calo indiferenciado, protoplastos, etc.), seguido da aplicação de um limiar máximo de três vezes o nível de seleção apropriado (dependendo do gene marcador selecionável) para se recuperar as células vegetais transformadas a partir de um grupo de massa de células não transformadas. Explantes são tipicamente transferidos para um suporte fresco do mesmo meio e cultivados rotineiramente. Subseqüentemente, as células transformadas são diferenciadas em brotos após a colocação em meio de regeneração suplementado com um nível limite máximo do agente de seleção de três vezes (ou seja, o herbicida). Os brotos são então transferidos para um meio de enraizamento seletivo recuperar os brotos com raiz ou plântulas. As plântulas transgênicas então crescem em plantas maduras e produzem sementes férteis (por exemplo,

Hiei et al. (1994) *Plant J.* 6: 271-282; Ishida et al. (1996) *Nat. Biotechnol.* 14: 745-750). Uma descrição geral de técnicas e métodos para gerar plantas transgênicas é encontrada em Ayres and Park (1994) *CRC Crit. Rev. Plant Sci.* 13: 219-239 e Bommineni and Jauhar (1997) *Maydica* 42: 107-120. Como o material transformado contém muitas células, tanto células transformadas e não transformadas estão presentes em qualquer pedaço do calo ou tecido ou grupo de células submetido a alvo. A habilidade de matar células não transformadas e permitir que as células transformadas proliferem resulta em culturas da planta transformadas. Freqüentemente, a habilidade para remoção de células não transformadas é uma limitação à recuperação rápida de células vegetais transformadas e geração com sucesso de plantas transgênicas. Métodos moleculares e bioquímicos podem ser utilizados para se confirmar a presença e integridade do polinucleotídeo(s) de interesse no genoma da planta transgênica.

A geração de plantas transgênicas pode ser realizada por um dentre diversos métodos, incluindo, mas não se limitando a, introdução de DNA heterólogo por *Agrobacterium* em células vegetais (transformação mediada por *Agrobacterium*), bombardeamento de células vegetais com DNA estrangeiro heterólogo aderido a partículas, e vários outros métodos diretamente-mediados e não particulados (por exemplo, Hiei et al. (1994) *Plant J.* 6: 271-282; Ishida et al. (1996) *Nat. Biotechnol.* 14: 745-750; Ayres and Park (1994) *CRC Crit. Rev. Plant Sci.* 13: 219-239; Bommineni and

Jauhar (1997) *Maydica* 42: 107-120) para transferência de DNA.

Existem três métodos comuns de transformação de célula vegetal com *Agrobacterium*. O primeiro método é o co-cultivo de *Agrobacterium* com protoplastos isolados cultivados. Este método requer um sistema de cultivo estabelecido que permita o cultivo de protoplastos e a regeneração da planta a partir dos protoplastos cultivados. O segundo método é a transformação das células ou tecidos com *Agrobacterium*. Esse método requer (a) que as células ou tecidos vegetais possam ser transformados por *Agrobacterium* e (b) que as células ou tecidos transformados possam ser induzidos a regenerar em plantas inteiras. O terceiro método é a transformação das sementes, ápices ou meristemas com *Agrobacterium*. Esse método requer micropropagação.

A eficiência de transformação por *Agrobacterium* pode ser intensificada através da utilização de um número de métodos conhecidos na técnica. Por exemplo, tem sido demonstrado que a inclusão de uma molécula natural de resposta a ferimento tal como acetosiringona (AS) a cultura de *Agrobacterium* intensifica a eficiência de transformação com *Agrobacterium tumefaciens* (Shahla et al. (1987) *Plant Molec. Biol.* 8: 291-298). Alternativamente, a eficiência de transformação pode ser aumentada ferindo-se o tecido alvo a ser transformado. O ferimento de um tecido de planta pode ser alcançado, por exemplo, esmurrando-se, macerando-se, bombardeando-se com microprojéteis, etc. Ver, por exemplo, Bidney et al. (1992) *Plant Molec. Biol.* 18: 301-313.

Ainda em outros avanços, as células vegetais são transfectadas com vetores via bombardeamento de partículas (ou seja, com uma arma genética). Métodos de transferência de genes mediados por partículas são conhecidos na técnica, 5 são disponíveis comercialmente, e incluem, mas não se limitam a, o instrumento de entrega de gene dirigido por gás descrito na Patente U.S. No. 5.584.807, todos os conteúdos da qual são aqui incorporados por referência em 10 sua íntegra. Este método envolve revestir a seqüência de polinucleotídeos de interesse com partículas de metal pesado, e acelerar as partículas revestidas sob a pressão de gás comprimido para a distribuição ao tecido alvo.

15 Outros métodos de bombardeamento de partículas também são disponíveis para a introdução de seqüências de polinucleotídeos heterólogas em células vegetais. Geralmente, estes métodos envolvem depositar a seqüência do polinucleotídeo de interesse sobre a superfície de 20 partículas pequenas, densas e de um material tal como ouro, platina, ou tungstênio. As partículas revestidas são então revestidas sobre uma superfície ainda mais rígida, tal como uma chapa de metal, ou sobre uma folha transportadora feita de um material frágil tal como Mylar. A folha revestida é 25 então acelerada rumo ao tecido biológico alvo. O uso da folha lisa gera um espalhamento uniforme das partículas aceleradas que maximiza o número de células recebendo partículas sob condições uniformes, resultando na

introdução da amostra de polinucleotídeo dentro do tecido alvo.

Sinais específicos de iniciação também podem ser
5 usados para se alcançar uma maior eficiência de tradução de
seqüências codificando o polipeptídeo de interesse. Tais
sinais incluem o códon de iniciação ATG e seqüências
adjacentes. Em casos onde as seqüências codificando o
polipeptídeo de interesse, seu códon de iniciação, e
10 seqüências acima sejam inseridas em um vetor de expressão
apropriado, nenhum sinal controle adicional para
transcrição ou tradução pode ser necessário. Entretanto, em
casos onde apenas a seqüência codificadora, ou uma porção
dessa, é inserida, sinais controle de tradução exógenos
15 incluindo o códon de iniciação ATG devem ser fornecidos.
Ademais, o códon de iniciação deve estar posicionado na
fase de leitura correta para garantir a tradução do inserto
inteiro. Elementos de tradução exógenos e códons de
iniciação podem ser de várias origens, tanto natural como
20 sintético. A eficiência da expressão pode ser reforçada
através da inclusão de intensificadores que sejam
apropriados para o sistema celular particularmente
utilizado, tal como estes descritos na literatura (Scharf
et al. (1994) *Results Probl Cell Differ.* 20 : 125).

25

Células que tenham sido transformadas com um polinucleotídeo codificando um domínio de polipeptídeo da invenção podem ser crescidas em plantas de acordo com caminhos convencionais. Ver, por exemplo, McCormick *et al.*

(1986) *Plant Cell Rep.* 5: 81-84. Essas plantas podem então ser crescidas, e polinizadas tanto pela mesma linhagem transformada ou por diferentes linhagens, tendo o híbrido resultante a expressão constitutiva do fenótipo desejado

5 caracteristicamente identificado. Duas ou mais gerações podem ser crescidas para garantir que a expressão do fenótipo característico desejado seja mantida e herdada de forma estável e, então, as sementes são colhidas para garantir que a expressão do fenótipo característico

10 desejado tenha sido alcançada. Dessa maneira, a presente invenção fornece sementes transformadas (também referidas como "sementes transgênicas") tendo um polinucleotídeo codificando um domínio de polipeptídeo da invenção, por exemplo, um cassete de expressão da invenção, incorporado

15 de forma estável dentro do genoma dessas.

C. Avaliação da Transformação da Planta

Após a introdução do DNA em células vegetais, a

20 transformação ou integração do polinucleotídeo no genoma vegetal é confirmada por vários métodos, tais como análise de polinucleotídeos, polipeptídeos e metabólitos associados com a seqüência integrada.

25 Análise de PCR é um método rápido de se rastrear células, tecidos ou brotos para a presença do gene incorporado no estágio anterior, antes do transplante para o solo (Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press,

Cold Spring Harbor, NY)). O PCR é efetuado usando-se *primers* de oligonucleotídeos específicos para o nucleotídeo de interesse ou no *background* do vetor de *Agrobacterium*, etc.

5

A introdução de DNA pode ser confirmada por análise de *Southern blot* do DNA genômico (Sambrook and Russell (2001) supra). Em geral, o DNA total é extraído da célula ou organismo, digerido com enzimas de restrição apropriadas, fracionado em um gel de agarose e transferido para uma membrana de náilon ou nitrocelulose. A membrana ou "*blot*" é então sondada com, por exemplo, fragmento de DNA rotulado radioativamente com ^{32}P para confirmação da integração do DNA introduzido no genoma da planta de acordo com técnicas padronizadas (Sambrook and Russell (2001) supra).

Na análise *Northern*, o RNA é isolado de tecidos específicos de uma célula ou organismo, fracionado em um gel de agarose com formaldeído e transferido para um filtro de náilon de acordo com procedimentos padronizados que são usados rotineiramente na técnica (Sambrook and Russell (2001) supra). A expressão de um RNA codificado pelo polinucleotídeo da presente invenção é então testada por hibridizando o filtro a uma sonda radioativa derivada da seqüência de interesse por métodos conhecidos na técnica (Sambrook and Russell (2001) supra).

Western blot, testes bioquímicos e similares podem ser conduzidos nas plantas transgênicas para se determinar a

presença de um polipeptídeo(s) codificado pelo polinucleotídeo(s) de interesse através de procedimentos padrão (Sambrook and Russell (2001) supra) usando-se anticorpos que se liguem a um ou mais epítomos presentes no polipeptídeo resistente a herbicida.

D. Métodos para controlar seletivamente as ervas daninhas em um campo de cultivo

Métodos para se controlar seletivamente as ervas daninhas em um campo contendo uma planta também são providos. Em um avanço, as sementes da planta ou as plantas são resistentes aos inibidores de glutamina sintetase herbicidas como resultado da inserção de um polinucleotídeo da presente invenção dentro da raiz da planta ou da planta. Em métodos específicos, a planta é tratada com uma concentração efetiva de um herbicida, onde a aplicação de herbicidas resulta em um controle seletivo das ervas daninhas ou outras plantas não transformadas. Por "concentração efetiva" pretende-se a concentração que controle o crescimento ou espalhamento das ervas daninhas ou outra planta não transformada sem afetar significativamente as plantas resistentes a herbicidas ou as raízes dessas plantas. Então, a quantidade pode ser pequena o suficiente para simplesmente retardar ou suprimir o crescimento ou desenvolvimento, ou a quantidade pode ser grande o suficiente para destruir irreversivelmente as plantas sensíveis. Tais concentrações efetivas para os herbicidas de interesse são geralmente conhecidas na

técnica. O herbicida pode ser aplicado tanto pré- ou pós emergência de acordo com técnicas usuais para aplicação do herbicida nos campos compreendendo as plantas ou sementes de plantas que tenham se tornado resistentes ao herbicida.

5

Os exemplos a seguir são fornecidos a título de ilustrações e não por meio de limitação.

EXEMPLOS EXPERIMENTAIS

10

Exemplo 1. Linhagens Resistentes a Glufosinato.

A linhagem bacteriana resistente a glufosinato (ATX 20345) foi isolada de uma amostra de solo através de uma suspensão de solo colocando-se aproximadamente 0,01 gramas de solo em 500 μ l de água destilada. A suspensão de solo foi agitada, e 20 μ l foram usados para inocular um volume de 2,5 ml de meio mínimo de cultura suplementado com 5 mM de glufosinato (Riedel-de Haen, disponível através de Sigma- Aldrich, St. Louis, MO). O meio mínimo contém os seguintes ingredientes (por 1 litro): 10 gramas de sacarose, 1 ml de 0,8M $MgSO_4$, 1 ml de 0,1M $CaCl_2$, 1 ml de metais traços, 2,38 gramas de KH_2PO_4 , 5,64 gramas de K_2HPO_4 . O pH é ajustado para 7,0, e a solução é esterilizada com filtros de 0,2 μ m. Os metais traços consistem de (por 100 ml) 0,1 g de $FeSO_4 \cdot 7H_2O$, 0,5 mg de $CuSO_4 \cdot 5H_2O$, 1,0 mg de H_3BO_3 , 1,0 mg de $MnSO_4 \cdot 5H_2O$, 7,0 mg de $ZnS(7H_2O)$, 1,0 mg de MoO_3 , 4,0 g de KCl. Nenhuma fonte adicional de nitrogênio foi fornecida no meio. As culturas foram crescidas a 21°C

num agitador rotativo por 3 dias e então transferidas para um meio mínimo fresco contendo 5 mM de glufosinato e incubadas a 21°C. Após 2 dias, as culturas foram usadas para inocular o meio mínimo fresco com 5 mM de glufosinato.

5 Após 2 dias, as culturas foram plaqueadas em Luria Bertani (LB) ágar e então re-plaqueadas para isolamento. ATX 20345 foi selecionada pela sua habilidade de crescer na presença de 5 mM de glufosinato. Colônias em LB ágar são róseo-avermelhadas, elevadas, circulares, e com 1-2 mm em

10 diâmetro. ATX20345 foi tipada por análise de ácido-graxo (conforme conhecido na técnica), e determinada como sendo uma linhagem de *Serratia marcescens*.

ATX 20345 foi capaz de crescer até altas densidades ópticas na presença de glufosinato sob as condições

15 descritas acima, apesar de esta ser incapaz de crescer sob as mesmas condições na ausência de glufosinato. Estes dados sugerem que o glufosinato está fornecendo nitrogênio para a bactéria, seja na forma de uma clivagem do glufosinato para liberar uma forma usável de nitrogênio ou pela

20 suplementação de amônia no complexo glufosinato-amônia fornecido pelo suplemento. A substituição do glufosinato por cloreto de amônia como uma fonte de nitrogênio, entretanto, permite crescimento, sugerindo que a amônia é uma fonte de nitrogênio adequada para esta linhagem. As

25 concentrações altas (5-50 mM), ATX 20345 cresce bem até 50 mM e, de fato, parece crescer melhor em concentrações crescentes de glufosinato de 5 mM até 50 mM de glufosinato. Crescimentos acima de 50 mM não foram testados.

Exemplo 2. Clonagem da glutamina sintase de ATX 20345

Para se obter o gene(s) responsável pela resistência da linhagem à glufosinato, uma biblioteca genômica de plasmídeos com pequenos insertos foi preparada a partir da
5 linhagem ATX 20345. O DNA genômico foi extraído de uma cultura fresca, crescida durante a noite em LB usando-se SDS, lise celular com proteinase K seguida de extrações com CTAB/NaCl, fenol-clorofórmio. Digestões parciais foram
10 realizadas no DNA genômico usando-se 0,1 unidades de Sau3A I a 37°C por 30 minutos seguidos da adição de EDTA e incubação a 65°C por 20 minutos para travar a reação. O DNA resultante tinha aproximativamente entre 4- 12kb de tamanho. O DNA foi purificado em gel, tratado com T4
15 polinucleotídeo quinase, e então ligado no vetor pUC18 digerido com BamH I. Ligações foram transformadas em DH5 α e plaqueadas diretamente em placas do meio mínimo M63 contendo 20mM de glufosinato, 100 μ g/ml de carbenicilina, e 0,1 mM de IPTG. Após vários dias, 4 colônias apareceram,
20 tentativamente denominadas TKH1- 4. Elas foram crescidas e o DNA plasmidial foi isolado e analisado por digestão com EcoR I+ Hind III, e Pst I/ Sac I para se comparar os insertos. Apesar de todos os quatro clones conterem insertos, TKH2 e TKH3 foram idênticos pela análise da
25 digestão com enzimas de restrição, e ambos continham insertos de aproximadamente 4,9 kb.

Para se determinar se algum dos quatro clones codificava a glutamina sintetase, todos foram transformados

para a linhagem celular M5004 *glnA*⁻ (*E. coli* Genetic Stock Center No. 5531, Mayer (1975) *Mol. Gen. Genet.* 137: 131-142), juntamente com o vetor controle pUC 18, e plaqueadas em LB/carbenicilina. As colônias resultantes para cada transformação foram plaqueadas em placas do meio mínimo M63 com ou sem a adição de glutamina. Apenas os clones contendo uma glutamina sintetase funcional deveriam crescer na ausência de glutamina. TKH2 e TKH3 recuperaram o fenótipo *glnA*⁻, enquanto que TKH1 e TKH4 não (Tabela 1). A habilidade de TKH2 e TKH3 em recuperar completamente o fenótipo *glnA*⁻ demonstrou que cada uma delas contém uma glutamina sintetase funcional.

Tabela 1. Complementação do fenótipo *glnA*⁻ por THK2 e THK3

	Crescimento em células <i>glnA</i>	
	Glutamina adicionada	Sem Glutamina
TKH1	+++	-
TKH2	+++	+++
TKH3	+++	+++
TKH4	+++	-

15

Para se reconfirmar que TKH2 e TKH3 conferem resistência a 20 mM de glufosinato, os plasmídeos purificados foram re-transformados para células DH5 α . pUC18 foi re-transformado como controle negativo. As misturas da transformação foram plaqueadas diretamente em placas de M63 mínimo com ou sem 20mM de glufosinato. Enquanto que pUC18,

20

TKH2 e TKH3 cresceram na ausência de glufosinato, apenas TKH2 e TKH3 cresceram na presença de glufosinato.

O seqüenciamento do DNA plasmidial de TKH2 e TKH3 com os iniciadores M1 3 *Forward* e M1 3 *Reverse*, e subsequente análise das seqüências, revelou que TKH2 e TKH3 são clones idênticos. Baseado na complementação, análise de seqüência, e demonstração da habilidade para conferir resistência à glufosinato em *E. coli*, nós concluimos que TKH2 e TKH3 codificam uma glutamina sintetase idêntica.

10

Exemplo 3. Seqüência da glutamina sintetase ags1

A seqüência de DNA do clone TKH foi determinada (aqui referida como ags1), e um quadro de iniciação de leitura com homologia para a família de glutamina sintetase foi identificado. Esta fase aberta de leitura foi amplificada por PCR usando-se polimerase de alta fidelidade (*high fidelity*), e clonando-se em pUC19 para produzir pAX685. A proteína codificada (AGS1) demonstra alta identidade de aminoácidos com glutamina sintases de bactérias GRAM negativas, incluindo *E. coli* (identidade de aminoácido de 90%), *Erwinia* (94%), *Pantoea* (93%) e *Yersinia* (96%).

20

Exemplo 4. Mutagênese de glutamina sintetase ags1 selvagem.

25

Mutantes do gene para glutamina sintetase bacteriana ags1 (SEQ ID NO:1) foram criados por mutagênese dirigida por erro usando-se o kit GENEMORPH® Random Mutagenesis, e também usando-se mutagênese sítio direcionada como

conhecido na técnica. Entretanto, muitos métodos são disponíveis para a criação de bibliotecas de mutantes. Os mutantes resultantes foram clonados dentro de um vetor pUC19, eletroporados tanto para células de *E. coli* XL-I ou DH5 alfa, e selecionadas para crescimento em meio M63+ ágar contendo antibióticos e 2, 10, ou 20 mM de glufosinato. Vinte e dois clones foram identificados como capazes de crescer em 20mM de glufosinato, e escolhidos para análises posteriores.

10

As seqüências dos clones resistentes a glufosinato foram determinadas. Após análise de seqüência, os clones 1 e 15; clones 2 e 3; clones 5 e 6; clones 9, 10, e 12; e os clones 16, 17, 19 e 20 foram verificados como sendo idênticos. Então, os clones 1, 3, 5, 7, 9, 12, 16, 19, e 20 não foram mais analisados.

Os clones 2, 10, 15, 17, e 21 demonstraram-se capazes de crescer no meio mínimo M63+ ágar com 2, 10, 20, 50, e 100 mM de glufosinato após eletroporação para células de *E. coli* DH5 α .

As seqüências de DNA dos clones resistentes a glufosinato foram traduzidas, e as seqüências da proteína resultante foram alinhadas com a seqüência da proteína AGS1 selvagem. Foram constatadas substituições de aminoácidos correspondentes às posições dos aminoácidos da AGS1 selvagem (SEQ ID NO:2) nestes clones resistentes à glufosinato. Estas substituições são ilustradas na Figura 1

e incluem S2T, V39I, S54A, G56A, A72V, F80S, F81S, E82D, D102M, V125M, V150M, A151T, D166N, G168C, P185S, V207I, H212N, V214M, V214A, V214E, G218S, V222M, D264V, S276Y, G289S, I303N, R345S, K395R, A420V, R447C. De particular
 5 interesse são as mutações em torno do ácido glutâmico na posição 213, que é o sítio catalítico da glutamina sintetase. Outro grupo interessante de mutações ocorre em torno do aminoácido 150.

10 Tabela 2. Clones Resistentes à Glufosinato Derivados de *agsl*

Gene	pAX#	Nucleotídeos SEQ ID NO:	Aminoácidos SEQ ID NO:	Designação do clone Original	Crescimento em 20mM de glufosinato	Crescimento em 100mM de glufosinato
<i>agsl</i> (w.1)	pAX685	1	2	N/A	.-	-
<i>agslm1</i>	pAX3421	3	4	Clone#2	+++	+++
<i>agslm2</i>	pAX3422	5	6	Pick#6	+++	+++
<i>agslm3</i>	pAX3427	7	8	Clone#4	+++	NT
<i>agslm4</i>	pAX3428	9	10	Clone#6	+++	NT
<i>agslm5</i>	pAX3430	11	12	Clone#8	+++	NT
<i>agslm6</i>	pAX3431	13	14	Clone#10	+++	+++
<i>agslm7</i>	pAX3432	15	16	Clone#11	+++	NT
<i>agslm8</i>	pAX3433	17	18	Clone#13	+++	NT
<i>agslm9</i>	pAX3434	19	20	Clone#14	+++	NT
<i>agslm10</i>	pAX3435	21	22	Clone#15	+++	+++
<i>agslm11</i>	pAX3436	23	24	Clone#17	+++	+++
<i>agslm12</i>	pAX3437	25	26	Clone#18	+++	NT
<i>agslm13</i>	pAX3438	27	28	Clone#21	+++	+++
<i>agslm14</i>	pAX3426	29	30	Clone#22	+++	NT

ags1m15	pAX3439	31	32	N/A	+++	+++
---------	---------	----	----	-----	-----	-----

NT= não testado

Exemplo 5. Complementação de um mutante glutamina sintase

5 Clones contendo agslm1 (pAX3421) e agslm2 (pAX3422) foram selecionados para novos trabalhos. Ambos os clones demonstraram-se capazes de complementar uma *E. coli* mutante em agsl com agslm1 crescendo em uma taxa maior do que agslm2.

10

Exemplo 6. Cinética de glutamina sintases resistentes a glufosinato

agsl1 (pAX3421) e agslm2 (pAX3422) foram subclonadas dentro do vetor de expressão em *E. coli* pRSFIB (Invitrogen) de forma a criar uma região N-terminal codificando uma marca (*tag*) de 6Xhis, purificado e cineticamente caracterizado. AGS1m2 ('pick6') teve atividade enzimática relativamente pequena em um ensaio de 5 minutos quando comparado com a AGS 1 tipo selvagem, mas pareceu chegar à atividade completa durante a noite. AGS1m1 ('pick 2') foi indistinguível da enzima selvagem na ausência de glufosinato, mas demonstrou atividade na presença de 100µM de glufosinato, uma concentração que inibe completamente a enzima glutamina sintetase tipo selvagem agsl.

Exemplo 7. Mutagênese de agslm2

agslm2 foi mutagenizado usando mutagênese de propensão a erro (error-prone) usando-se o kit GENEMORPH® II Random Mutagenesis kit (Stratagene) de acordo com as instruções do fabricante. O produto de PCR mutagenizado foi digerido com 5 Sac I e Hind III, e ligado em um vetor pUC, similarmente digerido com Sac I e Hind III. As ligações foram transformadas para células de *E. coli*, e plaqueadas em placas de M63+ contendo antibiótico, e 125mM de glufosinato. Os clones crescendo em placas com 125mM de 10 glufosinato foram re-testados em placas com 200mM de glufosinato, e comparados com plaqueamentos similares de agslm1 e agslm2. Enquanto que as células expressando AGS1m1 e AGS1m2 não cresceram em 200mM de glufosinato, um único clone, designado pAX3439, foi isolado por força de suas 15 habilidades para crescer em placas com 200mM de glufosinato. A seqüência de DNA da fase de leitura de ags em pAX3439 foi seqüenciada, e o gene designado como agslm16. A seqüência de DNA de agslm16 é representada aqui como SEQ ID NO: 31, e a seqüência de aminoácido é 20 representada aqui como SEQ ID NO: 32. AGS1m16 difere de AGS1m2 em um único aminoácido na posição H212 da proteína, que é modificado de Histidina ('H') para Asparagina ('N'), e contem um total de cinco mudanças de aminoácidos em relação à AGS1 tipo selvagem (ver Figura 1 e Tabela 3).

Tabela 3. Variantes de *ags1*

Gene	Mudanças de aminoácidos na proteína codificada em relação a AGS1
<i>ags1</i> (tipo-selvagem)	
<i>ags1m1</i>	V39I, V214M
<i>ags1m2</i>	F81S, P185S, G218S, I301N
<i>ags1m3</i>	E82D, V214A
<i>ags1m4</i>	A151T, V214M, S276Y
<i>ags1m6</i>	V222M, D264V, R345S,
<i>ags1m7</i>	G168C, V214M, K395R
<i>ags1m8</i>	R345S, R447C
<i>ags1m9</i>	S2T, A72V
<i>ags1m10</i>	V150M
<i>ags1m11</i>	G56A, V214E
<i>ags1m12</i>	V207I, V214M
<i>ags1m13</i>	D102N, V125M, V214M
<i>ags1m14</i>	V150M, D166N, V214M, G289S, A420V
<i>ags1m15</i>	S54A
<i>ags1m16</i>	F81S, P185S, H212N, G218S, I303N
<i>ags1m17</i>	F81S, P185S, H212T, V214A, G218S, I303N
<i>ags1m18</i>	F81S, P185S, H212T, V214S, G218S, I303N
<i>ags1m19</i>	F81S, P185S, H212S, V214A, G218S, I303N
<i>ags1m20</i>	F81S, P185S, H212M, V214H, G218S, I303N
<i>ags1m21</i>	N160S, G167R, V214M

AGS1m16 confere a maior resistência em *E. coli*. Células contendo AGS1m16 são capazes de crescer em concentrações de glufosinato de até 200mM. Colônias de células contendo AGS1m2, AGS1m11, e AGS1m4 crescem mais rapidamente em 50mM de glifosato do que outras variantes, exceto AGS1m16.

Exemplo 8. Variantes agslm17, agslm18, agslm19, agslm20, e agslm21

Baseado no conhecimento do mecanismo da reação de GS conhecido na técnica, e no alinhamento de AGS1 com outras enzimas GS, pode-se prever a localização do centro de reação de GS em AGS1 e variantes, agsl (m16) foi mutagenizada na região da proteína sugerida como sendo o centro de reação de GS, e foram identificadas diversas variantes que conferiram crescimento melhorado em placas com 225 mM de glufosinato mediante as células hospedeiras de E. coli. agslm17 (SEQ ID NO: 33) codifica a proteína AGS1m17 (SEQ ID NO:34). agslm18 (SEQ ID NO: 35) codifica a proteína AGS1m18 (SEQ ID NO: 36). agslm19 (SEQ ID NO: 37) codifica a proteína AGS1m19 (SEQ ID NO:38). agslm20 (SEQ ID NO: 39) codifica a proteína AGS1m20 (SEQ ID NO: 40). Verificou-se que os clones expressando AGS1m17, AGS1m18, AGS1m19, ou AGS1m20 foram todos capazes de crescer em placas contendo 375mM de glufosinato, enquanto que nenhum crescimento de clones expressando AGS1m16 foi observado em placas contendo 375mM de glufosinato.

agslm21 (SEQ ID NO: 41) codifica a proteína AGS1m21 (SEQ ID NO: 42). AGS1m21 é uma variante de AGS1 que contém uma modificação de aminoácido similar à identificada em outras variantes (V214M), assim como duas novas mutações (N160S e G167R). Esta enzima foi expressa em E. coli, purificada, e os valores cinéticos K_m (glutamato) e K_i (glufosinato) da variante foram medidos por ensaios

enzimáticos. Verificou-se que a enzima mutagenizada possui resistência aumentada à glufosinato, como mostrado na Tabela a seguir.

5 Tabela 4. Cinética de AGS1m21

Enzima	Km, mM	Ki, μ M
GlnA	2,7	nd
AGS1	3,4	13,0
AGS1m21	12,0	2000

Exemplo 9. Remoção de sítios de adenilação de AGS1 e variantes

10 É bem conhecido na técnica (Mehta et al. (2004) *J. Biol. Chem.* 279: 22477- 22482, aqui incorporado pela referência em sua íntegra) que em uma célula bacteriana, enzimas GS bacterianas são frequentemente sujeitas a regulação negativa por adenilação de um resíduo de tirosina
 15 em particular.

ags1(ad-) (SEQ ID NO: 43) codificando a proteína AGS1(AD-) (SEQ ID NO: 44), é uma variante de ags1 na qual o sítio de adenilação hipotético tenha sido removido por
 20 mutação da tirosina na posição 398 da AGS 1 para uma fenilalanina através de mutagênese sítio direcionada.

ags1m17(ad-) (SEQ ID NO: 45) codificando a proteína AGS1mH(AD-) (SEQ ID NO: 46), é uma variante de ags1m17 na
 25 qual o sítio de adenililação putativo foi removido por

mutação da tirosina na posição 398 da AGS1 para uma fenilalanina através de mutagênese sítio dirigida.

A enzima inalterada (AGS1) foi comparada com AGS1(AD-) e AGS1m17(AD-) através da realização de ensaios enzimáticos em ambas as enzimas seguido de purificação a partir de *E. coli*. A enzima glutamina sintetase de *E. coli* (GlnA) também foi testada. Os valores cinéticos obtidos para cada enzima são demonstrados na Tabela 5 abaixo.

10

Tabela 5. Cinética de AGS1 (AD-)

Enzima	K_m , mM	K_i , μM	V_{max} nmol/min/ μg	K_{cat} , seg^{-1}	$(K_{\text{cat}} * K_i)$ / K_m	MW (kD)
GlnA (<i>E. coli</i>)	2,7	ND	0,24	0,21	-	52
AGS1	3,4	13,0	0,10	0,08	0,32	52
AGS1 (AD-)	3,4	13,0	3,8	3,29	12,59	52
AGS1m17(AD-)	14,1	23,000	0,0078	0,07	110,04	

Exemplo 10. Clones resistentes a glufosinato são mutados perto do sítio ativo.

15

Um número surpreendente de variantes, e todos dentre os clones mais resistentes, continham mutações no sítio ativo da sintetase. Os dois ácidos glutâmicos no centro catalítico (correspondendo às posições 213 e 221 de AGS1, SEQ ID NO:2) são chave para a atividade da glutamina sintetase. A valina na posição 214 é variada em 13 das

variantes, e pode ser mutada para metionina (M), alanina (A), serina (S), histidina (H) e ácido glutâmico (E). Três variantes adicionais, incluindo AGS1m16, têm resíduos modificados nessa região.

5

Exemplo 11 Identificação de novas enzimas glutamina sintetase que são resistentes à glutamina sintetase herbicida.

10 Usando-se os métodos da invenção, pode-se identificar outras glutamina sintetases resistentes a herbicida pesquisando-se as bases de dados contendo enzimas glutamina sintetase, e/ou alinhando-se as seqüências de aminoácido das enzimas glutamina sintetase e analisando-se para ao
15 menos uma substituição de aminoácido dentro das posições correspondentes as posições 125 até 175 da SEQ ID NO: 2 ou entre as posições 200 e 250 da SEQ ID NO: 2. É sabido que alguma modificação destas regiões é tolerada na natureza sem perturbar a natureza da resistência a herbicida
20 conferida por estas regiões, e são, portanto equivalentes às seqüências listadas aqui. Portanto, é reconhecido que as enzimas que possuem cerca de 80%, cerca de 85%, cerca de 90%, cerca de 95%, 96%, 97%, 98% ou 99% de homologia à polipeptídeos da invenção podem conferir resistência a
25 inibidores de glutamina sintetase herbicidas.

Exemplo 12. Transformação de planta por bombardeamento de partículas.

Espigas de milho são mais bem colhidas entre 8-12 dias após a polinização. Os embriões são isolados a partir das espigas, e aqueles embriões de 0,8-1,5 mm de tamanho são usados preferencialmente na transformação. Os embriões são
5 plaqueados com o lado do escutelo para cima em um meio de incubação adequado, tal como o meio DN62A5S (3,98 g/L de Sais N6; 1 ml/L (de estoque 1000x) de 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 sacarose;
10 1 ml/L (de estoque 1 mg/ml) 2,4-D). Entretanto, outros meios e sais que não DN62A5S são adequados e são conhecidos na técnica. Os embriões são incubados durante a noite a 25°C no escuro. Entretanto, não é necessário por si, incubar os embriões durante a noite.

15

Os explantes resultantes são transferidos para malha quadrada (30-40 por placa), transferidos sobre meio osmótico por cerca de 30-45 minutos, e então transferidos para uma placa de bombardeamento ver, por exemplo,
20 Publicação do PCT No. WO/0138514 e Patente U.S. No. 5.240.842).

As construções de DNA projetadas para expressar enzimas glutamina sintetase da presente invenção em células
25 vegetais são aceleradas para dentro do tecido da planta usando-se um acelerador de feixe de aerosol, utilizando-se as condições essencialmente como descrito na Publicação do PCT No. WO/0138514. Após o bombardeamento os embriões são incubados por cerca de 30 min em meio osmótico, e colocados

sobre o meio de incubação durante a noite a 25°C no escuro. Para se evitar explantes indevidamente danificados, eles são incubados por ao menos 24 horas antes da transferência para o meio de recuperação. Os embriões são então
5 espalhados sobre meio de período de recuperação, por cerca de 5 dias, 25°C no escuro, e então transferidos para um meio de seleção. Os explantes são incubados em meio de seleção por até oito semanas, dependendo da natureza e das características da seleção particularmente utilizada. Após
10 o período de seleção, o calo resultante é transferido para o meio de maturação do embrião até que se observe a formação dos embriões somáticos maduros. Os embriões somáticos maduros resultantes são então colocados sob luz baixa, e o processo de regeneração é iniciado por métodos
15 conhecidos na técnica. Os brotos resultantes são deixados para enraizar no meio de enraizamento, e as plantas resultantes são transferidas para potes de viveiro e propagadas como plantas transgênicas. As plantas são analisadas para resistência melhorada para inibidores de
20 glutamina sintetase herbicidas.

Materiais

Meio DN62A5S

Componentes	Por litro	Fonte
Mistura básica de sais N6 de Chu (Prod. No C416)	3,98 g/l	Phytotechnology Labs
Solução de vitaminas N6 de Chu (Prod. No C149)	1 ml/l (de estoque 1000x)	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
Sacarose	50 g/l	Phytotechnology Labs
2,4-D (No. Prod. D-7299)	1mL/L (de solução estoque a 1mg/mL)	Sigma

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

Exemplo 13. Transformação de células vegetais por
 10 Transformação mediada por *Agrobacterium*

As espigas são mais bem colhidas entre 8-12 dias após a polinização. Os embriões são isolados de espigas, e aqueles embriões com 0,8-1,5 mm de tamanho são preferidos para uso na transformação. Os embriões são plaqueados com o

lado do escutelo para cima em um meio de incubação adequado, e incubados durante a noite a 25°C no escuro. Entretanto, não é necessário por si se incubar os embriões durante a noite. Os embriões são postos em contato com uma
5 linhagem de *Agrobacterium* contendo os vetores apropriados tendo uma seqüência da presente invenção para transferência mediada por plasmídeo Ti por cerca de 5-10 min, e então plaqueados sobre meio de co-cultivo por cerca de 3 dias (25°C no escuro). Após co-cultivo, os explantes são
10 transferidos para um meio de período de recuperação por cerca de cinco dias (a 25°C no escuro). Os explantes são incubados no meio seletivo por até oito semanas, dependendo da natureza e características da seleção utilizada em particular. Após o período de seleção, o calo resultante é
15 transferido para o meio de maturação de embrião, até se observar a formação de embriões somáticos maduros. Os embriões somáticos maduros resultantes são então colocados sob luz baixa, e o processo de regeneração é iniciado conforme conhecido na técnica. Os brotos resultantes são
20 deixados para enraizar no meio de enraizamento, e as plantas resultantes são transferidas para potes de viveiro e propagadas como plantas transgênicas.

Exemplo 14. Plantas transgênicas expressando AGS1m16

25

synags1m16 (SEQ ID NO:47) é uma seqüência de DNA alternada que codifica a proteína AGS1m16 (SEQ ID NO:32). *synags1m16* foi clonada dentro de um vetor de lançamento para guiar a super-expressão de AGS1m16 em milho. O vetor

dispõe a super-expressão de *synagslm16* sob o controle do promotor Trp5 (US application 11/377.318, preenchida em 16 de março de 2006 e aqui incorporada por referência em sua íntegra).

5

Nove plantas de milho transgênicas contendo *agslm16* foram geradas. A expressão da proteína AGSIm16 foi confirmada por *Western blot* para cada um destes eventos. Como controles, seis eventos que não continham *agslm16* foram gerados. A eficiência de utilização de nitrogênio foi avaliada para estes eventos através da determinação do conteúdo protéico das amostras das folhas isoladas das plantas do T₀ após quatro semanas de crescimento na estufa.

15

As proteínas presentes nas amostras de folhas foram quantificadas como segue: Cinquenta miligramas de material de folha (peso fresco, sem nervura central) foram liofilizados para determinação do peso seco. O tecido de folha desidratado foi então macerado na presença de água Milli Q fresca usando-se um MiniBeadbeater-96™ e esferas de aço inoxidável de 2.3 mm. O tecido de folha macerado foi filtrado através de um filtro de Fluoreto de Polivinilideno de 0.45 μ m (PVDF). O corante Bio-Rad Protein Dye foi adicionado às amostras de folha diluídas na água, e o método de Bradford foi realizado e lido no espectrofotômetro no comprimento de 595nm contra marcadores de proteína internos incluídos no método.

20

25

As concentrações de proteína solúvel foram divididas pelo peso seco de cada amostra para se obter a massa protéica por unidade de peso seco. Surpreendentemente, demonstrou-se que as plantas contendo agslm16 possuíam uma média 24% maior de conteúdo protéico do que as plantas controle que não continham agslm16.

Tabela 6. Proteína aumentada em plantas com AGS1m16

Evento	Genótipo	Proteína Total
1	agslm16	28,1
2	agslm16	8,0
3	agslm16	13,2
4	agslm16	18,4
5	agslm16	9,8
6	agslm16	13,9
7	agslm16	28,6
8	agslm16	10,9
9	agslm16	20,5
Média		16,8
C1	Controle	16,0
C2	Controle	14,1
C3	Controle	18,7
C4	Controle	10,1
C5	Controle	13,0
C6	Controle	9,1
Média		13,5

Todas as publicações e aplicações de patentes mencionadas na especificação são indicativas do nível de conhecimento daqueles especialistas na técnica a qual esta invenção pertence. Todas as publicações e aplicações de patentes são aqui incorporadas por referência na mesma extensão como se cada publicação ou aplicação de patente individual fosse especificamente e individualmente indicada para serem incorporadas por referência.

10 Apesar de a invenção precedente ter sido descrita em algum detalhe por meio de ilustração e exemplo para fins de clareza de entendimento, será óbvio que certas alterações e modificações podem ser praticadas no escopo das reivindicações em anexo.

15

LISTAGEM DE SEQUÊNCIAS

<110> Depositante: Athenix Corporation
 Nicholas B. Duck
 Todd K. Hinson
 Vadim Beilinson
 Laura Cooper Schouten
 Daniel John Tomso

<120> Título da Invenção: Glutamina Sintetases Bacterianas e Métodos de Uso

<130> Referência do documento: 45600/329367

<150> Número do Documento de Prioridade: 60/812.000
 <151> Data de Depósito do Documento de Prioridade: 08-06-2006

<160> Quantidade de SEQ ID NOs.: 47

<170> Software: FastSEQ for Windows Version 4.0

<210> SEQ ID NO: 1
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Serratia marcescens

<220> Características:
 <221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 1

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	

cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa gtg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	

370	375	380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg			1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu			
385	390	395	400
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac			1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp			
	405	410	415
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc			1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly			
	420	425	430
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa			1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys			
	435	440	445
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa			1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu			
	450	455	460
ctg tac tac agc gtc taa			1410
Leu Tyr Tyr Ser Val *			
465			

<210> SEQ ID NO: 2

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Serratia marcescens

<400> Sequência: 2

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

Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 3

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Seqüência Artificial

<220> Características:

<223> Outras Informações: agslml

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Seqüência: 3

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag ata aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Ile Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	

ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly 85 90 95	288
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala 100 105 110	336
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly 115 120 125	384
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser 130 135 140	432
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn 145 150 155 160	480
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val 165 170 175	528
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu 180 185 190	576
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu 195 200 205	624
gcg cac cac cac gaa atg gcg acc gcc ggt cag aac gaa gtg gca acc Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr 210 215 220	672
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys 225 230 235 240	720
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe 245 250 255	768
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His 260 265 270	816
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr 275 280 285	864
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys 290 295 300	912
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr 305 310 315 320	960
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser 1008	1008


```

      115                               120                               125
Pro  Glu  Pro  Glu  Phe  Phe  Leu  Phe  Asp  Asp  Ile  Arg  Phe  Gly  Ser  Ser
130
Ile  Arg  Gly  Ser  His  Val  Ala  Ile  Asp  Asp  Ile  Glu  Gly  Ala  Trp  Asn
145
Ser  Gly  Thr  Lys  Tyr  Asp  Gly  Gly  Asn  Lys  Gly  His  Arg  Pro  Ala  Val
165
Lys  Gly  Gly  Tyr  Phe  Pro  Val  Pro  Pro  Val  Asp  Ser  Ser  Gln  Asp  Leu
180
Arg  Ser  Thr  Met  Cys  Leu  Thr  Met  Glu  Glu  Met  Gly  Leu  Val  Val  Glu
195
Ala  His  His  His  Glu  Met  Ala  Thr  Ala  Gly  Gln  Asn  Glu  Val  Ala  Thr
210
Arg  Phe  Asn  Thr  Met  Thr  Lys  Lys  Ala  Asp  Glu  Ile  Gln  Ile  Tyr  Lys
225
Tyr  Val  Val  His  Asn  Val  Ala  His  Ala  Phe  Gly  Lys  Thr  Ala  Thr  Phe
245
Met  Pro  Lys  Pro  Met  Phe  Gly  Asp  Asn  Gly  Ser  Gly  Met  His  Cys  His
260
Met  Ser  Leu  Ser  Lys  Asn  Gly  Thr  Asn  Leu  Phe  Ala  Gly  Asp  Lys  Tyr
275
Gly  Gly  Leu  Ser  Glu  Thr  Ala  Leu  Phe  Tyr  Ile  Gly  Gly  Ile  Ile  Lys
290
His  Ala  Lys  Ala  Ile  Asn  Ala  Leu  Ala  Asn  Pro  Thr  Thr  Asn  Ser  Tyr
305
Lys  Arg  Leu  Val  Pro  Gly  Tyr  Glu  Ala  Pro  Val  Met  Leu  Ala  Tyr  Ser
325
Ala  Arg  Asn  Arg  Ser  Ala  Ser  Ile  Arg  Ile  Pro  Val  Val  Ala  Ser  Pro
340
Lys  Ala  Arg  Arg  Ile  Glu  Ala  Arg  Phe  Pro  Asp  Pro  Ala  Ala  Asn  Pro
355
Tyr  Leu  Cys  Phe  Ala  Ala  Leu  Leu  Met  Ala  Gly  Leu  Asp  Gly  Ile  Ile
370
Asn  Lys  Ile  His  Pro  Gly  Asp  Ala  Met  Asp  Lys  Asn  Leu  Tyr  Asp  Leu
385
Pro  Pro  Glu  Glu  Glu  Ala  Glu  Ile  Pro  Lys  Val  Ala  Gly  Ser  Leu  Asp
405
Glu  Ala  Met  Ala  Ala  Leu  Asn  Glu  Asp  Arg  Glu  Phe  Leu  Thr  Arg  Gly
420
Gly  Val  Phe  Thr  Asp  Asp  Ala  Ile  Asp  Ala  Tyr  Ile  Glu  Leu  Arg  Lys
435
Glu  Glu  Met  Asp  Arg  Val  Arg  Met  Thr  Pro  His  Pro  Val  Glu  Phe  Glu
450
Leu  Tyr  Tyr  Ser  Val
465

```

<210> SEQ ID NO: 5
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm2

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 5
 atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
tcc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca tcg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa gtg gcg acc gcc agt cag aac gaa gtg gca acc	672
Ala His His His Glu Val Ala Thr Ala Ser Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gca cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	

atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc aac aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	
370 375 380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg	1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu	
385 390 395 400	
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac	1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp	
405 410 415	
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc	1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly	
420 425 430	
ggc gtg ttc act gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa	1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys	
435 440 445	
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa	1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu	
450 455 460	
ctg tac tac agc gtc taa	1410
Leu Tyr Tyr Ser Val *	
465	

<210> SEQ ID NO: 6

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M2

<400> Sequência: 6

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

<211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm3

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 7

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

gcg cac cac cac gaa gcg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Ala Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac gcg tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	
370 375 380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg	1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu	
385 390 395 400	
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac	1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp	
405 410 415	
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc	1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly	
420 425 430	
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa	1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys	
435 440 445	
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa	1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu	
450 455 460	

ctg tac tac agc gtc taa
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 8
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: AGS1M3

<400> Sequência: 8
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80
 Phe Asp Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125
 Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
 130 135 140
 Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
 145 150 155 160
 Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
 165 170 175
 Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu
 180 185 190
 Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
 195 200 205
 Ala His His His Glu Ala Ala Thr Ala Gly Gln Asn Glu Val Ala Thr
 210 215 220
 Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
 225 230 235 240
 Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380

Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 9

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: agslm4

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Sequência: 9

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa 48
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg 96
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa 144
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45

atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct 192
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60

gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc 240
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80

ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc 288
 Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95

acc atg cag ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc 336
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110

gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg 384
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125

cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc 432
 Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
 130 135 140

atc cgc ggt tcc cac gtg acg atc gac gat atc gaa ggc gcc tgg aac 480


```

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
                               405                               410                               415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
                               420                               425                               430

ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
                               435                               440                               445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
                               450                               455                               460

ctg tac tac agc gtc taa 1410
Leu Tyr Tyr Ser Val *
465

```

```

<210> SEQ ID NO: 10
<211> Comprimento: 469
<212> Tipo: PRT
<213> Organismo: Sequência Artificial

```

```

<220> Características:
<223> Outras Informações: AGS1M4

```

```

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

```

			260						265				270			
Met	Ser	Leu	Tyr	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr	
		275					280					285				
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Ile	Lys	
		290				295					300					
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr	
305					310				315						320	
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser	
				325					330					335		
Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Arg	Ile	Pro	Val	Val	Ala	Ser	Pro	
				340				345					350			
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro	
		355					360					365				
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile	
	370					375					380					
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu	
385					390					395					400	
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp	
				405				410						415		
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly	
				420				425					430			
Gly	Val	Phe	Thr	Asp	Asp	Ala	Ile	Asp	Ala	Tyr	Ile	Glu	Leu	Arg	Lys	
		435				440						445				
Glu	Glu	Met	Asp	Arg	Val	Arg	Met	Thr	Pro	His	Pro	Val	Glu	Phe	Glu	
	450					455					460					
Leu	Tyr	Tyr	Ser	Val												
465																

<210> SEQ ID NO: 11

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: agslm6

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Sequência: 11

atg	tcc	gct	gaa	cac	gtt	ttg	acg	atg	ctg	aat	gag	cat	gaa	gtg	aaa	48
Met	Ser	Ala	Glu	His	Val	Leu	Thr	Met	Leu	Asn	Glu	His	Glu	Val	Lys	
1				5					10					15		
ttc	gta	gac	ctg	cgt	ttc	act	gac	acc	aag	ggt	aag	gaa	cag	cac	gtg	96
Phe	Val	Asp	Leu	Arg	Phe	Thr	Asp	Thr	Lys	Gly	Lys	Glu	Gln	His	Val	
			20					25					30			
act	atc	ccg	gct	cac	cag	gta	aac	gcc	gac	ttc	ttc	gaa	gaa	ggt	aaa	144
Thr	Ile	Pro	Ala	His	Gln	Val	Asn	Ala	Asp	Phe	Phe	Glu	Glu	Gly	Lys	
			35				40					45				
atg	ttt	gac	ggc	tcc	tct	atc	ggt	ggt	tgg	aag	ggc	atc	aac	gaa	tct	192
Met	Phe	Asp	Gly	Ser	Ser	Ile	Gly	Gly	Trp	Lys	Gly	Ile	Asn	Glu	Ser	
	50					55					60					
gac	atg	gtg	ctg	atg	ccg	gac	gcc	agc	acg	gcg	gtt	ctg	gat	ccg	ttc	240
Asp	Met	Val	Leu	Met	Pro	Asp	Ala	Ser	Thr	Ala	Val	Leu	Asp	Pro	Phe	
	65				70				75					80		
ttc	gaa	gaa	cct	acg	ctg	atc	att	cgc	tgt	gac	att	ctc	gag	ccg	ggc	288
Phe	Glu	Glu	Pro	Thr	Leu	Ile	Ile	Arg	Cys	Asp	Ile	Leu	Glu	Pro	Gly	

					85						90						95	
acc	atg	caa	ggc	tac	gat	cgc	gac	ccg	cgt	tcc	atc	tcc	aaa	cgc	gcc	336		
Thr	Met	Gln	Gly	Tyr	Asp	Arg	Asp	Pro	Arg	Ser	Ile	Ser	Lys	Arg	Ala			
					100						105						110	
gaa	gac	ttc	ctg	cgc	tcc	tcc	ggc	atc	gcg	gac	acc	gtg	ctg	ttc	ggg	384		
Glu	Asp	Phe	Leu	Arg	Ser	Ser	Gly	Ile	Ala	Asp	Thr	Val	Leu	Phe	Gly			
					115						120						125	
cca	gag	cct	gag	ttc	ttc	ctg	ttc	gac	gac	atc	cgc	ttc	ggc	agc	agc	432		
Pro	Glu	Pro	Glu	Phe	Phe	Leu	Phe	Asp	Asp	Ile	Arg	Phe	Gly	Ser	Ser			
					130						135						140	
atc	cgc	ggg	tcc	cac	gtg	gcg	atc	gac	gat	atc	gaa	ggc	gcc	tgg	aac	480		
Ile	Arg	Gly	Ser	His	Val	Ala	Ile	Asp	Asp	Ile	Glu	Gly	Ala	Trp	Asn			
					145						150						155	
tcc	ggc	aca	aaa	tac	gac	ggc	ggc	aac	aaa	ggc	cac	cgt	ccg	gcg	gtg	528		
Ser	Gly	Thr	Lys	Tyr	Asp	Gly	Gly	Asn	Lys	Gly	His	Arg	Pro	Ala	Val			
					165						170						175	
aaa	ggc	ggg	tac	ttc	ccg	ggt	cca	ccg	gtc	gac	tct	tcg	cag	gat	ctg	576		
Lys	Gly	Gly	Tyr	Phe	Pro	Val	Pro	Pro	Val	Asp	Ser	Ser	Gln	Asp	Leu			
					180						185						190	
cgt	tcc	acc	atg	tgt	ctg	acc	atg	gaa	gag	atg	ggc	ctg	gtg	ggt	gaa	624		
Arg	Ser	Thr	Met	Cys	Leu	Thr	Met	Glu	Glu	Met	Gly	Leu	Val	Val	Glu			
					195						200						205	
gcg	cac	cac	cac	gaa	gtg	gcg	acc	gcc	ggg	cag	aac	gaa	atg	gca	acc	672		
Ala	His	His	His	Glu	Val	Ala	Thr	Ala	Gly	Gln	Asn	Glu	Met	Ala	Thr			
					210						215						220	
cgc	ttc	aac	acc	atg	acc	aag	aaa	gcc	gac	gaa	att	cag	atc	tat	aag	720		
Arg	Phe	Asn	Thr	Met	Thr	Lys	Lys	Ala	Asp	Glu	Ile	Gln	Ile	Tyr	Lys			
					225						230						235	
tac	gtg	gtg	cac	aac	gtg	gcg	cac	gcc	ttc	ggg	aaa	acc	gcg	acc	ttc	768		
Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe			
					245						250						255	
atg	ccg	aag	ccc	atg	ttc	ggc	gtc	aac	ggg	tcc	ggc	atg	cac	tgc	cac	816		
Met	Pro	Lys	Pro	Met	Phe	Gly	Val	Asn	Gly	Ser	Gly	Met	His	Cys	His			
					260						265						270	
atg	tcg	ctg	tcc	aag	aac	ggc	acc	aac	ctg	ttc	gcc	ggc	gac	aaa	tac	864		
Met	Ser	Leu	Ser	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr			
					275						280						285	
ggc	ggc	ctg	tct	gaa	acc	gca	ctg	ttc	tac	atc	ggc	ggg	atc	atc	aag	912		
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Ile	Lys			
					290						295						300	
cac	gcc	aag	gcg	atc	aac	gcg	ctg	gcc	aac	cca	acc	acc	aac	tcg	tac	960		
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr			
					305						310						315	
aaa	cgt	ctg	gtg	cca	ggc	tac	gaa	gcg	ccg	gtg	atg	ctg	gct	tac	tcc	1008		
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser			
					325						330						335	
gcc	cgt	aac	cgc	tcc	gcg	tcc	atc	agt	atc	ccg	gtg	gtc	gcc	agc	ccg	1056		

Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Ser	Ile	Pro	Val	Val	Ala	Ser	Pro		
			340					345					350				
aaa	gcg	cgc	cgc	atc	gaa	gcc	cgc	ttc	ccg	gat	ccg	gcg	gct	aac	cca	1104	
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro		
		355					360					365					
tac	ctg	tgc	ttc	gcc	gca	ctg	ctg	atg	gcc	ggc	ctg	gac	ggc	atc	atc	1152	
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile		
	370					375					380						
aac	aag	atc	cac	cct	ggc	gac	gcc	atg	gac	aaa	aac	ctg	tac	gac	ctg	1200	
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu		
385					390					395					400		
ccg	ccg	gaa	gaa	gaa	gcc	gag	atc	cca	aaa	gtg	gcc	ggc	tcg	ctg	gac	1248	
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp		
				405				410						415			
gag	gcg	atg	gcc	gcg	ctg	aac	gaa	gac	cgc	gag	ttc	ctg	acc	cgc	ggc	1296	
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly		
			420				425						430				
ggc	gtg	ttc	acc	gac	gat	gcg	atc	gat	gcc	tac	atc	gaa	ctg	cgc	aaa	1344	
Gly	Val	Phe	Thr	Asp	Asp	Ala	Ile	Asp	Ala	Tyr	Ile	Glu	Leu	Arg	Lys		
		435				440						445					
gaa	gag	atg	gac	cgc	gtt	cgc	atg	acg	cca	cac	ccg	gtc	gag	ttc	gaa	1392	
Glu	Glu	Met	Asp	Arg	Val	Arg	Met	Thr	Pro	His	Pro	Val	Glu	Phe	Glu		
	450					455					460						
ctg	tac	tac	agc	gtc	taa											1410	
Leu	Tyr	Tyr	Ser	Val	*												
465																	

<210> SEQ ID NO: 12

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M6

<400> Sequência: 12

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

Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
 145 150 155 160
 Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
 165 170 175
 Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu
 180 185 190
 Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
 195 200 205
 Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Met Ala Thr
 210 215 220
 Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
 225 230 235 240
 Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Val Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Ser Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 13
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm7

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 13
 atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa 48
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg 96
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggt agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cat gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc tgc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Cys Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa atg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	

275	280	285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys 290 295 300			912
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr 305 310 315 320			960
aaa cgt ctg gtg cca ggc tac gaa gcg cct gtg atg ctg gct tac tcc Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser 325 330 335			1008
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro 340 345 350			1056
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro 355 360 365			1104
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile 370 375 380			1152
aac aag atc cac cct ggc gac gcc atg gac aga aac ctg tac gac ctg Asn Lys Ile His Pro Gly Asp Ala Met Asp Arg Asn Leu Tyr Asp Leu 385 390 395 400			1200
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp 405 410 415			1248
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly 420 425 430			1296
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys 435 440 445			1344
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu 450 455 460			1392
ctg tac tac agc gtc taa Leu Tyr Tyr Ser Val *			1410
465			

<210> SEQ ID NO: 14

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M7

<400> Sequência: 14

Met	Ser	Ala	Glu	His	Val	Leu	Thr	Met	Leu	Asn	Glu	His	Glu	Val	Lys
1				5					10					15	
Phe	Val	Asp	Leu	Arg	Phe	Thr	Asp	Thr	Lys	Gly	Lys	Glu	Gln	His	Val

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

<210> SEQ ID NO: 15

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: agslm8

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Sequência: 15

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

cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc agt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Ser Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	
370 375 380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg	1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu	
385 390 395 400	
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac	1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp	
405 410 415	
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc	1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly	
420 425 430	
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg tgc aaa	1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Cys Lys	
435 440 445	
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa	1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu	
450 455 460	
ctg tac tac agc gtc taa	1410
Leu Tyr Tyr Ser Val *	

465

<210> SEQ ID NO: 16
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Seqüência Artificial

<220> Características:
 <223> Outras Informações: AGS1M8

<400> Seqüência: 16

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

```

                405                410                415
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
                420                425                430
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Cys Lys
                435                440                445
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
                450                455                460
Leu Tyr Tyr Ser Val
465

```

<210> SEQ ID NO: 17
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm9

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 17

```

atg acc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa      48
Met Thr Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
  1                5                10                15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg      96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
                20                25                30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa      144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
                35                40                45

atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct      192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
  50                55                60

gac atg gtg ctg atg ccg gac gtc agc acg gcg gtt ctg gat ccg ttc      240
Asp Met Val Leu Met Pro Asp Val Ser Thr Ala Val Leu Asp Pro Phe
  65                70                75                80

ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc      288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
                85                90                95

acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc      336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
                100                105                110

gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg      384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
                115                120                125

cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc      432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
                130                135                140

atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac      480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
  145                150                155                160

```

tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa gtg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc gcc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc gcc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc gcc ctg tct gaa acc gca ctg ttc tac atc gcc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc gcc ctg gac gcc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	
370 375 380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg	1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu	
385 390 395 400	
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc gcc tcg ctg gac	1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp	
405 410 415	

Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 19

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Seqüência Artificial

<220> Características:

<223> Outras Informações: agslm10

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Seqüência: 19

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336

Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100	105 110
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115	120 125
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130	135 140
atc cgc ggt tcc cac atg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Met Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145	150 155 160
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165	170 175
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180	185 190
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195	200 205
gcg cac cac cac gaa gtg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210	215 220
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225	230 235 240
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245	250 255
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260	265 270
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275	280 285
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290	295 300
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305	310 315 320
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325	330 335
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340	345 350

aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365

tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tat gac ctg 1200
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430

ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460

ctg tac tac agc gtc taa 1410
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 20
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Seqüência Artificial

<220> Características:
 <223> Outras Informações: AGS1M10

<400> Seqüência: 20
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80
 Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125
 Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
 130 135 140
 Ile Arg Gly Ser His Met Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
 145 150 155 160
 Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val

atg ttt gac ggc tcc tct atc gct ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Ala Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa gag gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Glu Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc gcc cta tct gaa acc gca ctg ttc tac atc gcc ggt atc att aag	912

Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Ile	Lys	
	290					295					300					
cac	gcc	aag	gcg	atc	aac	gcg	ctg	gcc	aac	ccg	acc	acc	aac	tcg	tac	960
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr	
305					310				315						320	
aaa	cgt	ctg	gtg	cca	ggc	tac	gaa	gcg	ccg	gtg	atg	ctg	gct	tac	tcc	1008
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser	
				325					330					335		
gcc	cgt	aac	cgc	tcc	gcg	tcc	atc	cgt	atc	ccg	gtg	gtc	gcc	agc	ccg	1056
Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Arg	Ile	Pro	Val	Val	Ala	Ser	Pro	
			340					345					350			
aaa	gcg	cgc	cgc	atc	gaa	gcc	cgc	ttc	ccg	gat	ccg	gcg	gct	aac	cca	1104
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro	
		355					360					365				
tac	ctg	tgc	ttc	gcc	gca	ctg	ctg	atg	gcc	ggc	ctg	gac	ggc	atc	atc	1152
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile	
	370					375				380						
aac	aag	atc	cac	cct	ggc	gac	gcc	atg	gat	aaa	aac	ctg	tac	gac	ctg	1200
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu	
385					390				395						400	
ccg	ccg	gaa	gaa	gaa	gcc	gag	atc	cca	aaa	gtg	gcc	ggc	tcg	ctg	gac	1248
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp	
				405				410						415		
gag	gcg	atg	gcc	gcg	ctg	aac	gaa	gac	cgc	gag	ttc	ctg	acc	cgc	ggc	1296
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly	
			420				425						430			
ggc	gtg	ttc	acc	gac	gat	gcg	atc	gat	gcc	tac	atc	gaa	ctg	cgc	aaa	1344
Gly	Val	Phe	Thr	Asp	Asp	Ala	Ile	Asp	Ala	Tyr	Ile	Glu	Leu	Arg	Lys	
		435				440						445				
gaa	gag	atg	gac	cgc	gtt	cgc	atg	acg	cca	cat	ccg	gtc	gag	ttc	gaa	1392
Glu	Glu	Met	Asp	Arg	Val	Arg	Met	Thr	Pro	His	Pro	Val	Glu	Phe	Glu	
	450					455					460					
ctg	tac	tac	agc	gtc	taa											1410
Leu	Tyr	Tyr	Ser	Val	*											
465																

<210> SEQ ID NO: 22

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M11

<400> Sequência: 22

Met	Ser	Ala	Glu	His	Val	Leu	Thr	Met	Leu	Asn	Glu	His	Glu	Val	Lys	
1				5				10						15		
Phe	Val	Asp	Leu	Arg	Phe	Thr	Asp	Thr	Lys	Gly	Lys	Glu	Gln	His	Val	
		20					25						30			
Thr	Ile	Pro	Ala	His	Gln	Val	Asn	Ala	Asp	Phe	Phe	Glu	Glu	Gly	Lys	
		35				40						45				

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

<210> SEQ ID NO: 23

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: agslm12

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Sequência: 23

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gat ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg att gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Ile Glu	
195 200 205	
gcg cac cac cac gaa atg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	

225		230		235		240	
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc							768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe							
		245		250		255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac							816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His							
		260		265		270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac							864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr							
		275		280		285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag							912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys							
		290		295		300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac							960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr							
		305		310		315	320
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc							1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser							
		325		330		335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg							1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro							
		340		345		350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca							1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro							
		355		360		365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc							1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile							
		370		375		380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg							1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu							
		385		390		395	400
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac							1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp							
		405		410		415	
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc							1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly							
		420		425		430	
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa							1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys							
		435		440		445	
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa							1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu							
		450		455		460	
ctg tac tac agc gtc taa							1410
Leu Tyr Tyr Ser Val *							
465							

<210> SEQ ID NO: 24
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: AGS1M12

<400> Sequência: 24

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

Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 25
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslml3

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 25

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

aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cggtaccatgtgtctgaccatggaa gag atgggcctggtggttgaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac cac gaa atg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210 215 220	
cgcttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tacgtggtgcac aacgtggcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atgtcgtctgtcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	
370 375 380	
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg	1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu	
385 390 395 400	
ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac	1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp	
405 410 415	
gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc gcc	1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly	

420	425	430	
ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgt aaa			1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys			
435	440	445	
gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa			1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu			
450	455	460	
ctg tac tac agc gtc taa			1410
Leu Tyr Tyr Ser Val *			
465			

<210> SEQ ID NO: 26
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Seqüência Artificial

<220> Características:
 <223> Outras Informações: AGS1M13

<400> Seqüência: 26

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

```

305          310          315          320
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
          325          330          335
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
          340          345          350
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
          355          360          365
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
          370          375          380
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
385          390          395          400
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
          405          410          415
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
          420          425          430
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
          435          440          445
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
          450          455          460
Leu Tyr Tyr Ser Val
465

```

```

<210> SEQ ID NO: 27
<211> Comprimento: 1410
<212> Tipo: DNA
<213> Organismo: Sequência Artificial

```

```

<220> Características:
<223> Outras Informações: agslm14

```

```

<221> Nome/Chave: CDS
<222> Localização: (1)...(1410)

```

```

<400> Sequência: 27
atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa      48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
  1          5          10          15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg      96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
          20          25          30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa      144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
          35          40          45

atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct      192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
          50          55          60

gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc      240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
          65          70          75          80

ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc      288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
          85          90          95

acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc      336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
          100          105          110

```

gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115	120
125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130	135
140	140
atc cgc ggt tcc cac atg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Met Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145	150
155	160
tcc ggc aca aaa tac aac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asn Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165	170
175	
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu	
180	185
190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195	200
205	
gcg cac cac cac gaa atg gcg acc gcc ggt cag aac gaa gtg gca acc	672
Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr	
210	215
220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225	230
235	240
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245	250
255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260	265
270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275	280
285	
agc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag	912
Ser Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys	
290	295
300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305	310
315	320
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325	330
335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340	345
350	
aaa gcg cgc cgc atc gaa gcc cgc ttc cct gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355	360
365	

tat ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
370 375 380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg 1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
405 410 415

gag gcg atg gtc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
Glu Ala Met Val Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
420 425 430

ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
450 455 460

ctg tac tac agc gtc taa 1410
Leu Tyr Tyr Ser Val *

465

<210> SEQ ID NO: 28

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M14

<400> Sequência: 28

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

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

<210> SEQ ID NO: 29

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Seqüência Artificial

<220> Características:

<223> Outras Informações: agslm15

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Seqüência: 29

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc gct atc ggt ggt tgg aag ggc atc aac gaa tct	192

Met	Phe	Asp	Gly	Ser	Ala	Ile	Gly	Gly	Trp	Lys	Gly	Ile	Asn	Glu	Ser		
	50					55					60						
gac	atg	gtg	ctg	atg	ccg	gac	gcc	agc	acg	gcg	gtt	ctg	gat	ccg	ttc		240
Asp	Met	Val	Leu	Met		Asp	Ala	Ser	Thr		Ala	Val	Leu	Asp	Pro	Phe	
	65				70				75						80		
ttc	gaa	gaa	cct	acg	ctg	atc	att	cgc	tgt	gac	att	ctc	gag	ccg	ggc		288
Phe	Glu	Glu	Pro	Thr	Leu	Ile	Ile	Arg	Cys	Asp	Ile	Leu	Glu	Pro	Gly		
				85				90						95			
acc	atg	caa	ggc	tac	gat	cgc	gac	ccg	cgt	tcc	atc	tcc	aaa	cgc	gcc		336
Thr	Met	Gln	Gly	Tyr	Asp	Arg	Asp	Pro	Arg	Ser	Ile	Ser	Lys	Arg	Ala		
			100					105					110				
gaa	gat	ttc	ctg	cgc	tcc	tcc	ggc	atc	gcg	gac	acc	gtg	ctg	ttc	ggg		384
Glu	Asp	Phe	Leu	Arg	Ser	Ser	Gly	Ile	Ala	Asp	Thr	Val	Leu	Phe	Gly		
		115					120					125					
cca	gag	cct	gag	ttc	ttc	ctg	ttc	gac	gac	atc	cgc	ttc	ggc	agc	agc		432
Pro	Glu	Pro	Glu	Phe	Phe	Leu	Phe	Asp	Asp	Ile	Arg	Phe	Gly	Ser	Ser		
	130					135					140						
atc	cgc	ggt	tcc	cac	gta	gcg	atc	gac	gat	atc	gaa	ggc	gcc	tgg	aac		480
Ile	Arg	Gly	Ser	His	Val	Ala	Ile	Asp	Asp	Ile	Glu	Gly	Ala	Trp	Asn		
	145			150				155							160		
tcc	ggc	aca	aaa	tac	gac	ggc	ggc	aac	aaa	ggc	cac	cgt	ccg	gcg	gtg		528
Ser	Gly	Thr	Lys	Tyr	Asp	Gly	Gly	Asn	Lys	Gly	His	Arg	Pro	Ala	Val		
				165				170						175			
aaa	ggc	ggt	tac	ttc	ccg	gtt	cca	ccg	gtc	gac	tct	tcg	cag	gat	ctg		576
Lys	Gly	Gly	Tyr	Phe	Pro	Val	Pro	Pro	Val	Asp	Ser	Ser	Gln	Asp	Leu		
			180				185						190				
cgt	tcc	acc	atg	tgt	ctg	acc	atg	gaa	gag	atg	ggc	ctg	gtg	gtt	gaa		624
Arg	Ser	Thr	Met	Cys	Leu	Thr	Met	Glu	Glu	Met	Gly	Leu	Val	Val	Glu		
		195					200					205					
gca	cac	cac	cac	gaa	gtg	gcg	acc	gcc	ggt	cag	aac	gaa	gtg	gca	acc		672
Ala	His	His	His	Glu	Val	Ala	Thr	Ala	Gly	Gln	Asn	Glu	Val	Ala	Thr		
	210					215					220						
cgc	ttc	aac	acc	atg	acc	aag	aaa	gcc	gac	gaa	att	cag	atc	tat	aag		720
Arg	Phe	Asn	Thr	Met	Thr	Lys	Lys	Ala	Asp	Glu	Ile	Gln	Ile	Tyr	Lys		
	225				230				235						240		
tac	gtg	gtg	cac	aac	gtg	gcg	cac	gcc	ttc	ggt	aaa	acc	gcg	acc	ttc		768
Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe		
				245				250						255			
atg	ccg	aag	ccc	atg	ttc	ggc	gac	aac	ggt	tcc	ggc	atg	cac	tgc	cac		816
Met	Pro	Lys	Pro	Met	Phe	Gly	Asp	Asn	Gly	Ser	Gly	Met	His	Cys	His		
			260					265					270				
atg	tcg	ctg	tcc	aag	aac	ggc	acc	aac	ctg	ttc	gcc	ggc	gac	aaa	tac		864
Met	Ser	Leu	Ser	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr		
		275				280						285					
ggc	ggc	ctg	tct	gaa	acc	gca	ctg	ttc	tac	atc	ggc	ggt	atc	atc	aag		912
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Ile	Lys		
		290				295					300						

cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac 960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
305 310 315 320

aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc 1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
325 330 335

gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg 1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
340 345 350

aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
355 360 365

tac ctg tgc ttc gcc gca ctg cta atg gcc ggc ctg gac ggc atc atc 1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
370 375 380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg 1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
405 410 415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
420 425 430

ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
450 455 460

ctg tac tac agc gtc taa 1410
Leu Tyr Tyr Ser Val *
465

<210> SEQ ID NO: 30

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Seqüência Artificial

<220> Características:

<223> Outras Informações: AGS1M15

<400> Seqüência: 30

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

65					70					75				80	
Phe	Glu	Glu	Pro	Thr	Leu	Ile	Ile	Arg	Cys	Asp	Ile	Leu	Glu	Pro	Gly
				85					90					95	
Thr	Met	Gln	Gly	Tyr	Asp	Arg	Asp	Pro	Arg	Ser	Ile	Ser	Lys	Arg	Ala
			100					105						110	
Glu	Asp	Phe	Leu	Arg	Ser	Ser	Gly	Ile	Ala	Asp	Thr	Val	Leu	Phe	Gly
		115					120					125			
Pro	Glu	Pro	Glu	Phe	Phe	Leu	Phe	Asp	Asp	Ile	Arg	Phe	Gly	Ser	Ser
	130					135					140				
Ile	Arg	Gly	Ser	His	Val	Ala	Ile	Asp	Asp	Ile	Glu	Gly	Ala	Trp	Asn
145					150					155					160
Ser	Gly	Thr	Lys	Tyr	Asp	Gly	Gly	Asn	Lys	Gly	His	Arg	Pro	Ala	Val
				165					170						175
Lys	Gly	Gly	Tyr	Phe	Pro	Val	Pro	Pro	Val	Asp	Ser	Ser	Gln	Asp	Leu
			180						185					190	
Arg	Ser	Thr	Met	Cys	Leu	Thr	Met	Glu	Glu	Met	Gly	Leu	Val	Val	Glu
		195					200					205			
Ala	His	His	His	Glu	Val	Ala	Thr	Ala	Gly	Gln	Asn	Glu	Val	Ala	Thr
	210					215					220				
Arg	Phe	Asn	Thr	Met	Thr	Lys	Lys	Ala	Asp	Glu	Ile	Gln	Ile	Tyr	Lys
225					230					235					240
Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe
				245					250						255
Met	Pro	Lys	Pro	Met	Phe	Gly	Asp	Asn	Gly	Ser	Gly	Met	His	Cys	His
			260					265					270		
Met	Ser	Leu	Ser	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr
		275					280					285			
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Ile	Lys
	290					295					300				
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr
305					310					315					320
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser
				325					330						335
Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Arg	Ile	Pro	Val	Val	Ala	Ser	Pro
			340					345						350	
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro
		355					360					365			
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile
	370					375					380				
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu
385					390					395					400
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp
				405					410						415
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly
			420					425					430		
Gly	Val	Phe	Thr	Asp	Asp	Ala	Ile	Asp	Ala	Tyr	Ile	Glu	Leu	Arg	Lys
		435					440					445			
Glu	Glu	Met	Asp	Arg	Val	Arg	Met	Thr	Pro	His	Pro	Val	Glu	Phe	Glu
	450					455					460				
Leu	Tyr	Tyr	Ser	Val											
465															

<210> SEQ ID NO: 31

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: agslm16

<221> Nome/Chave: CDS

<222> Localização: (1)...(1410)

<400> Sequência: 31

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa	48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys	
1 5 10 15	
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg	96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val	
20 25 30	
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa	144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys	
35 40 45	
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct	192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser	
50 55 60	
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
tcc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca tcg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac aac gaa gtg gcg acc gcc agt cag aac gaa gtg gca acc	672
Ala His His Asn Glu Val Ala Thr Ala Ser Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gca cac gcc ttc ggt aaa acc gcg acc ttc	768

Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe	
				245					250					255		
atg	ccg	aag	ccc	atg	ttc	ggc	gac	aac	ggc	tcc	ggc	atg	cac	tgc	cac	816
Met	Pro	Lys	Pro	Met	Phe	Gly	Asp	Asn	Gly	Ser	Gly	Met	His	Cys	His	
			260					265					270			
atg	tcg	ctg	tcc	aag	aac	ggc	acc	aac	ctg	ttc	gcc	ggc	gac	aaa	tac	864
Met	Ser	Leu	Ser	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr	
		275					280					285				
ggc	ggc	ctg	tct	gaa	acc	gca	ctg	ttc	tac	atc	ggc	ggc	atc	aac	aag	912
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Asn	Lys	
	290					295					300					
cac	gcc	aag	gcg	atc	aac	gcg	ctg	gcc	aac	ccg	acc	acc	aac	tcg	tac	960
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr	
305				310					315						320	
aaa	cgt	ctg	gtg	cca	ggc	tac	gaa	gcg	ccg	gtg	atg	ctg	gct	tac	tcc	1008
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser	
			325						330					335		
gcc	cgt	aac	cgc	tcc	gcg	tcc	atc	cgf	atc	ccg	gtg	gtc	gcc	agc	ccg	1056
Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Arg	Ile	Pro	Val	Val	Ala	Ser	Pro	
			340					345					350			
aaa	gcg	cgc	cgc	atc	gaa	gcc	cgc	ttc	ccg	gat	ccg	gcg	gct	aac	cca	1104
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro	
		355					360					365				
tac	ctg	tgc	ttc	gcc	gca	ctg	ctg	atg	gcc	ggc	ctg	gac	ggc	atc	atc	1152
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile	
	370					375					380					
aac	aag	atc	cac	cct	ggc	gac	gcc	atg	gac	aaa	aac	ctg	tac	gac	ctg	1200
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu	
385					390					395					400	
ccg	ccg	gaa	gaa	gaa	gcc	gag	atc	cca	aaa	gtg	gcc	ggc	tcg	ctg	gac	1248
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp	
			405						410					415		
gag	gcg	atg	gcc	gcg	ctg	aac	gaa	gac	cgc	gag	ttc	ctg	acc	cgc	ggc	1296
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly	
			420					425					430			
ggc	gtg	ttc	act	gac	gat	gcg	atc	gat	gcc	tac	atc	gaa	ctg	cgc	aaa	1344
Gly	Val	Phe	Thr	Asp	Asp	Ala	Ile	Asp	Ala	Tyr	Ile	Glu	Leu	Arg	Lys	
		435					440					445				
gaa	gag	atg	gac	cgc	ggt	cgc	atg	acg	cca	cac	ccg	gtc	gag	ttc	gaa	1392
Glu	Glu	Met	Asp	Arg	Val	Arg	Met	Thr	Pro	His	Pro	Val	Glu	Phe	Glu	
	450					455					460					
ctg	tac	tac	agc	gtc	taa											1410
Leu	Tyr	Tyr	Ser	Val	*											
465																

<210> SEQ ID NO: 32

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M16

<400> Sequência: 32

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

450
Leu Tyr Tyr Ser Val
465

455

460

<210> SEQ ID NO: 33
<211> Comprimento: 1410
<212> Tipo: DNA
<213> Organismo: Sequência Artificial

<220> Características:
<223> Outras Informações: agslm17

<221> Nome/Chave: CDS
<222> Localização: (1)...(1410)

<400> Sequência: 33
atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa 48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
1 5 10 15
ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg 96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
20 25 30
act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa 144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
35 40 45
atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct 192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
50 55 60
gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc 240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
65 70 75 80
tcc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc 288
Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
85 90 95
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc 336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
100 105 110
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg 384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
115 120 125
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc 432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
130 135 140
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac 480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
145 150 155 160
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg 528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
165 170 175
aaa ggc ggt tac ttc ccg gtt cca tcg gtc gac tct tcg cag gat ctg 576
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu

180					185					190						
cgt	tcc	acc	atg	tgt	ctg	acc	atg	gaa	gag	atg	ggc	ctg	gtg	gtt	gaa	624
Arg	Ser	Thr	Met	Cys	Leu	Thr	Met	Glu	Glu	Met	Gly	Leu	Val	Val	Glu	
		195					200					205				
gcg	cac	cac	acc	gag	gca	gcg	acc	gct	agc	cag	aac	gaa	gtg	gca	acc	672
Ala	His	His	Thr	Glu	Ala	Ala	Thr	Ala	Ser	Gln	Asn	Glu	Val	Ala	Thr	
	210					215				220						
cgc	ttc	aac	acc	atg	acc	aag	aaa	gcc	gac	gaa	att	cag	atc	tat	aag	720
Arg	Phe	Asn	Thr	Met	Thr	Lys	Lys	Ala	Asp	Glu	Ile	Gln	Ile	Tyr	Lys	
225				230					235						240	
tac	gtg	gtg	cac	aac	gtg	gca	cac	gcc	ttc	ggc	aaa	acc	gcg	acc	ttc	768
Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe	
			245					250						255		
atg	ccg	aag	ccc	atg	ttc	ggc	gac	aac	ggc	ttc	ggc	atg	cac	tgc	cac	816
Met	Pro	Lys	Pro	Met	Phe	Gly	Asp	Asn	Gly	Ser	Gly	Met	His	Cys	His	
			260				265						270			
atg	tcg	ctg	tcc	aag	aac	ggc	acc	aac	ctg	ttc	gcc	ggc	gac	aaa	tac	864
Met	Ser	Leu	Ser	Lys	Asn	Gly	Thr	Asn	Leu	Phe	Ala	Gly	Asp	Lys	Tyr	
		275					280					285				
ggc	ggc	ctg	tct	gaa	acc	gca	ctg	ttc	tac	atc	ggc	ggc	gac	aaa	aag	912
Gly	Gly	Leu	Ser	Glu	Thr	Ala	Leu	Phe	Tyr	Ile	Gly	Gly	Ile	Asn	Lys	
	290					295					300					
cac	gcc	aag	gcg	atc	aac	gcg	ctg	gcc	aac	ccg	acc	acc	aac	tcg	tac	960
His	Ala	Lys	Ala	Ile	Asn	Ala	Leu	Ala	Asn	Pro	Thr	Thr	Asn	Ser	Tyr	
305					310					315					320	
aaa	cgt	ctg	gtg	cca	ggc	tac	gaa	gcg	ccg	gtg	atg	ctg	gct	tac	tcc	1008
Lys	Arg	Leu	Val	Pro	Gly	Tyr	Glu	Ala	Pro	Val	Met	Leu	Ala	Tyr	Ser	
			325						330					335		
gcc	cgt	aac	cgc	tcc	gcg	tcc	atc	cgt	atc	ccg	gtg	gtc	gcc	agc	ccg	1056
Ala	Arg	Asn	Arg	Ser	Ala	Ser	Ile	Arg	Ile	Pro	Val	Val	Ala	Ser	Pro	
		340					345						350			
aaa	gcg	cgc	cgc	atc	gaa	gcc	cgc	ttc	ccg	gat	ccg	gcg	gct	aac	cca	1104
Lys	Ala	Arg	Arg	Ile	Glu	Ala	Arg	Phe	Pro	Asp	Pro	Ala	Ala	Asn	Pro	
		355					360					365				
tac	ctg	tgc	ttc	gcc	gca	ctg	ctg	atg	gcc	ggc	ctg	gac	ggc	atc	atc	1152
Tyr	Leu	Cys	Phe	Ala	Ala	Leu	Leu	Met	Ala	Gly	Leu	Asp	Gly	Ile	Ile	
	370					375					380					
aac	aag	atc	cac	cct	ggc	gac	gcc	atg	gac	aaa	aac	ctg	tac	gac	ctg	1200
Asn	Lys	Ile	His	Pro	Gly	Asp	Ala	Met	Asp	Lys	Asn	Leu	Tyr	Asp	Leu	
385					390					395					400	
ccg	ccg	gaa	gaa	gaa	gcc	gag	atc	cca	aaa	gtg	gcc	ggc	tcg	ctg	gac	1248
Pro	Pro	Glu	Glu	Glu	Ala	Glu	Ile	Pro	Lys	Val	Ala	Gly	Ser	Leu	Asp	
			405						410					415		
gag	gcg	atg	gcc	gcg	ctg	aac	gaa	gac	cgc	gag	ttc	ctg	acc	cgc	ggc	1296
Glu	Ala	Met	Ala	Ala	Leu	Asn	Glu	Asp	Arg	Glu	Phe	Leu	Thr	Arg	Gly	
		420					425						430			
ggc	gtg	ttc	act	gac	gat	gcg	atc	gat	gcc	tac	atc	gaa	ctg	cgc	aaa	1344

Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460

ctg tac tac agc gtc taa 1410
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 34
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: AGS1M17

<400> Sequência: 34
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80
 Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125
 Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
 130 135 140
 Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
 145 150 155 160
 Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
 165 170 175
 Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu
 180 185 190
 Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
 195 200 205
 Ala His His Thr Glu Ala Ala Thr Ala Ser Gln Asn Glu Val Ala Thr
 210 215 220
 Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
 225 230 235 240
 Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335

Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 35
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm18

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 35
 atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa 48
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg 96
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa 144
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct 192
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc 240
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80
 tcc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc 288
 Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95
 acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc 336
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110
 gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg 384
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125

cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca tcg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac acg gaa tcc gcg acc gct agc cag aac gaa gtg gca acc	672
Ala His His Thr Glu Ser Ala Thr Ala Ser Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gca cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc aac aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc	1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser	
325 330 335	
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg	1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro	
340 345 350	
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca	1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro	
355 360 365	
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc	1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile	


```

      210                      215                      220
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
225                      230                      235
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
                      245                      250                      255
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
                      260                      265                      270
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
                      275                      280                      285
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys
                      290                      295                      300
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
305                      310                      315
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
                      325                      330                      335
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
                      340                      345                      350
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
                      355                      360                      365
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
                      370                      375                      380
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
385                      390                      395
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
                      405                      410                      415
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
                      420                      425                      430
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
                      435                      440                      445
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
                      450                      455                      460
Leu Tyr Tyr Ser Val
465

```

<210> SEQ ID NO: 37
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm19

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 37

```

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa      48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
  1                      5                      10                      15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg      96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
                      20                      25                      30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa      144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
                      35                      40                      45

atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct      192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
  50                      55                      60

```

gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc	240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe	
65 70 75 80	
tcc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc	288
Ser Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly	
85 90 95	
acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc	336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala	
100 105 110	
gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg	384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly	
115 120 125	
cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc	432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser	
130 135 140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac	480
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn	
145 150 155 160	
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg	528
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val	
165 170 175	
aaa ggc ggt tac ttc ccg gtt cca tcg gtc gac tct tcg cag gat ctg	576
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu	
180 185 190	
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa	624
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu	
195 200 205	
gcg cac cac tcc gag gcc gcg acc gct agc cag aac gaa gtg gca acc	672
Ala His His Ser Glu Ala Ala Thr Ala Ser Gln Asn Glu Val Ala Thr	
210 215 220	
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag	720
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys	
225 230 235 240	
tac gtg gtg cac aac gtg gca cac gcc ttc ggt aaa acc gcg acc ttc	768
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe	
245 250 255	
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac	816
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His	
260 265 270	
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac	864
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr	
275 280 285	
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc aac aag	912
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys	
290 295 300	
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac	960
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr	
305 310 315 320	

aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc 1008
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335

gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg 1056
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350

aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365

tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg 1200
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430

ggc gtg ttc act gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460

ctg tac tac agc gtc taa 1410
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 38

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M19

<400> Sequência: 38

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

```

Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
      100      105
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
      115      120
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
      130      135      140
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
      145      150      155
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
      165      170
Lys Gly Gly Tyr Phe Pro Val Pro Ser Val Asp Ser Ser Gln Asp Leu
      180      185
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
      195      200
Ala His His Ser Glu Ala Ala Thr Ala Ser Gln Asn Glu Val Ala Thr
      210      215      220
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
      225      230      235
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
      245      250
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
      260      265
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
      275      280
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys
      290      295      300
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
      305      310      315
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
      325      330
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
      340      345
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
      355      360
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
      370      375      380
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
      385      390      395
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
      405      410
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
      420      425
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
      435      440      445
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
      450      455      460
Leu Tyr Tyr Ser Val
465

```

<210> SEQ ID NO: 39
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm20

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 39
 atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa

Met	Ser	Ala	Glu	His	Val	Leu	Thr	Met	Leu	Asn	Glu	His	Glu	Val	Lys	
1				5					10					15		
ttc	gta	gac	ctg	cgt	ttc	act	gac	acc	aag	ggg	aag	gaa	cag	cac	gtg	96
Phe	Val	Asp	Leu	Arg	Phe	Thr	Asp	Thr	Lys	Gly	Lys	Glu	Gln	His	Val	
			20					25					30			
act	atc	ccg	gct	cac	cag	gta	aac	gcc	gac	ttc	ttc	gaa	gaa	ggg	aaa	144
Thr	Ile	Pro	Ala	His	Gln	Val	Asn	Ala	Asp	Phe	Phe	Glu	Glu	Gly	Lys	
		35					40					45				
atg	ttt	gac	ggc	tcc	tct	atc	ggg	ggg	tgg	aag	ggc	atc	aac	gaa	tct	192
Met	Phe	Asp	Gly	Ser	Ser	Ile	Gly	Gly	Trp	Lys	Gly	Ile	Asn	Glu	Ser	
	50					55					60					
gac	atg	gtg	ctg	atg	ccg	gac	gcc	agc	acg	gcg	ggt	ctg	gat	ccg	ttc	240
Asp	Met	Val	Leu	Met	Pro	Asp	Ala	Ser	Thr	Ala	Val	Leu	Asp	Pro	Phe	
	65				70				75						80	
tcc	gaa	gaa	cct	acg	ctg	atc	att	cgc	tgt	gac	att	ctc	gag	ccg	ggc	288
Ser	Glu	Glu	Pro	Thr	Leu	Ile	Ile	Arg	Cys	Asp	Ile	Leu	Glu	Pro	Gly	
			85					90						95		
acc	atg	caa	ggc	tac	gat	cgc	gac	ccg	cgt	tcc	atc	tcc	aaa	cgc	gcc	336
Thr	Met	Gln	Gly	Tyr	Asp	Arg	Asp	Pro	Arg	Ser	Ile	Ser	Lys	Arg	Ala	
		100						105					110			
gaa	gac	ttc	ctg	cgc	tcc	tcc	ggc	atc	gcg	gac	acc	gtg	ctg	ttc	ggg	384
Glu	Asp	Phe	Leu	Arg	Ser	Ser	Gly	Ile	Ala	Asp	Thr	Val	Leu	Phe	Gly	
		115					120					125				
cca	gag	cct	gag	ttc	ttc	ctg	ttc	gac	gac	atc	cgc	ttc	ggc	agc	agc	432
Pro	Glu	Pro	Glu	Phe	Phe	Leu	Phe	Asp	Asp	Ile	Arg	Phe	Gly	Ser	Ser	
	130					135					140					
atc	cgc	ggg	tcc	cac	gtg	gcg	atc	gac	gat	atc	gaa	ggc	gcc	tgg	aac	480
Ile	Arg	Gly	Ser	His	Val	Ala	Ile	Asp	Asp	Ile	Glu	Gly	Ala	Trp	Asn	
	145				150					155					160	
tcc	ggc	aca	aaa	tac	gac	ggc	ggc	aac	aaa	ggc	cac	cgt	ccg	gcg	gtg	528
Ser	Gly	Thr	Lys	Tyr	Asp	Gly	Gly	Asn	Lys	Gly	His	Arg	Pro	Ala	Val	
			165					170						175		
aaa	ggc	ggg	tac	ttc	ccg	ggt	cca	tcg	gtc	gac	tct	tcg	cag	gat	ctg	576
Lys	Gly	Gly	Tyr	Phe	Pro	Val	Pro	Ser	Val	Asp	Ser	Ser	Gln	Asp	Leu	
			180					185					190			
cgt	tcc	acc	atg	tgt	ctg	acc	atg	gaa	gag	atg	ggc	ctg	gtg	ggt	gaa	624
Arg	Ser	Thr	Met	Cys	Leu	Thr	Met	Glu	Glu	Met	Gly	Leu	Val	Val	Glu	
		195					200					205				
gcg	cac	cac	atg	gag	cat	gcg	acc	gct	agc	cag	aac	gaa	gtg	gca	acc	672
Ala	His	His	Met	Glu	His	Ala	Thr	Ala	Ser	Gln	Asn	Glu	Val	Ala	Thr	
		210				215					220					
cgc	ttc	aac	acc	atg	acc	aag	aaa	gcc	gac	gaa	att	cag	atc	tat	aag	720
Arg	Phe	Asn	Thr	Met	Thr	Lys	Lys	Ala	Asp	Glu	Ile	Gln	Ile	Tyr	Lys	
	225				230					235					240	
tac	gtg	gtg	cac	aac	gtg	gca	cac	gcc	ttc	ggg	aaa	acc	gcg	acc	ttc	768
Tyr	Val	Val	His	Asn	Val	Ala	His	Ala	Phe	Gly	Lys	Thr	Ala	Thr	Phe	
				245					250						255	

atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac 816
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270

atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac 864
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285

ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc aac aag 912
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Asn Lys
 290 295 300

cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac 960
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320

aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc 1008
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335

gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg 1056
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350

aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365

tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg tac gac ctg 1200
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
 385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430

ggc gtg ttc act gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460

ctg tac tac agc gtc taa 1410
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 40

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M20

<400> Sequência: 40

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

<210> SEQ ID NO: 41
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm21

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 41

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

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460

ctg tac tac agc gtc taa 1410
 Leu Tyr Tyr Ser Val *
 465

<210> SEQ ID NO: 42
 <211> Comprimento: 469
 <212> Tipo: PRT
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: AGS1M21

<400> Sequência: 42
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
 65 70 75 80
 Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
 85 90 95
 Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
 100 105 110
 Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
 115 120 125
 Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
 130 135 140
 Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Ser
 145 150 155 160
 Ser Gly Thr Lys Tyr Asp Arg Gly Asn Lys Gly His Arg Pro Ala Val
 165 170 175
 Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu
 180 185 190
 Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
 195 200 205
 Ala His His His Glu Met Ala Thr Ala Gly Gln Asn Glu Val Ala Thr
 210 215 220
 Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
 225 230 235 240
 Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315 320
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro

```

          355                      360                      365
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
   370                      375                      380
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Tyr Asp Leu
385                      390                      395                      400
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
   405                      410                      415
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
   420                      425                      430
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
   435                      440                      445
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
   450                      455                      460
Leu Tyr Tyr Ser Val
465

```

<210> SEQ ID NO: 43
 <211> Comprimento: 1543
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agsl(ad-)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 43

```

atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa      48
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
   1                      5                      10                      15

ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg      96
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
   20                      25                      30

act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa      144
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
   35                      40                      45

atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct      192
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
   50                      55                      60

gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc      240
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
   65                      70                      75                      80

ttc gaa gaa cct acg ctg atc att cgc tgt gac att ctc gag ccg ggc      288
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
   85                      90                      95

acc atg caa ggc tac gat cgc gac ccg cgt tcc atc tcc aaa cgc gcc      336
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
   100                      105                      110

gaa gac ttc ctg cgc tcc tcc ggc atc gcg gac acc gtg ctg ttc ggg      384
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
   115                      120                      125

cca gag cct gag ttc ttc ctg ttc gac gac atc cgc ttc ggc agc agc      432
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser

```

130	135	140	
atc cgc ggt tcc cac gtg gcg atc gac gat atc gaa ggc gcc tgg aac 480			
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn			
145	150	155	160
tcc ggc aca aaa tac gac ggc ggc aac aaa ggc cac cgt ccg gcg gtg 528			
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val			
	165	170	175
aaa ggc ggt tac ttc ccg gtt cca ccg gtc gac tct tcg cag gat ctg 576			
Lys Gly Gly Tyr Phe Pro Val Pro Val Asp Ser Ser Gln Asp Leu			
	180	185	190
cgt tcc acc atg tgt ctg acc atg gaa gag atg ggc ctg gtg gtt gaa 624			
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu			
	195	200	205
gcg cac cac cac gaa gtg gcg acc gcc ggt cag aac gaa gtg gca acc 672			
Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Val Ala Thr			
	210	215	220
cgc ttc aac acc atg acc aag aaa gcc gac gaa att cag atc tat aag 720			
Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys			
	225	230	235
tac gtg gtg cac aac gtg gcg cac gcc ttc ggt aaa acc gcg acc ttc 768			
Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe			
	245	250	255
atg ccg aag ccc atg ttc ggc gac aac ggt tcc ggc atg cac tgc cac 816			
Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His			
	260	265	270
atg tcg ctg tcc aag aac ggc acc aac ctg ttc gcc ggc gac aaa tac 864			
Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr			
	275	280	285
ggc ggc ctg tct gaa acc gca ctg ttc tac atc ggc ggt atc atc aag 912			
Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys			
	290	295	300
cac gcc aag gcg atc aac gcg ctg gcc aac ccg acc acc aac tcg tac 960			
His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr			
	305	310	315
aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc 1008			
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser			
	325	330	335
gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg 1056			
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro			
	340	345	350
aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104			
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro			
	355	360	365
tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152			
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile			
	370	375	380
aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg ttc gac ctg 1200			

Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Phe Asp Leu
385 390 395 400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
405 410 415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
420 425 430

ggc gtg ttc acc gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
435 440 445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
450 455 460

ctg tac tac agc gtc taa gccctacccg cgccgtctgc aaaggcggac 1440
Leu Tyr Tyr Ser Val *

465

ggcgcccaca attttctgca ggtcgacaag cttgcgggccg cactcgagtc tggtaaagaa 1500
accgctgctg cgaaatttga acgcagcac atggactcgt cta 1543

<210> SEQ ID NO: 44
<211> Comprimento: 469
<212> Tipo: PRT
<213> Organismo: Sequência Artificial

<220> Características:
<223> Outras Informações: AGS1(AD-)

<400> Sequência: 44
Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
1 5 10 15
Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
20 25 30
Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
35 40 45
Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
50 55 60
Asp Met Val Leu Met Pro Asp Ala Ser Thr Ala Val Leu Asp Pro Phe
65 70 75 80
Phe Glu Glu Pro Thr Leu Ile Ile Arg Cys Asp Ile Leu Glu Pro Gly
85 90 95
Thr Met Gln Gly Tyr Asp Arg Asp Pro Arg Ser Ile Ser Lys Arg Ala
100 105 110
Glu Asp Phe Leu Arg Ser Ser Gly Ile Ala Asp Thr Val Leu Phe Gly
115 120 125
Pro Glu Pro Glu Phe Phe Leu Phe Asp Asp Ile Arg Phe Gly Ser Ser
130 135 140
Ile Arg Gly Ser His Val Ala Ile Asp Asp Ile Glu Gly Ala Trp Asn
145 150 155 160
Ser Gly Thr Lys Tyr Asp Gly Gly Asn Lys Gly His Arg Pro Ala Val
165 170 175
Lys Gly Gly Tyr Phe Pro Val Pro Pro Val Asp Ser Ser Gln Asp Leu
180 185 190
Arg Ser Thr Met Cys Leu Thr Met Glu Glu Met Gly Leu Val Val Glu
195 200 205
Ala His His His Glu Val Ala Thr Ala Gly Gln Asn Glu Val Ala Thr
210 215 220

Arg Phe Asn Thr Met Thr Lys Lys Ala Asp Glu Ile Gln Ile Tyr Lys
 225 230 235 240
 Tyr Val Val His Asn Val Ala His Ala Phe Gly Lys Thr Ala Thr Phe
 245 250 255
 Met Pro Lys Pro Met Phe Gly Asp Asn Gly Ser Gly Met His Cys His
 260 265 270
 Met Ser Leu Ser Lys Asn Gly Thr Asn Leu Phe Ala Gly Asp Lys Tyr
 275 280 285
 Gly Gly Leu Ser Glu Thr Ala Leu Phe Tyr Ile Gly Gly Ile Ile Lys
 290 295 300
 His Ala Lys Ala Ile Asn Ala Leu Ala Asn Pro Thr Thr Asn Ser Tyr
 305 310 315
 Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
 325 330 335
 Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
 340 345 350
 Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
 355 360 365
 Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
 370 375 380
 Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Phe Asp Leu
 385 390 395 400
 Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
 405 410 415
 Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
 420 425 430
 Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
 435 440 445
 Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
 450 455 460
 Leu Tyr Tyr Ser Val
 465

<210> SEQ ID NO: 45
 <211> Comprimento: 1410
 <212> Tipo: DNA
 <213> Organismo: Sequência Artificial

<220> Características:
 <223> Outras Informações: agslm17(ad-)

<221> Nome/Chave: CDS
 <222> Localização: (1)...(1410)

<400> Sequência: 45
 atg tcc gct gaa cac gtt ttg acg atg ctg aat gag cat gaa gtg aaa 48
 Met Ser Ala Glu His Val Leu Thr Met Leu Asn Glu His Glu Val Lys
 1 5 10 15
 ttc gta gac ctg cgt ttc act gac acc aag ggt aag gaa cag cac gtg 96
 Phe Val Asp Leu Arg Phe Thr Asp Thr Lys Gly Lys Glu Gln His Val
 20 25 30
 act atc ccg gct cac cag gta aac gcc gac ttc ttc gaa gaa ggt aaa 144
 Thr Ile Pro Ala His Gln Val Asn Ala Asp Phe Phe Glu Glu Gly Lys
 35 40 45
 atg ttt gac ggc tcc tct atc ggt ggt tgg aag ggc atc aac gaa tct 192
 Met Phe Asp Gly Ser Ser Ile Gly Gly Trp Lys Gly Ile Asn Glu Ser
 50 55 60
 gac atg gtg ctg atg ccg gac gcc agc acg gcg gtt ctg gat ccg ttc 240

Asp 65	Met	Val	Leu	Met	Pro 70	Asp	Ala	Ser	Thr	Ala 75	Val	Leu	Asp	Pro	Phe 80		
tcc	gaa	gaa	cct	acg	ctg	atc	att	cgc	tgt	gac	att	ctc	gag	ccg	ggc	288	
Ser	Glu	Glu	Pro	Thr 85	Leu	Ile	Ile	Arg	Cys 90	Asp	Ile	Leu	Glu	Pro 95	Gly		
acc	atg	caa	ggc	tac	gat	cgc	gac	ccg	cgT	tcc	atc	tcc	aaa	cgC	gCC	336	
Thr	Met	Gln	Gly	Tyr 100	Asp	Arg	Asp	Pro	Arg 105	Ser	Ile	Ser	Lys	Arg 110	Ala		
gaa	gac	ttc	ctg	cgC	tcc	tcc	ggc	atc	gcg	gac	acc	gtg	ctg	ttc	ggg	384	
Glu	Asp	Phe 115	Leu	Arg	Ser	Ser	Gly 120	Ile	Ala	Asp	Thr	Val 125	Leu	Phe	Gly		
cca	gag	cct	gag	ttc	ttc	ctg	ttc	gac	gac	atc	cgC	ttc	ggc	agc	agc	432	
Pro	Glu	Pro	Glu	Phe 130	Phe	Leu 135	Phe	Asp	Asp	Ile	Arg 140	Phe	Gly	Ser	Ser		
atc	cgC	ggt	tcc	cac	gtg	gcg	atc	gac	gat	atc	gaa	ggc	gCC	tgg	aac	480	
Ile	Arg	Gly	Ser	His 145	Val 150	Ala	Ile	Asp	Asp	Ile 155	Glu	Gly	Ala	Trp 160	Asn		
tcc	ggc	aca	aaa	tac	gac	ggc	ggc	aac	aaa	ggc	cac	cgT	ccg	gcg	gtg	528	
Ser	Gly	Thr	Lys 165	Tyr	Asp	Gly	Gly	Asn 170	Lys	Gly	His	Arg	Pro	Ala 175	Val		
aaa	ggc	ggt	tac	ttc	ccg	gTt	cca	tcg	gtc	gac	tct	tcg	cag	gat	ctg	576	
Lys	Gly	Gly	Tyr 180	Phe	Pro	Val 185	Pro	Ser	Val	Asp	Ser	Ser	Gln 190	Asp	Leu		
cgT	tcc	acc	atg	tgt	ctg	acc	atg	gaa	gag	atg	ggc	ctg	gtg	gTt	gaa	624	
Arg	Ser	Thr 195	Met	Cys	Leu	Thr 200	Met	Glu	Glu	Met	Gly	Leu 205	Val	Val	Glu		
gcg	cac	cac	acc	gag	gca	gcg	acc	gct	agc	cag	aac	gaa	gtg	gca	acc	672	
Ala	His	His	Thr 210	Glu	Ala 215	Ala	Thr	Ala	Ser	Gln 220	Asn	Glu	Val	Ala	Thr		
cgC	ttc	aac	acc	atg	acc	aag	aaa	gcc	gac	gaa	att	cag	atc	tat	aag	720	
Arg	Phe	Asn	Thr 225	Met	Thr 230	Lys	Lys	Ala	Asp 235	Glu	Ile	Gln	Ile	Tyr 240	Lys		
tac	gtg	gtg	cac	aac	gtg	gca	cac	gcc	ttc	ggt	aaa	acc	gcg	acc	ttc	768	
Tyr	Val	Val	His 245	Asn	Val	Ala	His	Ala	Phe 250	Gly	Lys	Thr	Ala	Thr 255	Phe		
atg	ccg	aag	ccc	atg	ttc	ggc	gac	aac	ggt	tcc	ggc	atg	cac	tgc	cac	816	
Met	Pro	Lys 260	Pro	Met	Phe	Gly	Asp 265	Asn	Gly	Ser	Gly	Met	His 270	Cys	His		
atg	tcg	ctg	tcc	aag	aac	ggc	acc	aac	ctg	ttc	gcc	ggc	gac	aaa	tac	864	
Met	Ser	Leu 275	Ser	Lys	Asn	Gly 280	Thr	Asn	Leu	Phe	Ala	Gly 285	Asp	Lys	Tyr		
ggc	ggc	ctg	tct	gaa	acc	gca	ctg	ttc	tac	atc	ggc	ggt	atc	aac	aag	912	
Gly	Gly	Leu 290	Ser	Glu	Thr 295	Ala	Leu	Phe	Tyr	Ile 300	Gly	Gly	Ile	Asn	Lys		
cac	gcc	aag	gcg	atc	aac	gcg	ctg	gcc	aac	ccg	acc	acc	aac	tcg	tac	960	
His	Ala	Lys 305	Ala	Ile	Asn 310	Ala	Leu	Ala	Asn 315	Pro	Thr	Thr	Asn	Ser 320	Tyr		

```

aaa cgt ctg gtg cca ggc tac gaa gcg ccg gtg atg ctg gct tac tcc 1008
Lys Arg Leu Val Pro Gly Tyr Glu Ala Pro Val Met Leu Ala Tyr Ser
                325                      330                      335

gcc cgt aac cgc tcc gcg tcc atc cgt atc ccg gtg gtc gcc agc ccg 1056
Ala Arg Asn Arg Ser Ala Ser Ile Arg Ile Pro Val Val Ala Ser Pro
                340                      345                      350

aaa gcg cgc cgc atc gaa gcc cgc ttc ccg gat ccg gcg gct aac cca 1104
Lys Ala Arg Arg Ile Glu Ala Arg Phe Pro Asp Pro Ala Ala Asn Pro
                355                      360                      365

tac ctg tgc ttc gcc gca ctg ctg atg gcc ggc ctg gac ggc atc atc 1152
Tyr Leu Cys Phe Ala Ala Leu Leu Met Ala Gly Leu Asp Gly Ile Ile
                370                      375                      380

aac aag atc cac cct ggc gac gcc atg gac aaa aac ctg ttc gac ctg 1200
Asn Lys Ile His Pro Gly Asp Ala Met Asp Lys Asn Leu Phe Asp Leu
385                390                      395                      400

ccg ccg gaa gaa gaa gcc gag atc cca aaa gtg gcc ggc tcg ctg gac 1248
Pro Pro Glu Glu Glu Ala Glu Ile Pro Lys Val Ala Gly Ser Leu Asp
                405                      410                      415

gag gcg atg gcc gcg ctg aac gaa gac cgc gag ttc ctg acc cgc ggc 1296
Glu Ala Met Ala Ala Leu Asn Glu Asp Arg Glu Phe Leu Thr Arg Gly
                420                      425                      430

ggc gtg ttc act gac gat gcg atc gat gcc tac atc gaa ctg cgc aaa 1344
Gly Val Phe Thr Asp Asp Ala Ile Asp Ala Tyr Ile Glu Leu Arg Lys
                435                      440                      445

gaa gag atg gac cgc gtt cgc atg acg cca cac ccg gtc gag ttc gaa 1392
Glu Glu Met Asp Arg Val Arg Met Thr Pro His Pro Val Glu Phe Glu
450                455                      460

ctg tac tac agc gtc taa 1410
Leu Tyr Tyr Ser Val *
465

```

<210> SEQ ID NO: 46

<211> Comprimento: 469

<212> Tipo: PRT

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: AGS1M17(AD-)

<400> Sequência: 46

```

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

```

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

<210> SEQ ID NO: 47

<211> Comprimento: 1410

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: synagslml6

<400> Sequência: 47

```

atgtcagcag agcatgtgct caccatgctg aatgagcatg aggtgaagtt cgtggacctc 60
cgcttcaccg acaccaaggg caaggagcag catgtcacca tccctgctca tcaagtcaac 120
gccgacttct ttgaagaagg caagatgttt gatggaagct caattggagg atggaagggc 180
atcaatgaga gcgacatggt gctgatgcca gatgcttcga cggcgggtgct ggacccttc 240
tcagaagaac caacattgat catcagatgt gacatcctgg agcctggcac catgcaaggc 300

```

tatgatcgag	atccaagaag	catcagcaag	cgcgccgagg	acttcttgag	gagcagcggc	360
atcgccgaca	cogtgcctt	cgggccggag	cgggagttct	tcctcttcga	cgacatcaga	420
tttgatcaa	gcatcagagg	aagtcagtgt	gccatcgacg	acattgaagg	agcatggaac	480
agcggcacca	agtacgacgg	cggcaacaag	ggccaccggc	cggcggtgaa	ggcggctac	540
ttcccgggtc	cgtcggtgga	cagcagccaa	gatttgagga	gcaccatgtg	cctcacaatg	600
gaggagatgg	ggctggtggt	ggaagctcat	cacaacgagg	tggcgacggc	atcacaaaat	660
gagggtggcaa	caaggttcaa	cacatgacc	aagaaggctg	atgagatcca	gatctacaag	720
tatgtggtgc	acaatgttgc	tcatgccttc	ggcaagacgg	ccaccttcat	gcccagcca	780
atgttcggcg	acaatggaag	cggcatgcac	tgccacatga	gcttgagcaa	gaatggcacc	840
aacctatttg	ctggagacaa	gtacggcggc	ctttctgaga	cggcgctctt	ctacatcggc	900
ggcatcaaca	agcatgccaa	ggccatcaac	gcgctggcca	accccaccac	caacagctac	960
aagaggctgg	tgcttgata	tgaggcgccg	gtgatgctgg	catattcagc	aaggaacagg	1020
agcgcctcca	tcaggattcc	tgtggtggcc	tcgcccgaag	caagaagaat	tgaagcaaga	1080
tttccagatc	ccgccgcaa	cccttatta	tgcttcgccc	cgctgctgat	ggccggcctg	1140
gatggcatca	tcaacaagat	ccatcctgga	gatgcaatgg	acaagaacct	ctacgacctg	1200
ccgccagaag	aagaagctga	gatccccaa	gtggctggat	cattggatga	agcaatggcg	1260
gcgctcaatg	aagatcgaga	gttcctcacc	cgcgccggcg	tcttctactga	tgatgccatc	1320
gacgcctaca	tcgagctgag	gaaggaggag	atggacaggg	tgaggatgac	gccgcacccg	1380
gtggagtttg	agctctacta	ctccgtgtaa				1410

REIVINDICAÇÕES

1. Polinucleotídeo isolado compreendendo um variante de SEQ ID NO: 1, caracterizado pelo fato de que dito
5 polinucleotídeo variante é pelo menos 80% idêntico a SEQ ID NO: 1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida.

2. Polinucleotídeo isolado de acordo com a reivindicação
10 1, caracterizado pelo fato de que dito polipeptídeo compreende pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 a 250 correspondentes a SEQ ID NO:2.

- 15 3. Polinucleotídeo isolado de acordo com a reivindicação 2, caracterizado pelo fato de que é selecionado do grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 47, ou a seqüência de nucleotídeos de resistência a herbicidas do
20 inserto de DNA do plasmídeo depositado como No. de Acesso NRRL B-30930.

4. Polinucleotídeo isolado de acordo com a reivindicação
1, caracterizado pelo fato de que dito polinucleotídeo
25 codifica um polipeptídeo tendo pelo menos uma modificação

que resulta na perda de um sítio de adenilação em dito polipeptídeo.

5. Polinucleotídeo isolado de acordo com a reivindicação
5 4, caracterizado pelo fato de que é selecionado do grupo
consistindo de SEQ ID NO:43 e 45.

6. Polinucleotídeo isolado de acordo com a reivindicação
1, caracterizado pelo fato de que dito inibidor de
10 glutamina sintetase herbicida compreende glufosinato.

7. Polinucleotídeo isolado de acordo com a reivindicação
1, caracterizado pelo fato de que dito polinucleotídeo é
uma seqüência sintética que foi projetada para expressão em
15 uma planta.

8. Polinucleotídeo isolado caracterizado por ser de SEQ
ID NO: 1.

20 9. Vetor caracterizado por compreender o polinucleotídeo
de acordo com a reivindicação 1.

10. Vetor de acordo com a reivindicação 9, caracterizado por compreender, ainda, um polinucleotídeo codificando um polipeptídeo heterólogo.

5 11. Célula hospedeira caracterizada por conter o vetor de acordo com a reivindicação 9.

12. Célula hospedeira de acordo com a reivindicação 11, caracterizada por ser uma célula hospedeira bacteriana.

10

13. Célula hospedeira de acordo com a reivindicação 11, caracterizada por ser uma célula vegetal.

15 14. Planta transgênica caracterizada por compreender a célula hospedeira de acordo com a reivindicação 9.

20 15. Planta de acordo com a reivindicação 14, caracterizada por compreender, ainda, um ou mais de um polinucleotídeo codificando um gene de tolerância a herbicidas ou um polinucleotídeo codificando um gene de tolerância a insetos.

16. Planta de acordo com a reivindicação 15, caracterizada pelo fato de que dito gene de tolerância a herbicidas é um gene de EPSP sintase.

5 17. Planta de acordo com a reivindicação 15, caracterizada pelo fato de que dito gene de tolerância a insetos é uma endotoxina.

10 18. Planta de acordo com a reivindicação 14, caracterizada pelo fato de que dita planta é selecionada do grupo consistindo em milho, sorgo, trigo, girassol, tomate, crucíferas, pimentas, batata, algodão, arroz, soja, beterraba, cana-de-açúcar, tabaco, cevada e colza.

15 19. Semente transformada caracterizada por compreender o polinucleotídeo de acordo com a reivindicação 1.

20 20. Polipeptídeo isolado compreendendo um variante de SEQ ID NO:2, caracterizado pelo fato de que dito polipeptídeo variante é pelo menos 80% idêntico a SEQ ID NO:2 e é resistente a inibição por inibidor de glutamina sintetase herbicida.

21. Polipeptídeo isolado de acordo com a reivindicação 20, caracterizado pelo fato de que dito polipeptídeo compreende pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 a 250 correspondentes a SEQ ID NO:2.

22. Polipeptídeo isolado de acordo com a reivindicação 19, caracterizado por ser selecionado do grupo consistindo em SEQ ID NOS: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, ou um polipeptídeo que é codificado pela seqüência de nucleotídeos de resistência a herbicidas do inserto de DNA do plasmídeo depositado como No. de Acesso NRRL B- 30930.

23. Polipeptídeo isolado de acordo com a reivindicação 20, caracterizado pelo fato de que dito polipeptídeo compreende pelo menos uma modificação que resulta na perda de um sítio de adenilação em dito polipeptídeo.

24. Polipeptídeo isolado de acordo com a reivindicação 23, caracterizado por ser selecionado do grupo consistindo em SEQ ID NO:44 e 46.

25. Polipeptídeo isolado de acordo com a reivindicação 22, caracterizado pelo fato de que dito inibidor de glutamina sintetase herbicida compreende glufosinato.

5 26. Polipeptídeo de acordo com a reivindicação 23, caracterizado por compreender, ainda, uma seqüência de aminoácidos heteróloga.

10 27. Polipeptídeo isolado caracterizado por ser de SEQ ID NO: 2.

15 28. Método para conferir resistência a inibidores herbicidas de glutamina sintetase em uma planta caracterizado por compreender transformar dita planta com um polinucleotídeo compreendendo um variante de SEQ ID NO: 1, em que dito polinucleotídeo variante é pelo menos 80% idêntico a SEQ ID NO:1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida, e regenerar uma planta transformada.

20

25 29. Método de acordo com a reivindicação 28, caracterizado pelo fato de que dito polipeptídeo compreende pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 to 250 correspondentes a SEQ ID NO:2.

30. Método de acordo com a reivindicação 29, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15,
5 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, ou 47.

31. Método de acordo com a reivindicação 28, caracterizado pelo fato de que dito polinucleotídeo codifica um polipeptídeo tendo pelo menos uma modificação que resulta
10 na perda de um sítio de adenilação em dito polipeptídeo.

32. Método de acordo com a reivindicação 31, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NO:43 e 45.

15

33. Método de acordo com a reivindicação 28, caracterizado pelo fato de que dito inibidor de glutamina sintetase herbicida compreende glufosinato.

20 34. Método de acordo com a reivindicação 28, caracterizado pelo fato de que dito método compreende ainda a transformação de dita planta com um ou mais de um polinucleotídeo codificando um gene de tolerância a herbicidas ou um polinucleotídeo codificando um gene de
25 tolerância a insetos.

35. Método de acordo com a reivindicação 34, caracterizado pelo fato de que dito gene de tolerância a herbicidas é um gene de EPSP sintase.

5 36. Método de acordo com a reivindicação 34, caracterizado pelo fato de que dito gene de tolerância a insetos é uma endotoxina.

10 37. Planta caracterizada por ter estavelmente incorporada em seu genoma uma construção de DNA compreendendo um polinucleotídeo compreendendo um variante de SEQ ID NO: 1, em que dito polinucleotídeo variante é pelo menos 80% idêntico a SEQ ID NO:1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase
15 herbicida.

20 38. Planta de acordo com a reivindicação 37, caracterizada pelo fato de que dito polipeptídeo compreende pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 a 250 correspondentes a SEQ ID NO: 2.

39. Planta de acordo com a reivindicação 38, caracterizada pelo fato de que dito polinucleotídeo é selecionado do

grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, ou 47.

40. Planta de acordo com a reivindicação 37, caracterizada
5 pelo fato de que dito polinucleotídeo codifica um polipeptídeo tendo pelo menos uma modificação que resulta na perda de um sítio de adenilação em dito polipeptídeo.

41. Planta de acordo com a reivindicação 40, caracterizada
10 pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo em SEQ ID NO:43 e 45.

42. Planta de acordo com a reivindicação 37, caracterizada
15 pelo fato de que dito inibidor de glutamina sintetase herbicida compreende glufosinato.

43. Planta de acordo com a reivindicação 37, caracterizada pelo fato de que dita planta é uma célula vegetal.

20 44. Planta de acordo com a reivindicação 37, caracterizada pelo fato de compreender, ainda, um ou mais de um polinucleotídeo codificando um gene de tolerância a herbicidas ou um polinucleotídeo codificando um gene de tolerância a insetos.

45. Planta de acordo com a reivindicação 44, caracterizada pelo fato de que dito gene de tolerância a herbicidas é um gene de EPSP sintase.

5 46. Planta de acordo com a reivindicação 44, caracterizada pelo fato de que dito gene de tolerância a insetos é uma endotoxina.

10 47. Planta de acordo com a reivindicação 37, caracterizada pelo fato de que dita planta é selecionada do grupo consistindo de milho, sorgo, trigo, girassol, tomate, crucíferas, pimentas batata, algodão, arroz, soja, beterraba, cana-de-açúcar, tabaco, cevada e colza.

15 48. Método para controlar seletivamente ervas daninhas em um campo contendo uma planta tendo sementes plantadas ou plantas, caracterizado por compreender as etapas de: a) plantar as sementes ou plantas que são resistentes a inibidor de glutamina sintetase herbicida como resultado de
20 um polinucleotídeo compreendendo um variante de SEQ ID NO: 1 sendo inserido na semente ou planta, caracterizado pelo fato de que dito polinucleotídeo variante é pelo menos 80% idêntico a SEQ ID NO:1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase
25 herbicida; e, b) aplicar às plantas e ervas daninhas em um campo, uma concentração efetiva de um inibidor de glutamina

sintetase herbicida para controlar ervas daninhas sem afetar significativamente as plantas.

5 49. Método de acordo com a reivindicação 48, caracterizado pelo fato de que dito polipeptídeo compreende pelo menos uma modificação entre os aminoácidos 125 a 175 ou pelo menos uma modificação entre os aminoácidos 200 a 250 correspondentes a SEQ ID NO: 2.

10 50. Método de acordo com a reivindicação 49, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, ou 47.

15 51. Método de acordo com a reivindicação 48, caracterizado pelo fato de que dito polinucleotídeo codifica um polipeptídeo tendo pelo menos uma modificação que resulta na perda de um sítio de adenilação em dito polipeptídeo.

20 52. Método de acordo com a reivindicação 51, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NO:43 e 45.

53. Método de acordo com a reivindicação 50, caracterizado pelo fato de que dito inibidor de glutamina sintetase herbicida compreende glufosinato.

5 54. Método de acordo com a reivindicação 48, caracterizado pelo fato de que dita planta ou semente compreende ainda um ou mais de um polinucleotídeo codificando um gene de tolerância a herbicidas ou uma polinucleotídeo codificando um gene de tolerância a insetos.

10

55. Método de acordo com a reivindicação 54, caracterizado pelo fato de que dito gene de tolerância a herbicidas é um gene de EPSP sintase.

15 56. Método de acordo com a reivindicação 54, caracterizado pelo fato de que dito gene de tolerância a insetos é uma endotoxina.

20 57. Método para melhorar a produção em uma planta caracterizado pelo fato de que compreende introduzir, em dita planta, um polinucleotídeo codificando uma enzima glutamina sintetase derivada de uma bactéria.

58. Método de acordo com a reivindicação 57, caracterizado pelo fato de que dito polinucleotídeo compreende um variante de SEQ ID NO:1, caracterizado pelo fato de que dito polinucleotídeo variante é pelo menos 80% idêntico a SEQ ID NO: 1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida.

59. Método de acordo com a reivindicação 58, caracterizado pelo fato de que dito polinucleotídeo compreende um polinucleotídeo selecionado do grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 47, ou a seqüência de nucleotídeos de resistência a herbicidas do inserto de DNA do plasmídeo depositado como No. de Acesso NRRL B-30930.

15

60. Método de acordo com a reivindicação 57, caracterizado pelo fato de que dito polinucleotídeo codifica a polipeptídeo tendo pelo menos uma modificação que resulta na perda de um sítio de adenilação in dito polipeptídeo.

20

61. Método de acordo com a reivindicação 60, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NO:43 e 45.

62. Método para melhorar a utilização de nitrogênio em uma planta caracterizado pelo fato de que compreende introduzir, em dita planta, um polinucleotídeo codificando uma enzima glutamina sintetase derivada de uma bactéria.

5

63. Método de acordo com a reivindicação 62, caracterizado pelo fato de que dito polinucleotídeo compreende um variante de SEQ ID NO: 1, caracterizado pelo fato de que dito polinucleotídeo variante é pelo menos 80% idêntico a
10 SEQ ID NO: 1 e codifica um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida.

64. Método de acordo com a reivindicação 63, caracterizado pelo fato de que dito polinucleotídeo compreende um
15 polinucleotídeo selecionado do grupo consistindo de SEQ ID NOS: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, ou a seqüência de nucleotídeos de resistência a herbicidas do inserto de DNA do plasmídeo depositado como No. de Acesso NRRL B-30930.

20

65. Método de acordo com a reivindicação 62, caracterizado pelo fato de que dito polinucleotídeo codifica um polipeptídeo tendo pelo menos uma modificação que resulta na perda de um sítio de adenilação in dito polipeptídeo.

66. Método de acordo com a reivindicação 65, caracterizado pelo fato de que dito polinucleotídeo é selecionado do grupo consistindo de SEQ ID NO:43 e 45.

FIGURA 1C

```

AGS-1      FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3421m1  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3422m2  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3427m3  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3428m4  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3430m6  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3431m7  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3432m8  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3433m9  FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3434m10 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3435m11 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3436m12 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3437m13 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3438m14 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMV 420
pAX3426m15 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
pAX3439m16 FPDPAANFYLCFAALLMAGLDGIINKIHFGDAMDKNLYDLPPPEEEAEIPKVAGSLDEAMA 420
*****;*****.

```

```

AGS-1      ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3421m1  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3422m2  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3427m3  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3428m4  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3430m6  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3431m7  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3432m8  ALNEDREFLTRGGVFTDDAIDAYIELCKEEMDRVRMTPHPVEFELYYSV 469
pAX3433m9  ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3434m10 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3435m11 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3436m12 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3437m13 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3438m14 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3426m15 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
pAX3439m16 ALNEDREFLTRGGVFTDDAIDAYIELRKEEMDRVRMTPHPVEFELYYSV 469
***** *****

```

A/107/1953-4

RESUMO

POLINUCLEOTÍDEO ISOLADO E POLIPEPTÍDEO CONFERINDO RESISTÊNCIA A GLUTAMINA SINTETASE E MÉTODOS PARA PRODUZIR PLANTAS E CÉLULAS VEGETAIS TRANSGÊNICAS APRESENTANDO PRODUÇÃO E UTILIZAÇÃO DE NITROGÊNIO MELHORADAS

São providos composições e métodos para conferir resistência a herbicidas e melhorar a utilização de nitrogênio de bactérias, plantas, células vegetais, tecidos e sementes. Composições compreendendo uma seqüência codificadora para um polipeptídeo que confere resistência ou tolerância a inibidores de glutamina sintetase herbicidas são providas. As seqüências codificadoras podem ser usadas em construções de DNA ou cassetes de expressão para transformação e expressão em plantas. Composições também compreendem bactérias, plantas, células, tecidos e sementes vegetais transformadas. Em particular, polinucleotídeos isolados correspondendo a polinucleotídeos resistentes a inibidor de glutamina sintetase herbicida são providos. Adicionalmente, polipeptídeos correspondendo aos polinucleotídeos são englobados. Em particular, a presente invenção provê polinucleotídeos isolados compreendendo um variante de SEQ ID NO: 1, caracterizado pelo fato do polinucleotídeo variante codificar um polipeptídeo que é resistente a inibição por inibidor de glutamina sintetase herbicida.